Essays on Educational Inequality and Policy

Guilherme Strifezzi Leal

PhD in Economics

Essays on Educational Inequality and Policy

Guilherme Strifezzi Leal

PhD in Economics | Guilherme Strifezzi Leal

UNIVERSITAT DE
BARCELONA

PhD in Economics

Thesis title:

Essays on Educational Inequality and Policy

PhD candidate:

Guilherme Strifezzi Leal

Advisor:

Álvaro Choi de Mendizábal

Date:
May 2022

UNIVERSITAT DE
BARCELONA

Acknowledgements

First and foremost, I would like to express my deep gratitude to my advisor Álvaro Choi. Álvaro is, without a doubt, one of the most supportive, dedicated and kind persons I have met throughout my academic life. This research would certainly not have been possible without him. I am also sincerely thankful to the University of Barcelona Department of Economics, for providing an inspiring environment in which to conduct my thesis, and to all the professors and researchers who have helped me with valuable feedback on my research.

Secondly, I would like to thank all the colleagues I met along this journey, from those in my undergraduate courses in Brazil to the friends I made in Barcelona. All of this would have been much less meaningful without you.

Finally, I will be eternally grateful to my family and childhood friends from Brazil. Pursuing a PhD has been a dream of mine for a long time, as I always wanted to prepare myself to the fullest in order to contribute to the construction of a better world. There have been many difficult times, and my friends and family have always been my backbone, giving me the support and the strength that I needed to always keep pushing. I love you all.

Contents

List of Tables

List of Figures

1. Introduction

An efficient educational system not only creates conditions to promote sustainable economic development, but also helps to mitigate social inequalities through many different channels. First, education may help to reduce the income gap between the rich and the poor, once cash benefits are proportionally greater for low-income families. Second, it is also an important poverty reduction tool – according to UNESCO's Global Education Monitoring Report, it is estimated that extreme poverty might be halved if universal access to primary and secondary education were to be achieved. Finally, education could also be an engine of equality in other – less trivial - ways, for instance, by helping to reduce gender/racial barriers and to generate a greater political and civil engagement. (Walker et al. 2019).

Public policy in education might me targeted to different levels of the educational system and might be considerably different in nature. Some, as teacher training, changes in the educational curriculum and investments in school technology, could be designed to enhance the quality of educational provision, while others, such as student quotas and cash transfer programs targeted to vulnerable groups, aim to reduce educational inequalities. A specific matter that has been on the center of educators and policy makers debate in the past decades is the inequality of opportunity in access to higher education (HE). This is a problem that is present both in developed and emerging economies (Lynch and O'riordan 1998; Metcalf 2003; Kelly 2005; Alon 2009; Chesters and Watson 2013; Menezes Filho and Kirschbaum 2019) and that reverberates across a wide range of social dimensions, such as the labor market, crime, and the civic engagement of the population (Egerton 2003; Egwakhe and Osabuohlen 2009; Oreopoulos and Petronijevic 2011; Menezes Filho and Kirschbaum 2019). As a result, many different strategies have been implemented throughout recent years in an attempt to facilitate access to HE for underprivileged groups, and examples of public policies addressing this matter include college scholarships, student financing schemes, quotas for specific groups in higher education institutions, and measures to combat educational gaps in primary and secondary school levels.

In this dissertation, I study a set of policies that aim to mitigate the inequality of access in HE and their effects on educational attainment, welfare, and the

labor market. This inequality may arise due to different reasons, such as socioeconomic factors, gender, geography, ethnicity or race. The case of Brazil, a country with a particularly unequal educational system and in which a wide range of educational policies have been implemented since the 1990s, is used as a laboratory for this analysis.

This dissertation is composed of three independent but related research articles, each one studying a specific Brazilian policy/program targeted to enhance access of vulnerable groups to higher education. While the articles investigate considerably different policies, they all essentially address the same issue, which is how the public sector can efficiently use its tools to reduce the barriers to higher education, and if the evaluated – or similar – policies may or not serve as a guide to policy makers worldwide who wish to increase the efficiency and equality of opportunity of their educational systems.

In the first article (chapter 2 of this dissertation), 'Racial Quotas in Higher Education and Pre-College Academic Performance', I analyze the effects of a law that implemented racial quotas in Brazilian federal universities on the pre-college academic performance of non-white individuals (i.e., those eligible for the quotas). The objective of this chapter is to provide a better understanding of the incentives provided by affirmative action in education – a matter that is ambiguous from a theoretical perspective and lacks empirical evidence. To this end, I exploit the introduction of the Law of Quotas in 2013 by employing a difference-in-differences design to assess the impacts of the law on the scores of non-white individuals on the ENEM (Brazil's college-entrance exam). This research provides evidence that the Law of Quotas fostered incentives to pre-college human capital accumulation as it induced non-white students to attain higher scores on the ENEM. Therefore, this investigation indicates that the implementation of racial quotas in higher education not only promotes equity, but also brings about efficiency gains, as it encourages non-white students to close the performance gap with white students by the end of secondary education. Additional findings from the research include: i) the effects were stronger on quantitative-intensive subjects (Math and Natural Sciences) than on the remaining fields of the exam (Language, Social Sciences and the Essay); and ii) the Law of Quotas did not exert statistically significant differences by gender nor level of parental education.

In the second article (chapter 3 of this dissertation), 'Income-Based Scholarships and Access to Higher Education', I investigate the effects of a Brazilian federal program (the Prouni) that grants full and partial college scholarships to students from low-income families on access to HE. The objective of this chapter is to examine the effectiveness of financial aid to college students on promoting human capital accumulation – which is still a subject of discussion in the literature, especially as regards to emerging economies. To this end, I exploit the introduction of the Prouni in 2005 by employing a difference-in-differences methodology (similar to the one implemented in Chapter 2 – although in this chapter the results are also corroborated with an instrumental variables and a regression discontinuity approach), such as to evaluate the effects of the program on the higher education enrolment rates of low-income students. This research provides evidence that, by the third year after its implementation, the Prouni had increased the odds of attending higher education by 37% for individuals entitled to the full scholarship and by 20% for those entitled to the partial scholarship. Moreover, it is estimated that every US$ 100 million spent by the government with tax waivers form the Prouni (or, equivalently, US$1,000 per student) generated an approximate 0.8 percentage points increase in the HE enrolment rate of academic age individuals. The impact of the Prouni in terms of percentage increase in the HE enrolment rate is similar to those found in developed economies (namely in the US and the UK). The findings form this chapter also suggest that the impacts of the grants on access to higher education were greater for women and for non-white persons.

In the third article (chapter 4 of this dissertation), 'Welfare and Labor Supply Effects of Student Loan Schemes', I study the welfare and labor supply effects of different student loan schemes in higher education, by developing a partial equilibrium microsimulation model in which graduates maximize their expected utilities under wage uncertainty, risk aversion and elastic labor supply. The model predicts that shifting from a mortgage loan to an income-contingent loan (ICL) scheme shall (i) decrease labor supply; (ii) increase graduates' expected welfare; (iii) reduce repayment burdens, and (iv) increase the number of years until the debt is fully repaid. The model is then calibrated with real Brazilian data and the results confirm the predictions when changing the Brazilian government-backed student loan program to an ICL. I find that shifting to an ICL is especially welfare-enhancing for women and non-white people, two population groups who have lower initial

earnings, flatter income growth curves throughout their working lifetimes and who also face greater unemployment risks.

Finally, in the fifth chapter some concluding remarks are presented.

2. Racial Quotas in Higher Education and Pre-college Academic Performance

2.1. Introduction

Racial inequalities in education have equity and efficiency implications. While they exacerbate social inequalities and hinder intergenerational mobility, they also constitute a waste of human capital potential. Although racial inequalities in education have narrowed during the past decades, educational gaps between students from different racial and ethnic backgrounds are still wide in several countries (Marteleto 2012; O'Gorman 2010). Therefore, in order to close gaps in access and outcomes, affirmative actions with different designs have been implemented throughout recent years and have included measures such as preferential treatment in admission processes, race-specific financial aid and scholarship policies (Arcidiacono 2005; Ibarra 2001).

An alternative and sometimes complementary approach has been the introduction of racial and ethnic quotas. At the higher education level, this measure consists of pre-establishing a share of seats in institutions to specific racial and/or ethnic groups and has been applied in countries such as Brazil, India and Malaysia. Although there is substantial literature on the ex-post effects of educational quotas, that is, the effects on quota holders after college admissions outcomes are determined, the ex-ante effects of such policies have been significantly less explored, and a particularly relevant issue concerns the incentives that these quotas provide to pre-college human capital accumulation – indeed, the few empirical research works that have investigated these effects reached diverging results (for instance, Saeme 2014 and Assunção and Ferman 2015).

Moreover, there is a well-established literature on the importance of students' pre-college accumulated human capital in explaining the variation in college graduates' earnings (Walker and Zhu 2018; Dale and Krueger 2002 and 2014). Therefore, understanding the ex-ante effects of racial quotas is crucial not only to unravel the incentives provided by affirmative action but also because they play a key role in labor market success.

In this chapter, I evaluate the ex-ante effects of a law that implemented quotas in higher education in Brazil. Differences in access to higher education by race are significant in this country: in the year 2010, according to Censo IBGE (*Instituto Brasileiro de Geografia e Estatítica*) and the Higher Education Census, black and brown persons represented 51% of the total population, but accounted for only 34% of higher education enrolments, whereas white persons made up 48% of the population and 63% of enrolments. In an attempt to mitigate this inequality of access to higher education, the federal government created in 2013 what came to be known as the Law of Quotas, establishing that a proportion of seats in Brazilian federal universities should be filled by non-white and low-income students from public high schools. In this research, I study the effects of the law on students' pre-college academic performance; that is, whether the increase in enrolment by these students was due only to the existence of an increased number of reserved seats or whether the policy itself had a positive incentive effect on human capital accumulation (i.e., if there were efficiency gains). More specifically, this research focuses on evaluating the efficiency effects of the racial criteria of the law.

These effects are assessed by examining the extent to which the Law of Quotas affected the performance of students in the college entrance exam, the ENEM[1]. To this end, I employ a difference-in-differences approach by explicitly controlling for a set of student-specific variables contained in the ENEM's microdata. The repeated cross-sectional database provides information on socioeconomic factors, income, parental education, previous work experience and previous academic effort. To the best of my knowledge, this is the first research to take advantage of a major country-level quotas law to investigate such effects on a national scale. This research shows that the law fostered incentives to pre-college human capital accumulation as it induced eligible students to attain higher scores on the ENEM exam. Furthermore, I test for the existence of heterogeneous effects by subject, gender and parental education, and estimate both a two-periods difference-in-differences model and a regression with dynamic treatment effects.

This research shows that the effects of the law were greater in more quantitative-intensive subjects (Math and Natural Science) and that the

---

[1] Exame Nacional do Ensino Médio.

impact of the law increased throughout the first years after its implementation. I do not find any evidence, however, that the effects of the law were distinct between genders and between students with and without a college-educated parent.

The chapter is organized as follows. Section 2.2 provides a brief literature review on the effects of racial quotas. Section 2.3 expands on the institutional setting of the Brazilian educational system and of the Law of Quotas. Section 2.4 describes the data and the empirical strategy employed in the research. Section 2.5 presents the results of the difference-in-differences models. Section 2.6 presents a battery of robustness exercises. Section 2.7 discusses the main implications of the findings. Finally, section 2.8 concludes the chapter.

2.2. Educational quotas in higher education

The introduction of racial and ethnic quotas in college admissions has become a common practice in a number of countries. In the United States, even though such policies have not been prescribed by the Constitution, several guidelines issued by the U.S. education and justice departments have encouraged institutions to grant preferential treatment to applicants from minority groups in admissions to universities (Department of Education and Department of Justice, 2011). Moreover, in countries such as India, Malaysia and Brazil, a step further has been taken as these practices have been institutionalized by federal laws establishing that a certain percentage of seats in education institutions should be filled by specific racial or ethnic groups.

Economists have long been interested in understanding how these policies affect the college enrolment of students who benefit from the quotas (quota holders) as well as their effects on performance in higher education. Interestingly, an additional issue that has been less explored by the literature regards the incentives to human capital accumulation that these policies yield to targeted students prior to college admissions.

The incentive effects of affirmative action on pre-college human capital accumulation have been studied mostly from a theoretical perspective, and conclusions are ambiguous. On the one hand, such policies might lead to ex-post discrimination of minority groups, as argued by Coate and Loury (1993), Loury (1992), Milgron and Oster (1987) and Lundberg and Startz (1983), or even to more complacent students due to the high numbers of reservation

quotas, especially among the smartest section of the minority group, as stressed by Knight and Hebl (2005) and Assunção and Ferman (2015), which could encourage quota-eligible students to reduce skill acquisition during basic school. On the other hand, affirmative policies might mitigate the so-called "discouragement effects", dislocating the students to the margin of selection and increasing the willingness to re-allocate leisure time towards building human capital as pointed out by Cotton et al. (2016) and Furstenberg (2003).

Empirical investigations on the incentive effects of educational quotas, however, are still scarce and provide mixed results. Khanna (2020) evaluates the effects of reservation quotas for college seats and government jobs in pre-college years of schooling in India and finds that affirmative action incentivizes about 0.8 additional years of education for the average minority group student and 1.2 more years of education for a student from a marginal minority subgroup. In the U.S., Antonovics and Backes (2014) conclude that banning affirmative action policies in Californian public universities had no effect on pre-college academic performance (as measured by the students' SAT score and GPA), while Akhtari, Bau and Laliberté (2020) find that the reinstallation of affirmative action in three U.S. States reduced racial gaps across the pre-treatment test scores distribution

In the Brazilian context, the empirical literature has been mostly restricted to specific universities that have implemented racial quotas in their admission processes of their own will prior to the 2012 Law of Quotas, and results are again mixed. Saeme (2014) investigates the implementation of a 40% quota for black persons in the Federal University of São Carlos (UFScar) and finds that black students from public schools in São Paulo scored 1.54% higher on the ENEM as a result of the introduction of quotas in UFSCar admissions. Francis and Tanuri-Pianto (2012) evaluate the adoption of racial quotas at the University of Brasilia in 2004 and find that the quotas did not reduce pre-college effort (it might have even raised pre-college effort, although the evidence is tenuous), while Estevan, Gall and Morin (2017) investigate the effects of a program that awarded bonus points in the admission exam of the University of Campinas to targeted students and find no evidence of behavioral reactions regarding exam-preparation effort. Conversely, Assunção and Ferman (2015) evaluate the effects of the implementation of quotas in three public Universities in the States of Rio de Janeiro and Bahia

22

from 2002 to 2004 and find that these quotas induced targeted groups to attain lower high school scores.

The introduction of the Law of Quotas in 2012 in Brazil, which ensured the implementation of racial quotas in all of the federal universities in the country, created an advantageous setup to expand the understanding of the ex-ante effects of these quotas and finally provide more clarity in the direction of these incentives. Most of the research that has been undertaken to evaluate the impact of the law, however, has focused on its effects on college enrolment and ex-post college performance (Vidigal 2018; Queiroz et al. 2015; Ribas et al. 2015).

Indeed, the effects of the law on students' pre-college behavior have been largely ignored by the literature. While Mello (2019) investigated how the Law of Quotas impacted the ex-ante decision between attending a private or public high school, the effects of the law on pre-college academic performance are, to the best of my knowledge, yet to be examined. Therefore, this research contributes to the literature by presenting causal evidence of the impact of an affirmative policy on pre-college performance at the national level.

2.3. Social and Institutional Background

This section describes the social and institutional background relevant to this chapter. Subsection 2.3.1 describes some key demographic characteristics of the Brazilian population; subsection 2.3.2 outlines the structure of the Brazilian higher education system; subsection 2.3.3 provides information on the ENEM exam; and finally, subsection 2.3.4 describes the 2012 Law of Quotas.

2.3.1. Sociodemographic characteristics

The Brazilian population's racial composition stems from a confluence of many different ethnic backgrounds, from indigenous people, black Africans and Portuguese that represented the majority of Brazil's inhabitants in the colonization period to the subsequent waves of Europeans, Arabs and Asians that arrived in the country throughout the 20th century. Consequently, most of the Brazilian population possesses some degree of mixed-race ancestry, which has led researchers that investigate the racial dynamics in the country to focus on the so-called black-to-white continuum.

The black-to-white continuum encompasses 99% of the Brazilian population and the national institute responsible for collecting and reporting sociodemographic data (the IBGE, *Instituto Brasileiro de Geografia e Estatística*) uses three different racial terms to identify individuals among this continuum: white (*branco*, which represents 48% of the country's population), brown (*pardo*, 43% of the population) and black (*preto*, 8% of the population). The remaining 1% of the population is composed mainly of Asians and indigenous ethnicities. Table 2.1 - Racial Composition by Brazilian State shows that the racial composition varies widely across Brazilian States.

Table 2.1 - Racial Composition by Brazilian State

Region	State	White	Black	Asian	Brown	Indigenous
Brazil	**Brazil**	**48%**	**8%**	**1%**	**43%**	**0.4%**
North	Rondônia	35%	7%	1%	56%	1%
	Acre	24%	6%	2%	66%	2%
	Amazonas	21%	4%	1%	69%	5%
	Roraima	21%	6%	1%	61%	11%
	Pará	22%	7%	1%	70%	1%
	Amapá	24%	9%	1%	65%	1%
	Tocantins	25%	9%	2%	63%	1%
Northeast	Maranhão	22%	10%	1%	67%	1%
	Piauí	24%	9%	2%	64%	0%
	Ceará	32%	5%	1%	62%	0%
	Rio Grande do Norte	41%	5%	1%	52%	0%
	Paraíba	40%	6%	1%	53%	1%
	Pernambuco	37%	6%	1%	55%	1%
	Alagoas	32%	7%	1%	60%	0%
	Sergipe	28%	9%	1%	61%	0%
	Bahia	22%	17%	1%	59%	0%
Southeast	Minas Gerais	45%	9%	1%	44%	0%
	Espírito Santo	42%	8%	1%	49%	0%
	Rio de Janeiro	47%	12%	1%	39%	0%
	São Paulo	64%	6%	1%	29%	0%
South	Paraná	70%	3%	1%	25%	0%
	Santa Catarina	84%	3%	0%	12%	0%
	Rio Grande do Sul	83%	6%	0%	11%	0%
Central West	Mato Grosso do Sul	47%	5%	1%	44%	3%
	Mato Grosso	37%	8%	1%	52%	1%
	Goiás	42%	7%	2%	50%	0%
	Distrito Federal	42%	8%	2%	48%	0%

Source: 2010 Census - IBGE

Among the three largest racial groups in the country (white, black and brown persons), a common concern is the substantial educational disparity between them, especially in terms of access to higher education. According to the IBGE[2], in 2010, 13% of the white population had a college degree, whereas the same was true for only 4% of the black and brown population. By 2019, these figures had evolved to 21% for white persons and 9% for black and

---

[2] Census for 2010 data and PNAD (*Pesquisa Nacional por Amostra de Domicílio*) for 2019 data

brown persons. Moreover, the uneven playing field in the educational sphere also contributes to the perpetuation of inequality in income levels. In 2010, the average income for the white population was 1.9 times larger than it was for the black and brown population, while in 2019 it was 1.8 times larger.

2.3.2. Higher Education in Brazil

According to the 2019 Higher Education Census, the Brazilian Higher Education system serves 8.6 million students (in 2019, the average enrolment rate of individuals between 18 and 24 years old was 20.4%) and consists of 2,608 institutions, among which 2,306 (or 88%) are private and 302 (or 12%) are public. Private institutions, which are fee-paying, contain the vast majority of enrollments (6.5 million students in 2019, or nearly 76% of total enrollments). Each private institution has complete independence regarding tuition fees and runs its own admission process, which usually consists of written exams developed by the institution itself. Public institutions, in turn, are predominantly free of charge[3] and are managed by either the federal, state or municipal government. Federal (110) and State (132) HEIs (Higher Education Institutions) encompass most of the public enrollments (62% and 32%, respectively), while Municipal institutions (60) account for only 6% of public enrollments.

Public HEIs are generally more prestigious and since they are mostly tuition-free, these institutions have the most competitive selection processes in the country[4]. Until 2010, the admission process to public HEIs was highly decentralized and most institutions developed their own exam –indeed, some of them used the ENEM as part of the selection criteria. This structure led to tests with widely different contents and to a highly localized higher education

---

[3] Institutions maintained by federal and state levels of governments are forbidden by law to charge tuition fees, but municipal institutions are allowed and usually charge some tuition fees.

[4] Federal and State universities have higher average scores in the *Índice Geral de* Cursos (IGC), a quality index developed by the Ministry of Education, and comprise most of the higher ranked institutions in the *Ranking Universitário Folha* (RUF), an annual evaluation of the HEIs in Brazil developed by the *Folha de São Paulo* newspaper. According to Binelli et al. (2008), there were on average 9 applicants per seat at public institutions in 2003, while this ratio was 1.5 in private institutions.

market, since it induced students to restrict their study and preparation to admission processes for specific universities. On that count, in 2010 the Ministry of Education created the *Sistema de Seleção Unificada* (SISU), an online platform where Federal and State universities could use the grades of the students in the national standardized exam (the ENEM) for their admission processes. In order to be eligible for admission, students who take the ENEM exam must then complete a SISU application. By 2015, the system was being used by 108 public institutions, among which 92 were federal HEIs. larger.

2.3.3. The *Exame Nacional do Ensino Médio* (ENEM)

The ENEM is a national non-mandatory standardized exam organized by the INEP (*Instituto Nacional de Estudo e Pesquisas Educaionais Anísio Teixeira*) within the Ministry of Education that takes place once a year in Brazil, and it is one of the largest national exams in the world with a yearly average, between 2010 and 2019, of 6.5 million test takers. Created in 1998 with the purpose of evaluating high school students' performance and learning, it now plays a multiple role: it is a mandatory exam for the SISU application (and therefore serves as an entrance exam to many HEIs in the country); one of the selection criteria in the Prouni (*Universidade para Todos*), a federal scholarship program established in 2005; and it is also used to evaluate and compare the quality of high school institutions in the country.

The ENEM is comprised of one multiple choice exam and one essay. The multiple choice (or objective) exam consists of four different subjects: natural sciences, social sciences, languages and math. In the essay (or written exam), candidates must discourse upon a topic of public interest (usually about Brazilian social, political and/or economic issues). A detailed description of the ENEM is provided in the ENEM's Act (*Edital do ENEM*) and Syllabus (*Matriz de Referência ENEM*).

2.3.4. The 2012 Law of Quotas

Access to higher education is significantly unequal in Brazil and has historically lacked representation of non-white students. Moreover, the inequality of access between students who attended private high schools and those who attended a public institution has also been an ongoing concern. According to IBGE's *Síntese de Indicadores Sociais*, in 2017, 79% of private

high school graduates progressed to tertiary education, while the same was true for only 36% of graduates from public schools.

In view of that picture, a handful of Brazilian public universities began to implement racial quotas in their admission processes in the early 2000s. Finally, in August 2012, the Brazilian federal government established the Law 12.711/2012, which later came to be known as the Law of Quotas. The law stated that at least 50% of places in Federal HEIs should be filled by students that had attended the entire high school period (in Brazil, this consists of 3 years) in a public institution. Among this group, at least 50% (that is, 25% of the total) should be filled by students from public schools whose per capita family income amounts to at most one and a half times the minimum wage (approximately US$300 per month in 2020), and at least X% (that is, X*50% of the total) should be filled by black, brown, and indigenous students[5] (from this point forward, I shall refer to this group as non-white) from public schools, where X represents the share of non-white students in the respective HEI's State population according to the 2010 Census. Figure 2.1 summarizes the rules of the law in a diagram.

---

[5] Although an individual's racial information is self-stated, successful candidates that were accepted to Universities through the quotas are subject to have their profiles evaluated by the University's Commission and/or by a Public Prosecutor based on the Statute of Racial Equality, mitigating any incentives for white students to declare themselves as non-white.

Figure 2.1 - The 2012 Law of Quotas
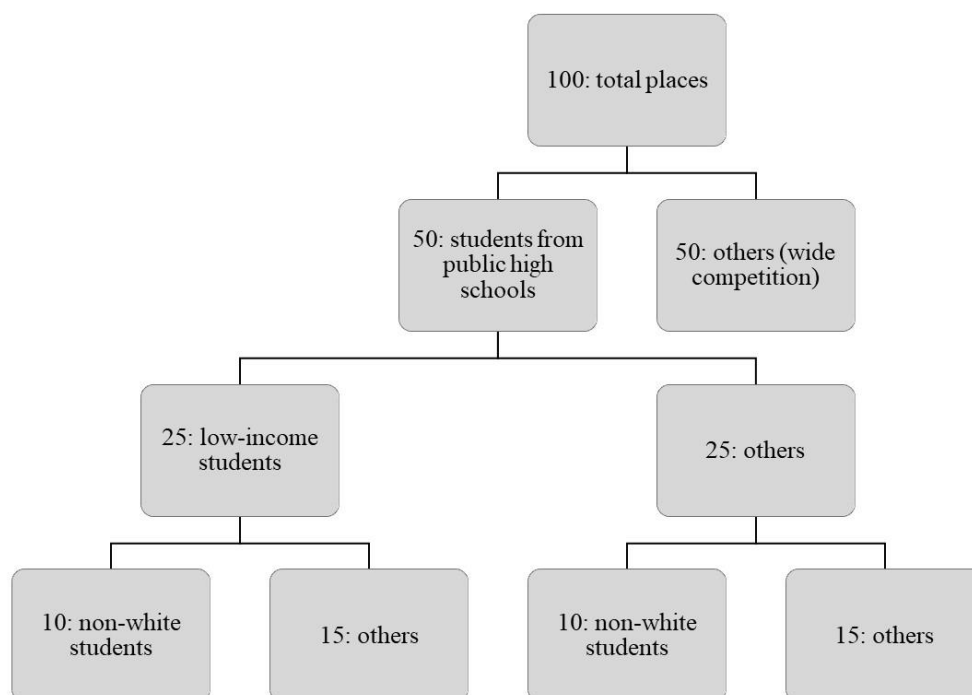
Note: The diagram above presents a simulation of the reserved places in a hypothetical federal university in which the total number of places equals 100 and in which the State's non-white individuals' percentage is of 40% (hence, 40% of 25 = 10).

Although it was announced in 2012, the law stated that HEIs had until 2016 to fully implement the quotas, but a minimum of 25% of the reserved seats should be implemented in each year from 2013 onwards. Therefore, universities had to reserve at least 12.5% of their seats in 2013, at least 25% in 2014, at least 37.5% in 2015 and finally the pre-established share of 50% in 2016, at the latest. Figure 2.2 illustrates the rate at which the law was implemented by universities.

Figure 2.2 - Average Percentage of Reserved Places in Federal Universities throughout the Years
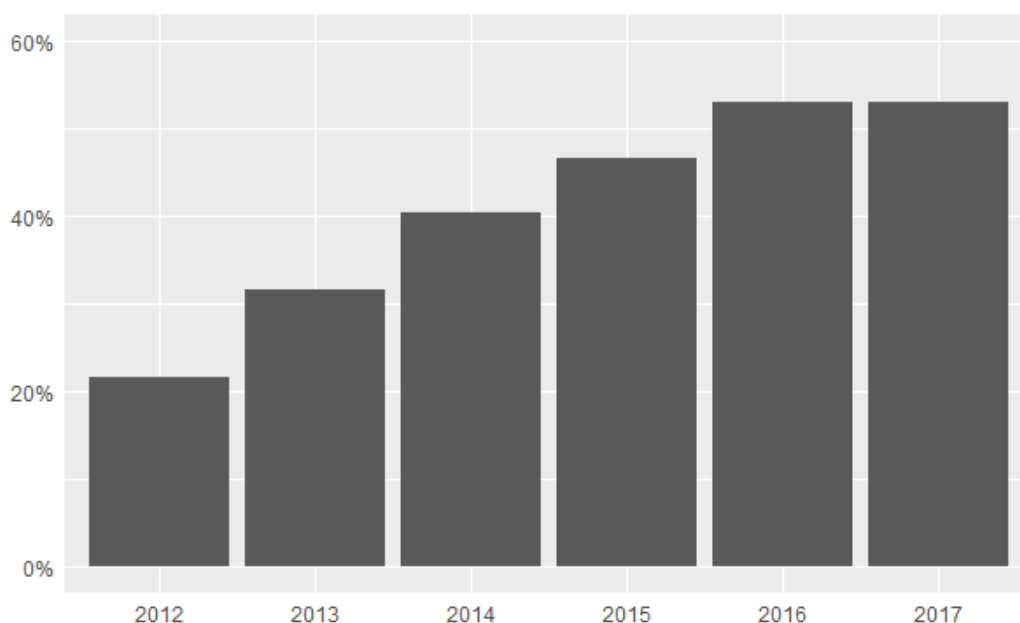
2.4. Data and methodology

2.4.1. Data

This research uses publicly available ENEM microdata, which has been published yearly since 1998 by the INEP (*Instituto Nacional de Estudo e Pesquisas Educacionais Anísio Teixeira*), an agency linked to the Ministry of Education. This repeated cross-sectional database provides information on the ENEM scores and key individual and household level variables of all students who sat the exam. Until 2008, the ENEM consisted of one essay and 63 multiple choice questions, with a score ranging from 0 to 100. In 2009, however, the exam was completely reformulated and, from that year onwards, the number of multiple choice questions increased to 180 (divided into 4 categories: natural science, social science, mathematics and languages) and all of the scores, including the essay, were measured on a scale from 0 to 1000, using the Item Response Theory[6]. Additionally, from 2009 onwards, the scores between different years have become comparable (*ENEM – Guia*

---

[6] The probability of obtaining a correct answer is assessed according to its difficulty, the probability that a student could guess a correct answer, and its ability to discriminate against students.

do participante) and students were consequently allowed to use scores from previous years in the SISU application.

The ENEM's microdata on individual and household characteristics comes from a mandatory self-declared questionnaire that all the candidates must fill out when signing up for the exam. The survey contains questions on basic socioeconomic factors (such as race, gender, age, marital status, city of residence, etc.), level of family income, parental education, high school record (if the candidate has ever been held back or dropped out of school), work-related factors (if and how much had the candidates worked during their lifetime), school type in which the candidate was enrolled during high school and fundamental school (public or private), among others. It is important to note that the ENEM's microdata is not a panel data, since the set of students that take the exam changes every year, and even if a student takes the ENEM exam more than once, the database does not allow us to track this student's performance over time.

This research focuses on the years from 2010 to 2016, the latter being the final year for the Law of Quotas to be fully implemented by all the institutions. Years prior to 2010 or after 2016 will not be incorporated in the model, since the format and many of the mandatory questions from the ENEM's survey changed as of those years. I select students who had already completed high school or were to complete it in the year of the exam and those who actually attended the test. Hence, students who only signed up for the exam but did not take it and those who were taking it as a practice test before graduating high school are excluded. I shall focus the analysis on the school and racial criteria of the law, since the ENEM's microdata only discloses income information on intervals of minimum wage, which hampers the evaluation of the effects of the law on individuals who are on the threshold of the income criteria.

Table 2.2 - Key Variables and Table 2.3- Summary Statistics provide the definitions and descriptive statistics of the variables included in the models to be presented in subsection 2.4.2. As displayed in Table 2.3- Summary Statistics, the percentage of non-white applicants on the ENEM exam increased significantly throughout the entire timespan of the database, especially after the implementation of the Law of Quotas (first put into effect in 2013). Furthermore, the variables used in the study were chosen such that the largest amount of available data could be preserved (that is, priority was

given to the mandatory questions of the ENEM' questionnaire). Except for the Race/Ethnicity variable, all of them have no missing values, and, as shown in Table 2.3- Summary Statistics, the percentage of missing information for Race/Ethnicity is low and this was dropped from the analysis.

Table 2.2 - Key Variables

Variables	Description
Age	Numerical
Gender	Masculine or feminine
Marital status	Single, married, divorced or widowed
State	State of residence (27 federative units of Brazil)
Degree of ruralization	Percentage of rural households in the city of residence
High school type	Entirely in public school, entirely in private school or mixed
Average income	Average per capita family income in intervals of minimum wages
Race/Ethnicity	White, black, brown, indigenous or other
Parental education	Parental higher degree of education (6 categories)
Work factor	Dummy: 1 if student has ever worked before
Dropout/Grade repetition	Dummy: 1 if student has ever been held back or dropped out in HS

Source: ENEM's Microdata – INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira)

Table 2.3- Summary Statistics

Variables	2010	2011	2012	2013
Age - mean (sd)	22.5 (7.2)	22.2 (7.0)	22.3 (7.2)	22.7 (7.5)
Gender (M; F)	40%; 60%	40%; 60%	41%; 59%	42%; 58%
Marital status (S; M)	86%; 14%	86%; 14%	87%; 12%	86%; 12%
Ruralization - mean (sd)	11.5% (17%)	12% (17%)	12.1% (17%)	12.4% (17%)
High school (pub.; priv. + mix.)	79%; 21%	79%; 21%	79%; 21%	80%; 20%
Average income - mean (sd)	1 mw (1.5)	0.7 mw (1.1)	0.8 mw (1.05)	0.7 mw (1.02)
Work factor	55% Y; 45% N	54% Y; 46% N	59% Y; 41% N	61% Y; 39% N
Dropout/Grade Repetition	19% Y; 81% N	19% Y; 81% N	18% Y; 82% N	19% Y; 81% N
missing	*0%*	*0%*	*0%*	*0%*
Race (white, brown, black)	45%; 40%; 12%	43%; 41%; 12%	43%; 42%; 12%	40%; 44%; 13%
missing	*3.3%*	*2.4%*	*1.7%*	*1.6%*

Variables	2014	2015	2016	2010-2016
Age - mean (sd)	23.1 (7.7)	22.5 (7.3)	22.3 (7.2)	22.5 (7.3)
Gender (M; F)	42%; 58%	42%; 58%	42%; 58%	41%; 59%
Marital status (S; M)	85%; 13%	88%; 10%	89%; 9%	87%; 12%
Ruralization - mean (sd)	12.3% (17%)	11.9% (17%)	12% (18%)	12% (17%)
High school (pub.; priv. + mix.)	83%; 17%	81%; 19%	81%; 19%	81%; 19%
Average income - mean (sd)	0.7 mw (1.02)	0.8 mw (1.1)	0.7 mw (1.02)	0.7 mw (1.1)
Work factor	62% Y; 38% N	59% Y; 41% N	54% Y; 46% N	58% Y; 42% N
Dropout/Grade Repetition	17% Y; 83% N	16% Y; 84% N	15% Y; 85% N	17% Y; 83% N
missing	*0%*	*0%*	*0%*	*0%*
Race (white, brown, black)	39%; 44%; 13%	38%; 46%; 13%	36%; 47%; 14%	40%; 44%; 13%
missing	*1.4%*	*1.7%*	*1.7%*	*1.9%*

Source: ENEM's Microdata – INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira)

2.4.2. Methodology

The fact that the Law of Quotas only applied to certain students and the presence of "clean" individuals who were unaffected by it makes difference-in-differences an appropriate methodology to evaluate the causal effects of the law on students' ENEM scores. Let us previously recall that the law applies to all students who have attended public high schools, but it provides special benefits for those who are non-white, since among the reserved seats based on the school criteria, there is a pre-established share of seats based on the racial criteria (Figure 2.1). I will then separately estimate the effects of each one of these components of the Law of Quotas, which shall be called the school component and the racial component, on the students' ENEM scores. However, for reasons that shall be explained forthwith, I shall only venture further into the results of the racial component model.

I then estimate the difference-in-differences model in two steps. First, I investigate solely the impact of the school component of the law. Hence, the ENEM scores are observed before and after the introduction of the Law of Quotas in 2013, and between two groups: a treatment group composed of white students from public high schools (therefore impacted by the school component but not by the racial component of the law); and a control group composed of private high school students[7] from all ethnicities (i.e., not eligible for any quotas).

Secondly, I assess the impact of the racial component of the law; that is, the effect on non-white public school students. In this case, the ENEM scores are again observed before and after the introduction of the law; however, the treatment group is now composed of non-white students from public high-schools, and the control group is composed of white students from public high schools (that is, both groups are affected by the school component of the law – since they are both public high school students - but only the treatment group is affected by the racial component). Table 2.4 summarizes the grouping of the models.

---

[7] Those who attended only a part of high school in a private institution also compose the control group, since they are not eligible for the quotas. For simplicity, I shall refer to this group as private high school (or simply private school) students

Table 2.4 - Model's grouping summary

	Model	
	School component model	Racial component model
Treatment Group	White public school students	Non-white public school students
Control Group	Private school students (all ethnicities)	White public school students

Standard approaches for causal inferences in difference-in-differences are valid only under the assumption that the treatment and control groups display parallel trends before the policy intervention – which additionally is assumed to be a good post-treatment counterfactual. The previous trends for the control and treatment groups in the school component model are presented in Figure 2.3, while Figure 2.4 presents these trends for the racial component model. The figures show that, prior to the implementation of the Law of Quotas in 2013, the ENEM objective scores of the control and treatment groups present relatively similar trends. Nonetheless, so as to strengthen the validity of the results, I use the method suggested by Rambachan and Roth (2019) for robust inference in difference-in-differences settings valid even when the parallel trends assumption does not hold exactly. This method consists of evaluating robustness of the difference-in-difference results to some degree of deviation from the pre-existing difference in trends[8]

---

[8] Rambachan and Roth (2019) introduce a parameter M which governs the maximum amount by which the slope of the pre-treatment difference in trends can change between consecutive periods. See Rambachan and Roth (2019) for further details.

Figure 2.3 - Yearly Average ENEM Objective Score (2010 = 100). Control Group: Private School Students; Treatment Group: White Public School Students
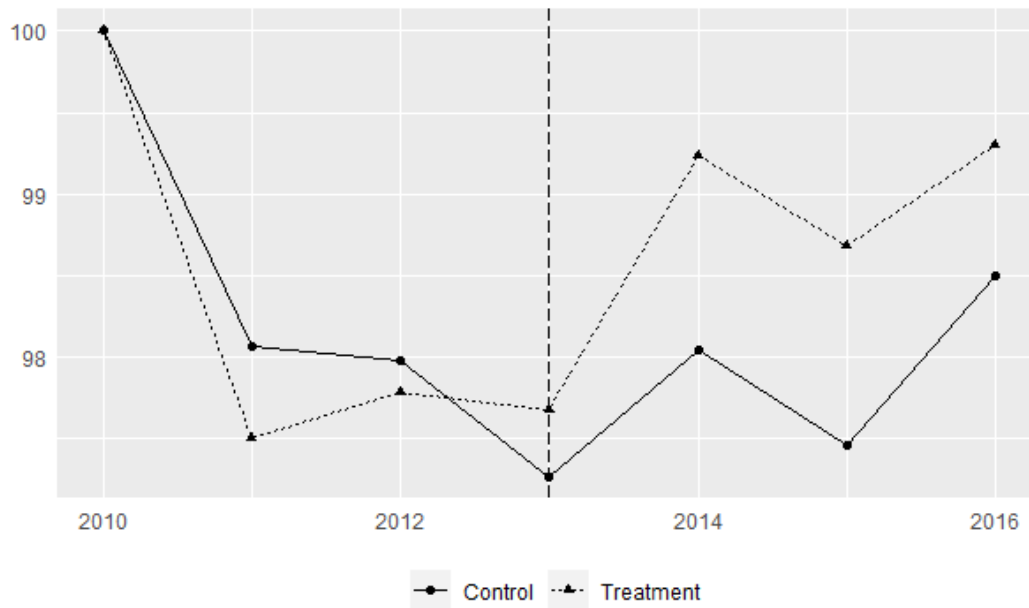
Figure 2.4 - Yearly Average ENEM Objective Score (2010 = 100). Control Group: White Public School Students; Treatment Group: Non-white Public School Students
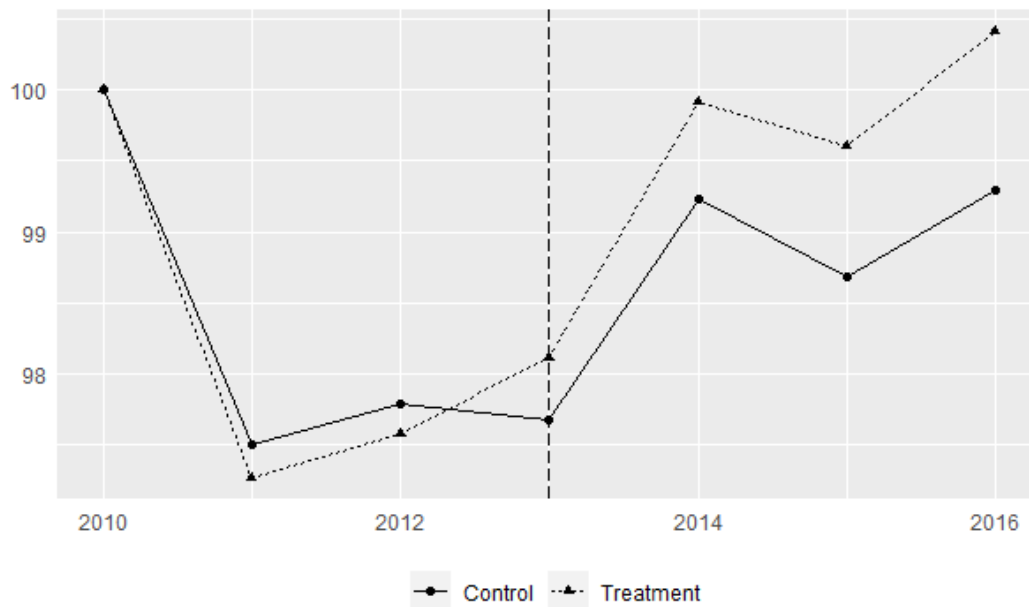
In order to apply the methodology proposed by Rambachan and Roth (2019), the following dynamic event-study regression is estimated:

$$\log{(S_{it})} = c_0 + \phi_t + \lambda D_i^t + \gamma X_{it}$$
$$+ \sum_{s \neq 2012} \beta_s \times \mathbb{1}[t = s] \times D_i^s + \varepsilon_{it} \qquad (2.1)$$

where S_{it} denotes the ENEM total score on the multiple choice exam of a student i who took the exam in a specific year t; D_i^t is a dummy variable that equals one if the individual belong to the treatment group; $\hat{\phi}$ and $\hat{\lambda}$ measure the time-specific and group-specific fixed effects, respectively; X_{it} includes the student-specific control variables described in Table 2.2; and the coefficients $\{\hat{\beta}\}$ account for the event-study coefficients (which measure the causal effect of the treatment plus the difference in trends between the treatment and control groups), where $\hat{\beta}_{2012}$ is normalized to zero.

The results of the estimations will be presented in Section 2.5 as follows. First, I present the results of Rambachan and Roth (2019) analysis for both the school component and racial component model. For reasons that will be explained hereafter, for the school component model only, I shall explore these results in further depth.

First, a standard two periods difference-in-differences regression (2PDD), following equation 2.2, is estimated:

$$\log{(S_{it})} = c_0 + \Phi W + \lambda D_i^t + \beta D_i^r W + \varepsilon_{it} \qquad (2.2)$$

where W, which is a dummy variable that equals one if $t \geq 2013$ (that is, if it belongs to the post-treatment period), is introduced, hence the coefficient $\hat{\beta}$ measures the average treatment effect throughout the entire post-treatment period.

Second, I use the 2PDD framework in order to assess the heterogeneous effects of the law on the different subjects of the exam – which is done by simply using the score in each subject as the dependent variable in the regressions – and between genders and students with and without college-educated parents. For the two latter, the following regressions are estimated:

$$\log{(S_{it})} = c_0 + \Phi W + \Phi^m W D_i^m + \lambda D_i^r + \lambda^m D_i^r D_i^m$$
$$+ \gamma X_{it} + \gamma^m X_{it} D_i^m + \beta D_i^r W + \beta^m D_i^r D_i^m W \qquad (2.3)$$
$$+ \varepsilon_{it}$$

$$\log\left(S_{it}\right) = c_0 + \Phi W + \Phi^c W D_i^c + \lambda D_i^r + \lambda^c D_i^r D_i^c + \gamma X_{it}$$
$$+ \gamma^c X_{it} D_i^c + \beta D_i^r W + \beta^c D_i^r D_i^c W + \varepsilon_{it} \tag{2.4}$$

where D_i^m is a dummy variable that equals one if the student is male and $\Phi^m, \lambda^m, \gamma^m$ and β^m refer to the incremental time-fixed, group-fixed, control variables and treatment effects, respectively, for male individuals. Analogously, D_i^c is a dummy variable that equals one if either the mother or the father of the student holds a college degree, and $\Phi^c, \lambda^c, \gamma^c$ and β^c refer to the incremental time-fixed, group-fixed, control variables and treatment effects, respectively, for these individuals.

Finally, I estimate a regression with dynamic treatment effects, following equation 2.1 (using the total score on the ENEM's multiple choice exam as the response variable), in which all of the pre-treatment betas (i.e. $\hat{\beta}_{2010}$ to $\hat{\beta}_{2012}$) are normalized to zero[9].

2.5. Results

Figure 2.5 and Figure 2.6 plot set-identified estimations of the treatment effect $\hat{\beta}_{2016}$ among different deviations of the pre-existing differential trend for the school component and racial component model, respectively. The year of 2016 is used as reference in this analysis since it is the latest year of the database as well as the final year for the Law to be fully implemented.

---

[9] I have also estimated a model with dynamic treatment effects in which all the pre-treatment betas are allowed to vary, but it does not induce any significant changes in the treatment effect coefficients

Figure 2.5 - School Component Model Treatment Effect ($\hat{\beta}_{2016}$) Sensitivity to Parallel Trends Violations. Original = OLS estimation; FLCI = Optimal Fixed Length Confidence Interval
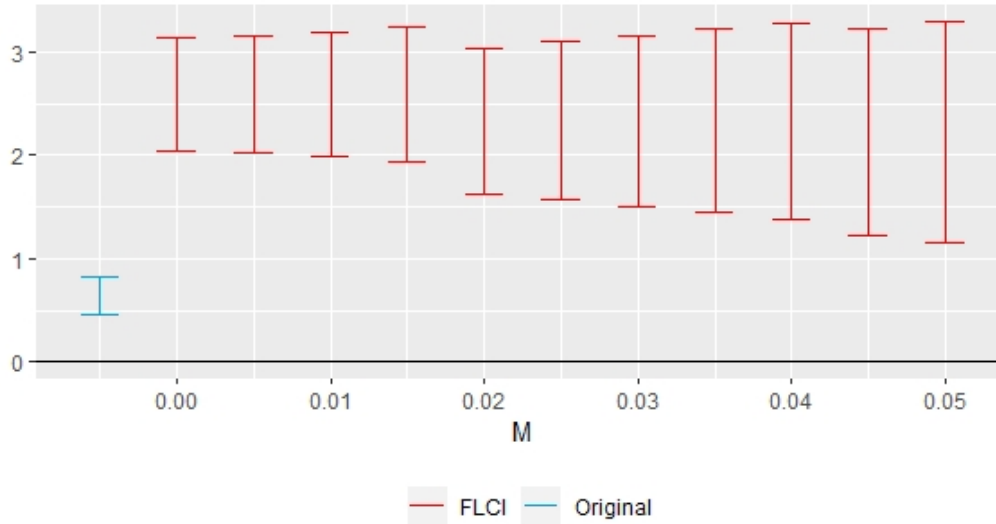
Figure 2.6 - Racial Component Model Treatment Effect ($\hat{\beta}_{2016}$) Sensitivity to Parallel Trends Violations. Original = OLS estimation; FLCI = Optimal Fixed Length Confidence Interval
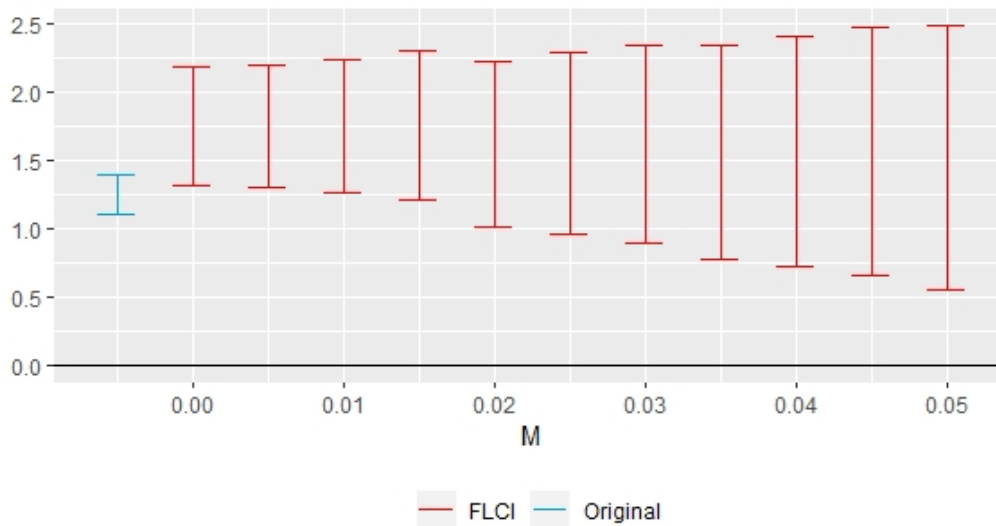
In the figures, the original OLS estimate shows the estimated treatment effect assuming that the parallel trends hold exactly, while the remaining estimates consider linear extrapolations of the pre-treatment differential trend ($M = 0$ considers an exact linear extrapolation, while $M > 0$ accounts for changes in the slope of the pre-treatment differential trend).

39

Figure 2.5 shows that the effect of the school component on the students' ENEM scores is positive under the entire set of violations considered (up to $M = 0.05$). In fact, M would need to be as large as 0.11 (that is, the slope of the pre-treatment differential trend would have to change by 0.11 between years in the post-treatment period) in order to reject the null hypothesis that the treatment effect is significant (and, in this case, positive). Therefore, it can be argued with reasonable confidence that the school component of the law did indeed have a positive impact on the eligible students' ENEM scores, a result that is robust to significant variations in the extrapolation of the pre-treatment difference in trends (within the range of M evaluated in Figure 2.5, a treatment effect that ranges from a little above 1% to almost 4% is estimated). However, due to the great difference in assuming the OLS original estimate and any estimate with positive M, I shall not go into further details on this model.

Figure 2.6 shows that the racial component of the law also had a positive effect on the treatment group's ENEM scores (varying from approximately 0.5% to 2.5% within the range of M considered in the analysis). Again, this effect is robust to significant deviations of the pre-treatment differential trend's linear extrapolation. Furthermore, there is no statistical difference between the OLS original estimation and the set-identified estimation with positive M. The section thus proceed as follows. First, I present the results of a standard two periods difference-in-differences model, in which the average treatment effect for the entire post-treatment period is estimated (2013 to 2016) – more precisely, the results of a model with and without the set of control variables are contrasted. Then, I test for heterogeneous effects by subject of the ENEM exam, gender and parental education. Third, I estimate a difference-in-differences model with dynamic treatment effects to investigate whether the impact of the law varied throughout the years. Finally, the results for a set of robustness checks are presented in section 6. Except when stated otherwise, all the models make use of the same control and treatment groups (i.e., the racial component model control and treatment groups).

I start by estimating a 2PDD model, contrasting its' results both with and without the set of control variables. The first column in Table 2.5 presents the results of the regression without controls, while the second column contains the results of the regression with controls. The table shows a positive

and significant estimated treatment effect that does not change significantly between the short and the augmented models, despite a substantial increase in the R-squared. Oster (2019) suggests a test for unobservable variable bias based on Altonji, et al. (2005) which makes use precisely of these two information (the change in coefficients and R-squared between the regression with and without control variables). Following Oster's recommended specification with $R_{max} = 1$, the bounding set becomes $[\tilde{\beta}, \beta^*(1,1)]$, where $\beta^*(1,1) = \tilde{\beta} - \frac{(\dot{\beta}-\tilde{\beta})(1-\tilde{R})}{\tilde{R}-\dot{R}}$ [10]. The recommended bounding set in this case is [0.0078, 0.0104], which safely excludes zero, thus providing evidence that the significant estimated treatment effects observed in Table 2.5 are not driven by non-observable factors.

In the augmented model, an average treatment effect of 1.04% is estimated. In other words, this model suggests that the racial component of the Law of Quotas induced eligible students to attain a 1.04% higher score in the ENEM exam, on average, during 2013 to 2016.

Table 2.5- Racial Component Model without control variables: Standard Two Periods Regression (all coefficients multiplied by 100)

Independent Variables	(1)	(2)
Group Fixed Effect	-4.475 *** (0.037)	-2.088 *** (0.039)
Time Fixed Effect	0.440 *** (0.039)	0.320 *** (0.036)
Treatment Effect	1.093 *** (0.050)	1.044 *** (0.046)
Control	No	Yes
Observations	1,124,157	1,124,157
R²	2.31%	18.11%

Robust standard errors in parenthesis
* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

---

[10] $\tilde{\beta}$ and $\tilde{R}$ account for the treatment coefficient and R-squared in the augmented regression, $\dot{\beta}$ and $\dot{R}$ for the treatment coefficient and R-squared in the short regression, and R_{max} is the R-squared of a hypothetical model that includes both observed and unobserved controls. See Oster (2019) for further details.

I now use the same two-periods design to investigate whether there were heterogeneous effects of the racial component of the law by (i) subject of the exam; (ii) gender; and (iii) level of parental education.

First, I examine the effects per subject. As mentioned in subsection 2.3.3, the ENEM consists of one multiple choice exam (containing four different disciplines: natural sciences, social sciences, languages and math) and one essay. Thus, five separate regressions are estimated, so that in each model, the response variable is the student's score in each of the five tests. Figure 2.7 displays the estimated treatment effect coefficient per subject plus the one for the overall objective exam already presented.

The effects of the racial component of the law were greater in more quantitative-intensive fields (natural sciences and mathematics). Actually, the estimated treatment effect coefficient for the essay is not statistically significant, indicating that the racial component of the law might not have had any effect on this specific part of the exam.

Figure 2.7 - Treatment Effect Coefficient (multiplied by 100) per ENEM Subject
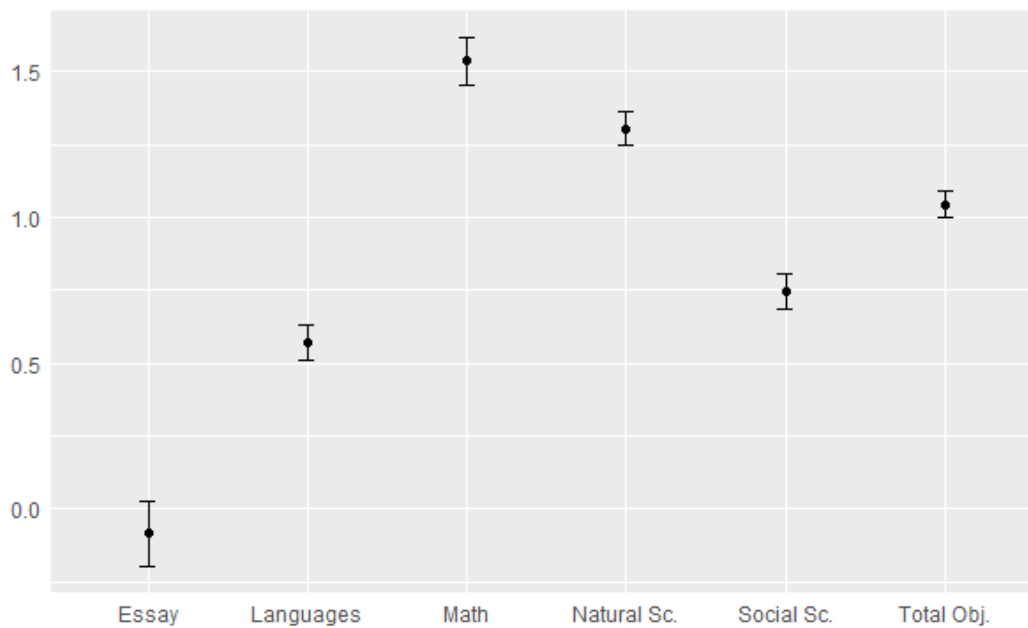
Following, I assess the possibility of heterogeneous effects of the racial component between genders and students with and without college-educated parents, by estimating the regression from equations 2.3 and 2.4. Table A.2.1 and Table A.2.2 in the appendix contain the results of these estimations, which show that the difference-in-difference coefficients are not statistically

different between genders or between students whose parents hold different levels of education[11].

Finally, I investigate whether the effect of the racial component of the Law of Quotas evolved over time. The dynamic treatment effect coefficients are presented in Figure 2.8 , which indicates that the effect of the law on the students' ENEM scores in the objective exam increased throughout the first years after its implementation (the coefficients $\hat{\beta}_{2014}$ and $\hat{\beta}_{2015}$ are not statistically different at a 5% significance level, but the remaining coefficients in fact are).

Figure 2.8 - Treatment Effect (multiplied by 100) throughout the Years
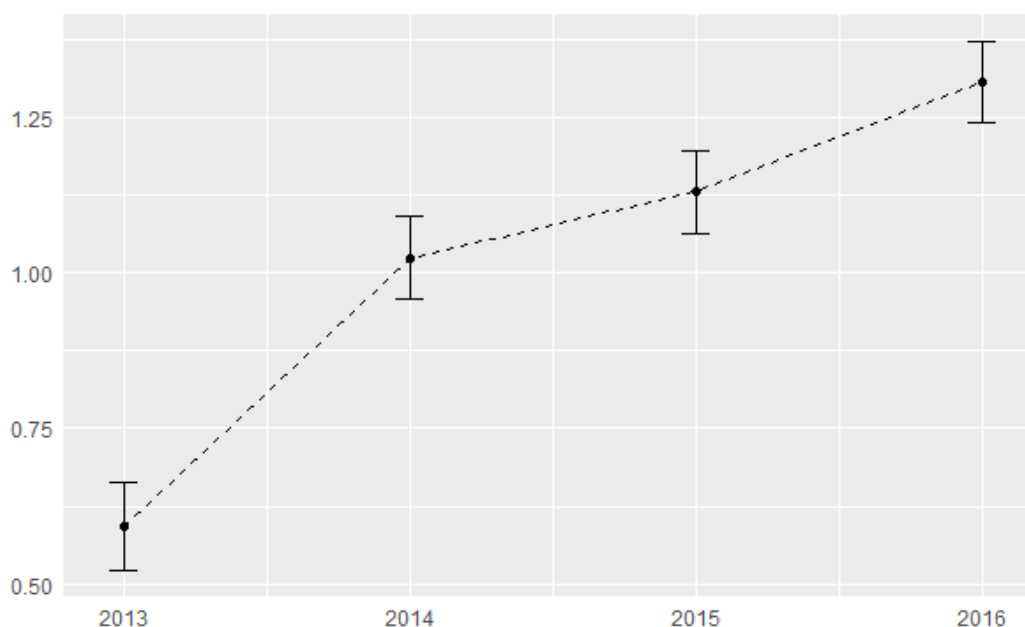
2.6. Robustness

In this section, a battery of robustness exercises are conducted in order to further qualify the findings from the racial component model. First, I check for the anticipatory effects of the law. Second, I perform a placebo test, in which only private school students (who are not eligible for the Law of Quotas at all) are used in the model's population. Third, I carry out another placebo test, using 2017, in which the law had already been fully implemented, as the cutting year in the difference-in-differences model.

---

[11] The estimated treatment effects are, however, different among levels of parental education at a 10% significance level

Fourth, I reevaluate the results including extra years to the regression. Fifth, I reevaluate the results with re-standardized ENEM scores. Sixth, I assess robustness of the results to a pre-processed dataset using an Entropy Balancing methodology. Finally, I use an alternative database to estimate the effects of the Law of Quotas on non-white students' high school completion rate – using the same two-periods difference-in-differences framework presented in previous subsections.

I start by checking for anticipatory effects of the treatment, that is, whether the law had any effect on students' ENEM scores before it was implemented (prior to 2013). First, it should be noted that it is unlikely that there should be any anticipatory effects, since (i) the law was published on August 29th, 2012, two months before that year's exam, and (ii) the law stated that the quotas (or at least a share of them) should be implemented by the institutions only from 2013 onwards. A visual inspection of Figure 2.4 suggests that an increase in the scores of the treatment group (supposedly due to the treatment effect) was found only in 2013, which would rule out the possibility of anticipatory effects. Nevertheless, I also estimate a two periods model following equation 2.2 but excluding the years right before and right after the law's implementation (2012 and 2013) from the regression. The treatment effect coefficient in this setting (equal to 1.17 –Table A.2.3 in the Appendix) remains significant and very close to the coefficient estimated in the augmented model from Table 2.5, which strengthens the hypothesis that there were indeed no anticipatory effects.

In the second robustness exercise, I perform a placebo test using only students from private high schools in both the treatment and control groups. The concern here is that the increase in the ENEM score of non-white students is driven by some other factor other than the Law of Quotas, such as noisy data or some unobserved racial driver. Therefore, a 2PDD model is estimated using non-white students from private schools as the treatment group and white students from private schools as the control group. Despite the racial difference between the groups, both of them are private school students, which means that they are not eligible for the quotas and we should not see any significant treatment effect. The treatment effect coefficient in this case is insignificant (p-value of 28% - Table A.2.4 in the Appendix), suggesting that indeed there has been no effect of the law on private school students.

Third, I perform another placebo test, in which 2017 is used as the cutting year in the difference-in-differences model. In section 2.5, it was found that the effects of the racial component of the law increased in each year up to 2016. From 2017 forward, however, it would be reasonable to see a stabilization of the effects of the law, due to the number of years since its implementation and to the fact that by 2017 the quotas had already been fully implemented by all institutions. Therefore, I estimate a two periods difference-in-differences regression in which the post-treatment period comprises the years of 2017 and 2018. These results shall be taken carefully since this model does not include the work factor and previous academic effort factor variables due to changes in the ENEM's questionnaire. Nevertheless, the estimated treatment effect coefficient is insignificant in this case (p-value of 71% - Table A.2.5 in the Appendix), which suggests both that the results obtained were not merely a placebo effect and that by 2017 the impacts of the law had completely stabilized.

Fourth, I reestimate the model including extra years to the analysis. Since many of the ENEM's survey questions changed throughout the years, some restrictions to the timespan had to be imposed to the main models. However, as an additional robustness check, I reestimate the 2PDD racial component model including one pre-treatment year (2009, the year in which the exam was reformulated) and two post-treatment years (2017 and 2018), with the caveat that the work factor and the dropout/grade repetition control variables were not added to this exercise. Table A.2.6 in the Appendix presents the estimated coefficients from this regression, in which the treatment effect coefficient is very similar to the one in Table 2.5.

Fifth, since there is a substantial drop in scores from 2010 to 2011 (as seen in Figure 2.3 and Figure 2.4), which could raise some red flags concerning the reliability of the estimations, I re-standardize the ENEM scores around zero, so that the average grade is the same in all years. The reader may note, however, that this is a re-standardization of the scores, since the ENEM is already designed such as its' grades are comparable between years. Table A.2.7 in the Appendix presents the results of this model, in which a positive and significant treatment effect coefficient is found.

Next, I assess robustness of the model's results to a pre-processed and rebalanced database. Since the Law of Quotas was not randomly assigned, the causal conclusions derived from such observational data might be

polluted by covariate imbalance. Therefore, the control group in the racial component model is reweighted using Entropy Balancing (Hainmueller 2012), a method which intends to match the covariate moments for the different experimental groups and is double robust with respect to linear outcome regressions (Zhao and Percival 2017). The results of the 2PDD racial component model estimation with EB is displayed in Table A.2.8 in the Appendix, which shows that the treatment effect coefficient is again significant and similar, although slightly lower, than the one presented in Table 2.5.

Lastly, I perform a final exercise with a similar difference-in-differences framework as the one from equation 2.2 (a two-periods regression) so as to evaluate the impacts of the Law of Quotas on the high school completion rate of non-white individuals. For this estimation, however, rather than working with the ENEM microdata, I make use of IBGE's PNAD (*Pesquisa Nacional por Amostra de Domicílios*) – Brazil's national household sample survey, a yearly repeated cross-sectional database with information on housing, demography, migration, education, labor and income at both individual and household levels. This analysis focuses on young individuals of high-school graduate age (18 to 24 years old) and at the years from 2011 to 2015[12]. I then estimate a logistic difference-in-differences regression in which the output of interest is a dichotomous variable indicating whether the student completed high school. The treatment group in the regression is composed of non-white individuals and the control group by white individuals. Finally, a set of control variables is added to the model, similar –although not identical- to the ones described in Table 2.2 - Key Variables, from the PNAD database. Figure A.2.1 in the Appendix exhibits the evolution of the high school completion rate for the treatment and control groups, while Table A.2.9 presents the results of the estimation. Results show that the treatment effect is positive and significant, providing evidence that the Law of Quotas had a positive effect on non-white students' pre-college effort (as measured by

---

[12] I abstract from 2010 since the survey was not conducted in that year due to the 2010 Census and from 2016, since from that year onwards the PNAD was replaced by its latest version, the PNAD *Contínua*

their high school completion rate), and hence yielding further robustness to the previous findings.

2.7. Discussion

The results obtained suggest that both the school component and the racial component of the Law of Quotas induced eligible students to attain higher scores on the ENEM exam. The possible presence of some pre-treatment difference in trends in the school component model hampers a more detailed evaluation of the impacts of this element of the law, but the sensitivity analysis indicates that the significance of the treatment effect is robust to substantial violations in the parallel trends assumption. Since there is no certainty on the direction and magnitude of the post-treatment differential trend, I shall not venture further into this result.

For the racial component model, however, it was possible to explore the results in more depth. First, I estimated a standard two-periods difference-in-differences model, which indicated that the racial component of the Law of Quotas induced eligible students to attain a 1.04% higher score in the ENEM exam, on average, during 2013 to 2016.

Second, I checked for heterogeneous effects, the main findings being: i) the effect of the law was stronger in quantitative-intensive subjects (Math and Natural Sciences) than it was in the remaining fields (Language, Social Sciences and the Essay); and ii) the racial component of the Law of Quotas did not exert statistically significant differences by gender nor parental education. A possible explanation to the former might be that quantitative-intensive subjects might be less dependent on socioeconomic background (in other words, hours of self-study for the ENEM exam in mathematics are less conditioned to the students' social and home environment). Indeed, a number of research studies have suggested that math achievements tend to be more sensitive to teachers and schools' efficiency gains, while reading/linguistic achievements might be more dependent on socioeconomic status and parental occupation and/or involvement at school (Perry and McConney 2013; Cheadle 2008; Rimm-Kaufman et al. 2007; Sui-Chu and Willms 1996).

Third, I have also estimated a difference-in-differences model with dynamic treatment effects in order to evaluate whether the impact of the law evolved throughout the years after its implementation. It was found that the treatment effect indeed increased from 2013 to 2016 and, therefore, this appears to be

47

a case in which the effect of the policy intervention depends on the length of exposure to it. That is, while a student from the treatment group that took the exam in 2016 had four years to absorb the effects of the treatment and increase their investment in human capital, an individual that took the exam in 2013 had only one year to do so. An alternative and perhaps complementary explanation is that the increasing treatment effect is due to the design of the Law of Quotas. Since the law stated that universities had until 2016 to fully implement the quotas, the share of reserved seats presented an upward trend from 2013 to 2016 (see Figure 2.2), which could explain part of the dynamic observed in Figure 2.8. In any case, the "incentive" effect of the policy clearly outweighed the possible "relaxation" effect on students of the increase in the number of seats.

A possible concern that could arise from the estimation of the school component model is that the Law of Quotas could have increased competition for seats among private school students and therefore have impacted their pre-college performance as well, which would put the suitability of the control group at stake. However, the reduction of available seats for these individuals due to the law's implementation was attenuated by an overall increase in the number of seats in federal universities by 41% from 2012 to 2016. As can be seen in Figure 2.3, the average ENEM score among private school students did not present significant changes after the introduction of the law (it may have decreased from 2010 to 2011 but remains reasonably stable thereafter). In order to qualify this hypothesis, I also estimated a regression for private school students only from 2011 to 2016 (excluding the drop in scores from 2010 to 2011 - before the Law of Quotas) with a time-fixed effect dummy that equals one from 2013 onwards (that is, after the law was implemented) and found that this coefficient is insignificant (p-value of 33% - Table A.2.10 in the Appendix), which corroborates the hypothesis that the law did not affect the private students' scores. In the same manner, it is assumed that this 41% increase in the overall number of seats from 2012 to 2016 also mitigated any increase in competition for seats among white students from public schools that might have arisen from the racial component of the law.

Finally, the results suggest that the positive incentives provided by affirmative action, such as the mitigation of the discouragement effects described by Cotton et al. (2016) and Furstenberg (2003), have prevailed over

any negative incentive effect that might have stemmed from the policy. Although the empirical investigations that have been previously performed were limited and mainly focused on specific universities, the majority of them pointed towards a positive effect of higher education quotas on pre-college effort and academic performance as well. Hence, the results from this study both corroborate and strengthen these previous findings.

These results have strong policy implications as they indicate that educational quotas not only enhance the participation of disadvantaged groups in higher education directly through an increased number of seats but also by encouraging these individuals to invest in human capital and close the performance gap by the end of secondary education. Therefore, this behavioral response to the implementation of quotas should not be overlooked and should be taken into account by policymakers, especially in developing economies with a high level of inequality in education.

2.8. Conclusion

Several different measures have been implemented in recent years in an attempt to mitigate racial inequalities in education. One sort of intervention has been the establishment of reserved seats in higher education to specific racial groups and, although there is a rich body of evidence that investigates the ex-post effects of these quotas, little research has been done with respect to the effects that they have on pre-college academic performance. I contribute to this literature by evaluating how the Brazilian 2012 Law of Quotas affected the performance of students on the college-entrance exam, the ENEM.

The results from this chapter suggest that both the school component and the racial component of the Law of Quotas fostered incentives to pre-college human capital accumulation as it induced eligible students to attain higher scores on the ENEM exam. Additionally, the positive effects of the racial component of the law increased throughout the first years after its implementation.

Furthermore, I have also tested for the presence of heterogeneous effects of the racial component of the law across a set of different dimensions. While racial quotas had a larger effect on the scores of quantitative-intensive subjects than it had on linguistic/humanities related subjects, no evidence of heterogeneous effects was found by gender or parental education.

Although robustness exercises scaffold the validity of these results, I acknowledge some limitations in the employed strategy. First, an indirect strategy for controlling for the income criteria of the quotas had to be taken, due to data restrictions. Second, I have controlled for a set of observable individual and socioeconomic characteristics, others remaining as non-observable. Third, since the ENEM's microdata does not disclose information on each candidate's SISU's application, it was not possible to control for the actual university the students finally enrolled at (or at least were accepted in). Nevertheless, sensitivity analyses allow us to provide strong evidence that the Law of Quotas implemented in Brazil did indeed encourage eligible students to increase their pre-college academic performance (i.e., that the introduction of quotas in higher education not only promotes equity, but also brings about efficiency gains). Thus, this research helps to shed some light on the incentives provided by quotas in higher education and hence might serve as a guide to educators and policy makers whose aim is not only to increase the equality of educational opportunity, but also the efficiency of their educational system.

Appendix

Figure A.2.1 - High School Completion Rate (PNAD classification - 2010 = 100). Control Group: White Individuals; Treatment Group: Non-white individuals
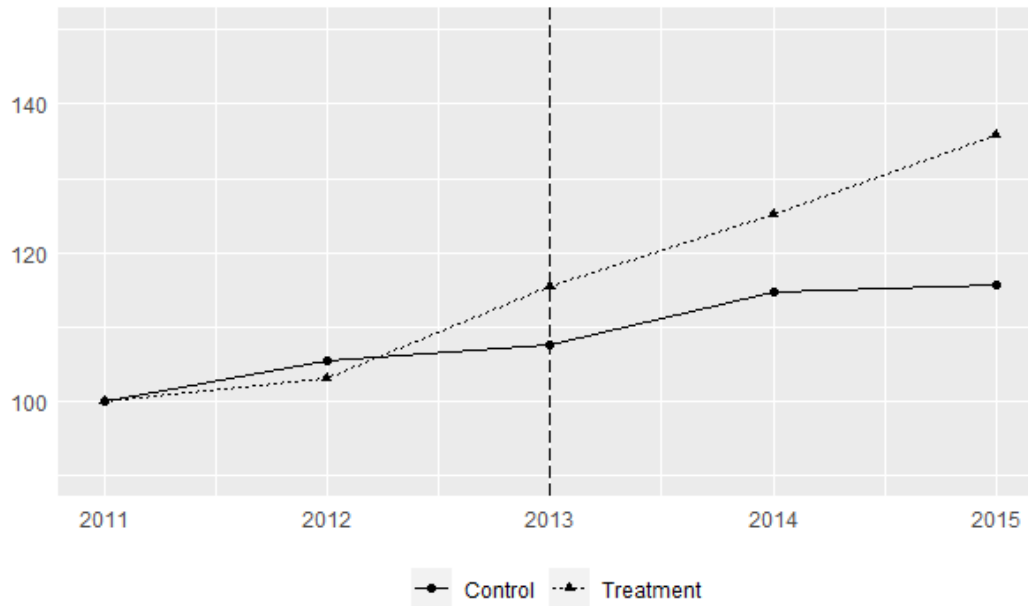
Table A.2.1- Racial Component Model per Gender (all coefficients multiplied by 100)

Independent Variables			
Male		Female	
Marital Status	Yes ***	Marital Status	Yes ***
State	Yes ***	State	Yes ***
Parental Education	Yes ***	Parental Education	Yes ***
Age	0.027 *** (0.004)	Age	-0.071 *** (0.002)
Average Income	3.701 *** (0.050)	Average Income	4.326 *** (0.062)
Ruralization	-6.712 *** (0.117)	Ruralization	-4.295 *** (0.083)
Work Factor	0.477 *** (0.045)	Work Factor	0.884 *** (0.031)
Dropout/Grade Repetition	-2.870 *** (0.046)	Dropout/Grade Repetition	-2.551 *** (0.037)
Group Fixed Effect	-2.275 *** (0.065)	Group Fixed Effect	-1.967 *** (0.048)
Time Fixed Effect	-0.397 *** (0.045)	Time Fixed Effect	0.806 *** (0.045)
Observations	1,124,157		
R²	18.38%		

Robust standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Note: For ease of exposition, I present the net coefficient for each gender (i.e., in the "Female" column I present the coefficients from Equation 5 that do not contain the "m" suffix, whereas in the "Male" column I present these same coefficients plus the incremental coefficients for males)

Table A.2.2 - Racial Component Model per Level of Parental Education (all coefficients multiplied by 100)

Independent Variables			
With College Degree		Without College Degree	
Marital Status	Yes ***	Marital Status	Yes ***
State	Yes ***	State	Yes ***
Parental Education	Yes ***	Parental Education	Yes ***
Age	0.081 *** (0.009)	Age	-0.041 *** (0.002)
Average Income	3.197 *** (0.063)	Average Income	4.332 *** (0.049)
Ruralization	-10.661 *** (0.226)	Ruralization	-4.537 *** (0.071)
Work Factor	-1.290 *** (0.082)	Work Factor	0.983 *** (0.027)
Dropout/Grade Repetition	-3.219 *** (0.107)	Dropout/Grade Repetition	-2.635 *** (0.030)
Group Fixed Effect	-2.597 *** (0.126)	Group Fixed Effect	-1.988 *** (0.040)
Time Fixed Effect	-0.132 (0.105)	Time Fixed Effect	0.387 *** (0.039)
Observations	1,124,157		
R²	18.32%		

Robust standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Note: For ease of exposition, I present the net coefficient for each gender (i.e., in the "Without College Degree" column I present the coefficients from Equation 6 that do not contain the "c" suffix, whereas in the "With College Degree" column I present these same coefficients plus the incremental coefficients for students with a college-educated parent)

Table A.2.3 - Robustness Check: Anticipation of Treatment Effect Regression (all coefficients multiplied by 100)

Independent Variables			
Marital Status	Yes ***	Ruralization	-4.876 *** (0.079)
State	Yes ***	Work Factor	0.997 *** (0.03)
Parental Education	Yes ***	Dropout/Grade Repetition	-2.693 *** (0.034)
Age	-0.040 *** (0.002)	Group Fixed Effect	-2.170 *** (0.047)
Gender (M=1)	3.681 *** (0.027)	Time Fixed Effect	0.467 *** (0.043)
Average Income	3.807 *** (0.043)	Treatment Effect	1.172 *** (0.056)
Observations	807,063		
R²	17.71%		

Robust standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.2.4 - Robustness Check: Placebo Test with Private School Students Regression (all coefficients multiplied by 100)

Independent Variables			
Marital Status	Yes ***	Ruralization	-8.215 *** (0.26)
State	Yes ***	Work Factor	-2.136 *** (0.069)
Parental Education	Yes ***	Dropout/Grade Repetition	-7.353 *** (0.104)
Age	0.023 ** (0.007)	Group Fixed Effect	-2.199 *** (0.092)
Gender (M=1)	2.845 *** (0.053)	Time Fixed Effect	-0.591 *** (0.066)
Average Income	1.309 *** (0.016)	Treatment Effect	0.122 (0.112)
Observations	214,501		
R²	24.82%		

Robust standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.2.5 - Robustness Check: Placebo Test with 2017 as the Treatment Year Regression
(all coefficients multiplied by 100)

Independent Variables			
Marital Status	Yes ***	Ruralization	-4.477 *** (0.063)
State	Yes ***	Work Factor	No
Parental Education	Yes ***	Dropout/Grade Repetition	No
Age	-0.035 *** (0.002)	Group Fixed Effect	-1.799 *** (0.048)
Gender (M=1)	3.27 *** (0.022)	Time Fixed Effect	3.543 *** (0.039)
Average Income	3.787 *** (0.053)	Treatment Effect	0.018 (0.048)
Observations	1,122,718		
R²	18.16%		

Robust standard errors in parenthesis
* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.2.6 - Robustness Check: Racial Component Model Encompassing 2020 to 2018
(all coefficients multiplied by 100)

Independent Variables			
Age	-0.043 *** (0.002)	Ruralization	-5.202 *** (0.057)
Gender (M=1)	3.533 *** (0.020)	Average Income	3.865 *** (0.035)
Marital Status	Yes ***	Group Fixed Effect	-2.128 *** (0.034)
State	Yes ***	Time Fixed Effect	1.526 *** (0.032)
Parental Education	Yes ***	Treatment Effect	1.061 *** (0.040)
R²	16.97%		

Robust standard errors in parenthesis
* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.2.7 - Robustness Check: Racial Component Model with Re-standardized Scores

Independent Variables			
Age	-0.432 *** (0.009)	Ruralization	-0.353 *** (0.005)
Gender (M=1)	-0.269 *** (0.002)	Work Factor	0.032 *** (0.002)
Marital Status	Yes ***	Dropout/Grade Repetition	-0.179 *** (0.002)
State	Yes ***	Group Fixed Effect	-0.120 *** (0.002)
Average Income	0.283 *** (0.001)	Time Fixed Effect	0.008 *** (0.002)
Parental Education	Yes ***	Treatment Effect	0.028 *** (0.003)
R²	18.40%		

Robust standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Note: Since the ENEM scores were re-standardized around zero, the treatment effect coefficient in this regression is not comparable to the one in Table 5 in level, only in terms of significance

Table A.2.8 - Robustness Check: Racial Component Model with Pre-Processed Data using EB (all coefficients multiplied by 100)

Independent Variables			
Age	-0.055 *** (0.002)	Ruralization	-4.582 *** (0.073)
Gender (M=1)	-3.810 *** (0.025)	Work Factor	0.827 *** (0.028)
Marital Status	Yes ***	Dropout/Grade Repetition	-2.559 *** (0.031)
State	Yes ***	Group Fixed Effect	-1.740 *** (0.039)
Average Income	5.163 *** (0.047)	Time Fixed Effect	0.605 *** (0.038)
Parental Education	Yes ***	Treatment Effect	0.786 *** (0.048)

Robust standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.2.9 - Robustness Check: 2PDD Logistic Regression – Output = High School Completion

Independent Variables			
State	Yes ***	Work Factor	0.077 *** (0.013)
Ruralization	Yes ***	Group Fixed Effect	-0.970 *** (0.021)
Per Capita Family Income	Yes ***	Time Fixed Effect	0.154 *** (0.017)
Age	0.236 *** (0.003)	Treatment Effect	0.118 *** (0.026)
Gender (F=1)	0.390 *** (0.013)		
Observations	205,285		
Nagelkerke R²	14.35%		

Standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.2.10 - Model with only Private School Students without Treatment Effect and without the Year 2010 (all coefficients multiplied by 100)

Independent Variables			
Marital Status	Yes ***	Ruralization	-8.116 *** (0.28)
State	Yes ***	Work Factor	-2.197 *** (0.075)
Parental Education	Yes ***	Dropout/Grade Repetition	-7.268 *** (0.104)
Age	0.026 ** (0.008)	Group Fixed Effect	-2.199 *** (0.111)
Gender (M=1)	2.817 *** (0.058)	Time Fixed Effect	-0.089 (0.091)
Average Income	1.456 *** (0.031)		
Observations	183,632		
R²	24.73%		

Robust standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

3. Income-based Scholarships and Access to Higher Education

3.1. Introduction

Income constraints and restrictions on access to credit lead to substantial entry barriers to higher education (HE) among disadvantaged groups. This dynamic not only widens educational inequalities but also hinders social mobility and contributes to the perpetuation of inequality in income levels (Lisboa and Menezes-Filho 2001; Barros et al. 2010). Therefore, in order to enhance the participation of underprivileged individuals in HE, a number of student aid programs have been implemented both in OECD and non-OECD countries, such as merit-based and income-based student loans, scholarships and maintenance grants.

Although a significant amount of evidence suggests that these programs have been effective in enabling access to HE and in mitigating educational inequalities, a share of the literature was unable to find statistically significant effects of financial aid policies on HE participation (Long 2004; Baumgartner and Steiner 2006; Tangkitvanich and Manasboonphempool 2010). Additionally, some studies have indicated that college enrolment might be more sensitive to long-run family and school factors than to short-term credit constraints (Cameron and Heckman 2001; Keane and Wolpin 2001; Carneiro and Heckman 2002). Therefore, understanding the extent to which student aid programs contribute to the increase in higher education enrolment among low-income individuals is imperative in order to unravel the effectiveness of such investments on promoting human capital accumulation.

Furthermore, while there has been rapidly accumulating evidence on the effects of these programs on developed economies (especially in the US), studies on non-OECD nations are still limited, an issue largely due to data availability restrictions in less-developed countries. However, understanding the role played by student aid policies in non-OECD economies is critical since barriers to higher education not only exacerbate educational and social inequalities but also generate important obstacles to economic development (Canton and Blom 2004).

In this study, I assess the effects of a student aid policy on access to HE in a non-OECD economy by exploiting the introduction of a public income-based scholarship program in Brazil. The inequality in access to higher education is large in this country: In 2004, according to PNAD (*Pesquisa Nacional por Amostra de Domicílio*), only 6% of individuals with a per capita family income equal to or below 3 minimum wages were enrolled in or had completed higher education, whereas the same was true for 46% of individuals whose per capita family income was of more than 3 minimum wages. In an attempt to alleviate this issue, in 2005 the Brazilian government created the Prouni, a federal program that grants full scholarships (covering 100% of tuition fees) to individuals attending private higher education institutions and whose monthly per capita family income amounts to at most 1.5 minimum wages, as well as partial scholarships (covering 50% or 25% of tuition fees) to those whose monthly per capita family income lies between 1.5 and 3 minimum wages.

In this research, I study the extent to which the Prouni contributed to the expansion of access to higher education in Brazil among low-income individuals. More precisely, I separately estimate the causal effects of the program on the higher education participation of individuals who were eligible for the full and partial scholarships. To this end, a difference-in-differences approach is employed in which I explicitly control for a set of individual-specific variables contained in the PNAD's database. To the best of my knowledge, this is the first research to evaluate the causal effects of an income-based scholarship program on participation in HE in a developing country.

Moreover, I also estimate the results by population subgroups – gender and race – in order to assess the existence of heterogenous effects, and a set of robustness exercises is conducted in order to further qualify the findings from this investigation - which include the estimation of the treatment effects using alternative strategies, namely, an instrumental variables methodology and a regression discontinuity design.

The results from this chapter indicate that, by the year 2007, the Prouni had increased the odds of an individual entitled to the full scholarship enrolling in higher education by 37%, an approximate impact of 1.4 percentage points on this group's higher education enrolment rate. As for the individuals entitled to the partial grant, the program increased the odds of attending

higher education by 20%, with an estimated impact of 3.4 percentage points on their higher education enrolment rate. Besides, the estimations suggest that every US$100 million spent with Prouni's tax waivers generated an approximate 0.5 percentage points increase in the HE enrolment rate of academic age individuals (or, equivalently, every US$1,000 per student increased this rate in 1.3 percentage points), and that the impacts of the Prouni on the students' higher education enrolment were greater for women and non-white individuals.

This chapter is organized as follows. Section 3.2 provides a brief literature review on the effects of student aid programs. Section 3.3 expands on the institutional setting of the Brazilian educational system and of the Prouni. Section 3.4 describes the data and the empirical strategy employed in the research. Section 3.5 presents the results of the models. Section 3.6 presents the robustness exercises. Section 3.7 discusses the main implications of the findings. Finally, section 3.8 concludes the chapter.

3.2. Literature

There is a substantial body of evidence suggesting that financial aid to college students is effective on enhancing access to HE. The vast majority of these studies have focused on US policies and programs (Dynarski 2000 and 2003; Cornwell et al. 2006; Kane 2003 and 2007; Abraham and Clark 2006; and Nguyen 2020), although a few of them have investigated these effects on other developed economies, such as Dearden et al. (2014) in the UK and Nielsen et al. (2010) in Denmark. In general, these empirical studies have found that a US$1,000 increase in grant aid generates an average increase of 3-5 percentage points in HE participation (Dearden et al. 2014).

The effectiveness of such programs on promoting HE enrolment, however, is not that trivial. Indeed, a handful of empirical investigations was unable to find statistically significant effects of student loans (Tangkitvanich and Manasboonphempool 2010), financial assistance schemes (Baumgartner and Steiner 2006) and tax credits (Long 2004) on HE enrolment rates. As stressed by Carneiro and Heckman (2002), there are two –not mutually exclusive– explanations for the gap in college attendance between individuals of different income classes: (i) credit constraint limiting the resources required to finance college education; and (ii) long run family and school factors crystallized in ability. It is therefore crucial to examine the extent to which

financial aid programs, which only aim to alleviate factor (i), are effective on promoting college enrolment.

Moreover, evidence of such effects in developing countries are still scarce. Solis (2017) and Canton and Blom (2004) investigate the impacts of limited access to credit on higher education enrolments by examining the implementation of student loan programs in Chile and Mexico, respectively, and both find evidence that the programs had strong positive effects on access to HE. Similarly, Gurgand et al. (2011) compare university enrolment rates in South Africa among students who were granted loans to cover registration fees and those who were not and conclude that credit constraints lead to a significant decrease in enrolments. These investigations, however, study the impacts of loan programs and credit on HE, and, as pointed out by Lepine (2018), it is not clear whether or not the findings from the abovementioned studies would generalize to the case of non-refundable aids.

The closest study to have investigated the impacts of an income-based scholarship program on access to HE in a developing country is Vélez et al. (2020), which examines the effects of the Ser Pilo Paga program in Colombia. In their study, the authors estimate that financial eligibility for the scholarship raised immediate enrollment by 56.5 to 86.5 percent, depending on the complier population. Nonetheless, there is a crucial difference between the Colombian program and the Prouni. The Ser Pilo Paga was not only an income-based, but also a merit-based program, as the scholarships were awarded only to the highest performers on the country's high school exit exam. As argued by Bernal and Penney (2019), the introduction of this program in Colombia not only enhanced access to HE, but also incentivized eligible students to improve their pre-college human capital accumulation – and the merit criteria of the program played a key role in that – which in turn might also have encouraged low-income individuals to enroll in HE. Therefore, the effects of an income-based scholarship program - that is, in which income is the sole criterion for scholarship eligibility - on a developing economy remains unexplored.

The establishment of the Prouni in 2005 in Brazil created an advantageous setup to expand the understanding of the effects of such programs on access to higher education in a non-OECD economy. The studies that were developed so far to evaluate the impacts of the Prouni, however, have focused on its' effects on students' higher education performance. Lepine (2018), for

instance, used a propensity score matching methodology to argue that students who receive the scholarship perform better in college and take less time to graduate, while Becker and Mendonça (2019) stated that the program positively impacted the Prouni beneficiaries' scores at the ENADE (the college-exit exam).

The effects of the Prouni on the participation of low-income individuals in HE has, to the best of my knowledge, not yet been assessed. Therefore, in this chapter I contribute to the literature by (i) estimating the causal effects of an income-based scholarship program on HE enrolment in a developing country, Brazil; (ii) providing further subsidies for the discussion on the effectiveness of student aid programs on access to HE; and (iii) comparing the effects of these programs between OECD and non-OECD economies.

3.3. Institutional background

This section describes the institutional background relevant to this research. Subsection 3.3.1 describes the structure of the higher education system in Brazil, while subsection 3.3.2 provides further information on the Prouni program.

3.3.1. Higher Education in Brazil

According to the 2019 Higher Education Census, the Brazilian Higher Education system serves 8.6 million students (in 2019 the average enrolment rate of individuals between 18 and 24 years old was 20.4%) and consists of 2,608 institutions, among which 2,306 (or 88%) are private and 302 (or 12%) are public. Private institutions, which are fee-paying, contain the vast majority of enrolments (6.5 million students in 2019, or nearly 76% of total enrollments). Public institutions, in turn, are predominantly free of charge[13] and are managed by either the federal, state or municipal government. Federal (110) and State (132) Higher Education Institutions (HEIs) encompass most of the public enrolments (62% and 32%, respectively), while Municipal institutions (60) contain only 6% of public enrolments.

A particularly relevant dysfunctionality of the Brazilian higher education system regards the inversion in the quality gap between private and public

---

[13] Institutions maintained by federal and state levels of governments are forbidden by law to charge tuition fees, but municipal institutions are allowed and usually charge some tuition.

institutions when evaluated in basic and higher education levels. While public HEIs are generally more prestigious and have the most competitive selection processes in the country[14], the opposite is true in primary and secondary school levels, where public institutions are less efficient (Sampaio and Guimaraes, 2009) and have been historically outperformed by private ones in standardized tests[15]. This dynamic amplifies the entry barriers to higher education among disadvantaged students and nourishes a cycle of inequality in higher education. Students from wealthier families who have access to higher-quality private basic education have better conditions to get into public, prestigious and tuition-free universities, while low-income individuals who had previously attended public basic schools often have no choice but to attend private and fee-paying HEIs.

Therefore, in order to democratize access to higher education in the country, a set of federal policies and programs were implemented in Brazil, especially from the beginning of the 21st century onwards (the Prouni, which shall be described in the next subsection, being among the most prominent ones). Indeed, the number of higher education enrolments has significantly risen in Brazil in recent years (Figure 3.1). From 2001 to 2019, total enrolments in private higher education institutions increased by 212%, whereas in public institutions, this number increased by 120%.

---

[14] Federal and State universities have higher average scores in the *Índice Geral de* Cursos (IGC), a quality index developed by the Ministy of Education, and comprise most of the higher ranked institutions in the *Ranking Universitário Folha* (RUF), an annual evaluation of the HEIs in Brazil developed by the *Folha de São Paulo* newspaper. According to Binelli et al. (2008), there were on average 9 applicants per seat at public institutions in 2003, while this ratio was of 1.5 in private institutions.

[15] Among the 50 highest ranked schools in the 2019 ENEM (Brazil's college-entrance exam), only 3 institutions were public (INEP).

Figure 3.1 - Number of Higher Education Enrolments in Brazilian Institutions (in thousands)
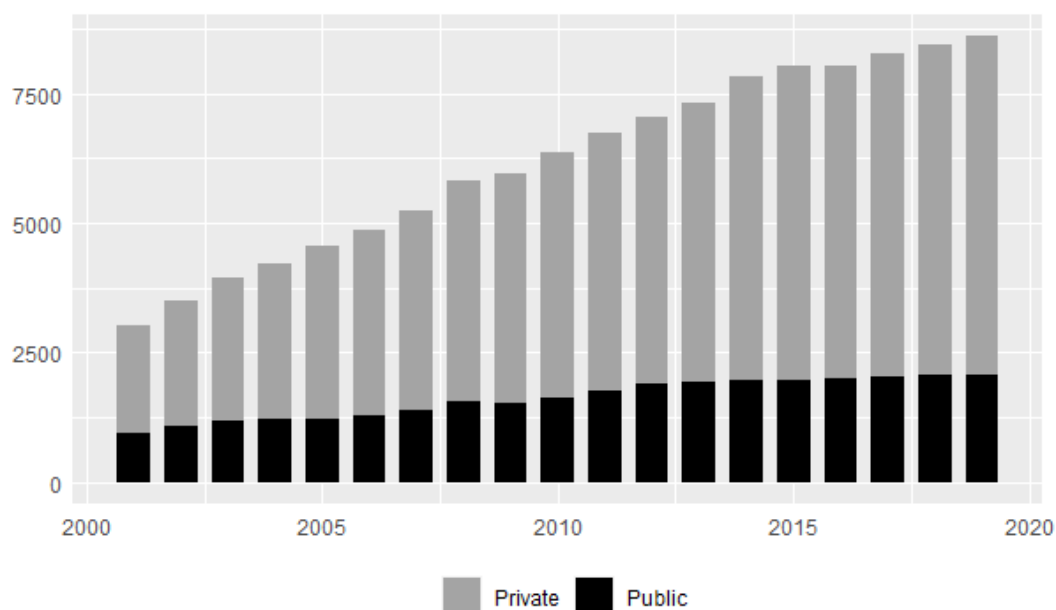
Source: Higher Education Census – INEP

3.3.2. The Prouni

The Prouni (*Programa Universidade para Todos*) is a federal scholarship program which was implemented by the Brazilian government in 2005 in an attempt to expand the enrolment of low-income young adults in higher education in the country. The program grants the students two different types of scholarships to private HEIs: a full scholarship (covering 100% of tuition fees, awarded to students whose monthly per capita family income amounts to at most 1.5 minimum wages) and a partial scholarship (covering 50% or 25% of tuition fees, awarded to those whose monthly per capita family income lies between 1.5 and 3 minimum wages). Additionally, to be eligible to the program, the student must meet at least one of the following criteria: (i) having attended high school at a public institution; (ii) having attended high school at a private institution with full scholarship; (iii) having a disability; or (iv) being an active professor at the public elementary or middle school network. Furthermore, a share of these grants is designated to non-

white students –classified into blacks, browns and indigenous-, according to the share of each race/ethnicity in each Brazilian state[16].

At the other end, participation by HEIs is voluntary and those joining the program agree to reserve a certain fraction of places to Prouni students in exchange for tax exemptions. From 2005 to 2015, a total of 1.75 million Prouni scholarships were awarded in the country. The evolution of the full and partial Prouni scholarships granted by the institutions is displayed in Figure 3.2. The figure shows an increasing number of granted Prouni scholarships, especially from 2008 onwards. Some remarks are in order regarding this dynamic. First, this trend does not seem to stem from an increasing number of available Prouni seats, since the number of granted scholarships was not restricted by any supply bottleneck over the analyzed period. From 2005 to 2015, the occupation percentage of Prouni reserved places was around 85% for full scholarships and 60% for partial ones (i.e. there were no queues in the access to scholarships). Second, Brazil's demographic trend rules out the possibility that the dynamic from Figure 3.2 derives from increasing cohorts. All the same, the overall number of HE candidates in Brazil increased by 127% from 2000 to 2011 (Neves, 2015), which indicates that the main driver behind the dynamic observed in Figure 2 might have been the increasing demand for higher education in the country throughout these years[17].

---

[16] For instance, in the state of Bahia, 76% of the population is non-white (either black, brown or indigenous). Therefore, the HEIs from Bahia that join the Prouni program must reserve 76% of scholarships to non-white persons.

[17] A detailed investigation on the evolution of the demand for higher education in Brazil and its' causes during 2000-2015 is presented in Neves (2015).

Figure 3.2 - Number of Prouni Scholarships granted by Institutions (in thousands)
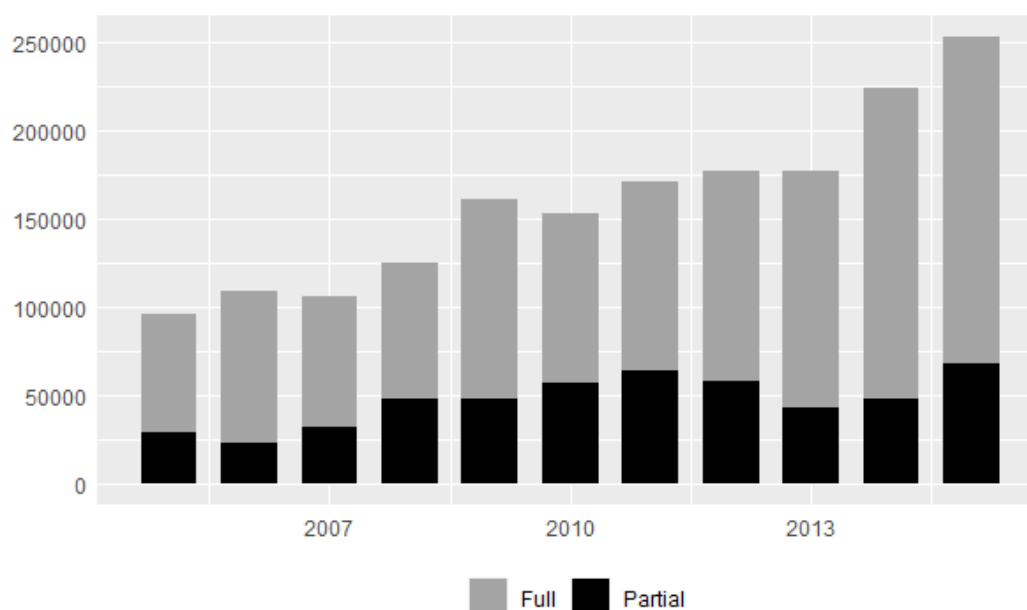
Source: Brazil's Ministry of Education

Students who meet the criteria and wish to apply for the Prouni program must go through an online centralized process, in which they are ranked according to their ENEM scores, and then select a set of desired HEIs as well as complete the specific selection process of each institution. Moreover, once a student is awarded the Prouni scholarship, they must pass at least 75% of their classes at the HEI in order to keep the funding. In 2008, the government implemented the "*Bolsa Complementar*", a different arrangement in which individuals that are eligible for the partial scholarship could receive a 25% Prouni grant, while the remaining 75% of the tuition fees would be covered by the FIES (*Fundo de Financiamento ao Estudante do Ensino Superior* - a federal student loan program). Since 2009, there is an extra requirement that candidates must fulfil in order to be eligible for the Prouni: they must score above a threshold in the ENEM. Anyway, this fixed threshold is relatively low and more than half the students taking the ENEM score above it (Lepine, 2018).

3.4. Data and methodology

3.4.1. Data

This research uses publicly available PNAD microdata (Brazil's national household sample survey), which was published in a yearly basis from 1967 to 2016 by the IBGE (*Instituto Brasileiro de Geografia e Estatística*)[18]. Created with the objective of providing subsidies to the study of Brazil's socioeconomic development, this repeated cross-sectional database provides information on housing, demography, migration, education, labor and income at both individual and household levels. The subjects of the survey are selected through a probabilistic household sample and information is collected by official IBGE's interviewers.

This study will focus on young individuals of academic age (17 to 24 years old) and on the years from 2001 to 2007. This timespan was selected so as to avoid the presence of concomitant educational policies that could pollute the findings, such as the creation of the FIES (*Fundo de Financiamento ao Estudanto do Ensino Superior)* in 1999 –a federal student loan program– and the SISU (*Sistema de Seleção Unificada*) in 2010 –an online platform that centralizes the admission processes to public universities-. From 2001 to 2007, the Prouni was the only major higher education program to be implemented in Brazil. The year 2007 was also strategically selected as the final year of the analysis since it does not contain the subsequent modifications on the Prouni's design (the implementation of the "*Bolsa Complementar*" and the ENEM threshold criterion). A concern that could naturally arise from this analysis is that a change in the FIES loan volume (the only program to have the same income threshold as the Prouni scholarship in the selected timespan) could bias its estimations. However, over the analyzed period, the amount of credit contracts executed within the FIES remained reasonably stable[19].

Table 3.1 and Table 3.2 provide the definitions and descriptive statistics of the variables from the PNAD database that were included in the models (to

---

[18] The PNAD survey was not carried out in the years the Census was conducted; and from 2016 onwards, the PNAD was replaced by its latest version, the *PNAD Contínua.*

[19] The number of yearly FIES contracts signed was also tested as an additional control variable (in the regressions to be presented in sections 2.4 and 2.5) but it did not exert significant changes in the estimated treatment coefficients.

be presented in section 3.5). Table 3.2 shows that, from 2001 to 2007, the participation in higher education in the data's sample (i.e., the percentage of individuals enrolled in higher education) increased from 7.7% to 11.5%. As can be seen in Table 3.2, the percentage of missing information in the database is considerably low (below 2%); these observations were dropped from the analysis.

Table 3.1 - Variables Description

Variables	Description
Age	Numerical (years)
Gender	Dummy. Woman = 1.
State	State of residence (27 federative units of Brazil)
Race/Ethnicity	White, black, brown, indigenous, or Asian
Work factor	Dummy = 1 if individual was not engaged in wage earning activity
Average income	Monthly per capita family income in minimum wages (US$ 200 in 2020)
Ruralization	Degree of ruralization of household's census area (8 categories)
HE participation	Dummy = 1 if individual was enrolled in HE

Source: PNAD (*Pesquisa Nacional por Amostra de Domicílio*)

Table 3.2 - Descriptive Statistics

Variables	2001	2002	2003	2004
Age - mean (sd)	20.4 (2.3)	20.4 (2.3)	20.4 (2.3)	20.4 (2.3)
missing	*0%*	*0%*	*0%*	*0%*
Gender (M; F)	49%; 51%	50%; 50%	50%; 50%	50%; 50%
missing	*0%*	*0%*	*0%*	*0%*
Race (white, brown, black)	47%; 47%; 6%	46%; 47%; 6%	46%; 47%; 6%	44%; 49%; 6%
missing	*0%*	*0%*	*0%*	*0%*
Work factor	54% Y; 46% N	55% Y; 45% N	54% Y; 46% N	56% Y; 44% N
missing	*0%*	*0%*	*0%*	*0%*
Average income - mean (sd)	1.5 mw (2.4)	1.5 mw (2.4)	1.3 mw (2.0)	1.3 mw (2.4)
missing	*2.9%*	*2.8%*	*2.8%*	*2.8%*
HE participation	7.7%	8.7%	9.2%	9.2%
missing	*1.9%*	*1.7%*	*1.5%*	*1.5%*
Sample size	56,968	57,929	57,676	59,104

Variables	2005	2006	2007	2001-2007
Age - mean (sd)	20.5 (2.3)	20.5 (2.3)	20.5 (2.3)	20.4 (2.3)
missing	*0%*	*0%*	*0%*	*0%*
Gender (M; F)	50%; 50%	50%; 50%	50%; 50%	50%; 50%
missing	*0%*	*0%*	*0%*	*0%*
Race (white, brown, black)	43%; 49%; 7%	43%; 49%; 7%	42%; 49%; 8%	45%; 48%; 7%
missing	*0%*	*0%*	*0%*	*0%*
Work factor	56% Y; 44% N	56% Y; 44% N	57% Y; 43% N	55% Y; 45% N
missing	*0%*	*0%*	*0%*	*0%*
Average income - mean (sd)	1.3 mw (1.9)	1.2 mw (1.9)	1.2 mw (1.9)	1.3 mw (2.1)
missing	*2.3%*	*2.5%*	*2.9%*	*2.7%*
HE participation	9.8%	10.8%	11.5%	9.6%
missing	*1.3%*	*1.0%*	*1.0%*	*1.4%*
Sample size	60,702	59,786	56,368	408,533

Source: PNAD (*Pesquisa Nacional por Amostra de Domicílio*)

3.4.2. Methodology

In order to study the causal effects of the implementation of the Prouni on individuals' HE enrolment, a difference-in-differences methodology is employed. The idea behind this approach is fairly simple. Outcomes are observed before and after a specific treatment and between two groups, a treatment group that was exposed to the treatment and a control group that

was not exposed to it. The treatment effect is then estimated by comparing the change in outcome between the two groups, while a set of control variables is added to the model in order to control for individuals' specific characteristics. Since the Prouni was applicable only to individuals below a certain income threshold, it was possible to construct two groups that are substantially similar to each other with the crucial difference that the treatment group is entitled to the Prouni scholarship while the control group is unaffected by it.

More precisely, I estimate two independent difference-in-differences regressions in order to separately assess the impacts of the full and the partial Prouni scholarships on HE enrolment. In the partial Prouni scholarship model, the HE enrolments are observed before and after the Prouni's implementation in 2005 and between two groups: a treatment group composed of individuals whose monthly per capita family income lies between 1.5 and 3 minimum wages -hence, eligible for the partial scholarship- and a control group composed of individuals whose monthly per capita family income lies between 3 and 4 minimum wages –that is, individuals that belong to a slightly greater income class although not eligible for any scholarships at all-. In the full Prouni scholarship model, the HE enrolments are again observed before and after the Prouni's implementation in 2005 and using the same control group. However, in this case, the treatment group is composed of individuals whose monthly per capita family income amounts to at most 1.5 minimum wages –that is, those eligible for the full scholarship-. Table 3.3 summarizes the grouping of the models.

Table 3.3 - Model's grouping summary

| | Model | |
	Partial Prouni Scholarship	Full Prouni Scholarship
Treatment Group	$1.5 <$ p.c. family income ≤ 3	p.c. family income ≤ 1.5
Control Group	$3 <$ p.c. family income ≤ 4	$3 <$ p.c. family income ≤ 4

It should be noted that per capita family income is being used as the sole criterion for scholarship eligibility when building the treatment groups. Nonetheless, as stated in subsection 3.3.2, in order to be eligible for the scholarship the student must also meet at least one of the following criteria:

(i) having attended high school at a public institution; (ii) having attended high school at a private institution with full scholarship; (iii) having a disability; or (iv) being an active professor at the public elementary or middle school network. These additional criteria are, however, barely restrictive - especially due to criterium (i)-. From 2005 to 2007 (the treatment period in the models), 91% of high school students with at most 3 minimum wages of per capita family income indeed attended public institutions.

Prior to estimating the treatment effects of the models, I first address the validity of the parallel trends' assumption. The previous trends for the control and treatment groups in the partial Prouni scholarship model are presented in Figure 3.3, while these trends for the full Prouni scholarship model are presented in Figure 3.4.

Figure 3.3 - Partial Prouni Scholarship HE Enrolment Evolution
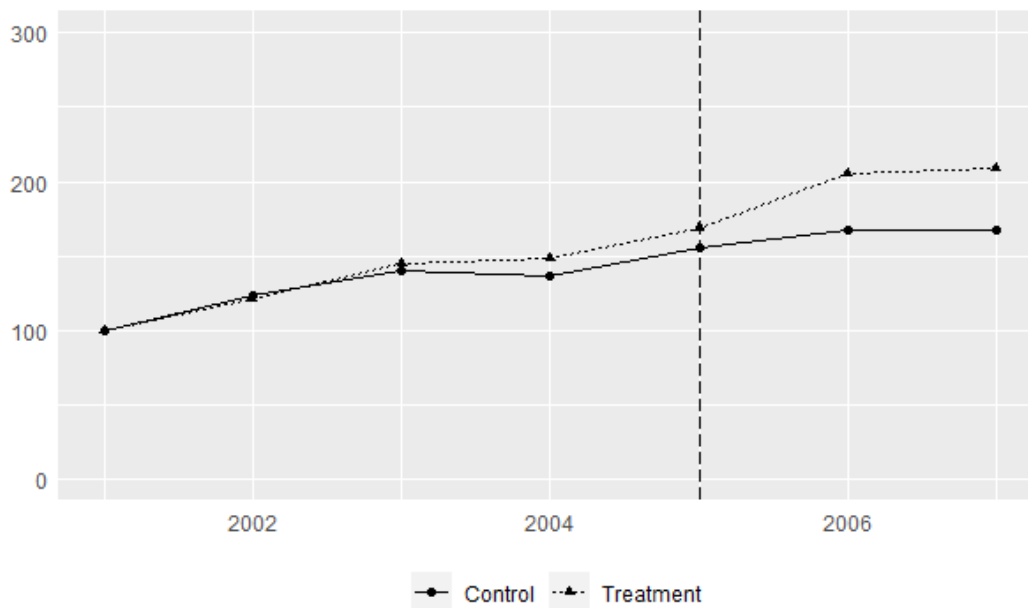
Figure 3.4 - Full Prouni Scholarship HE Enrolment Evolution
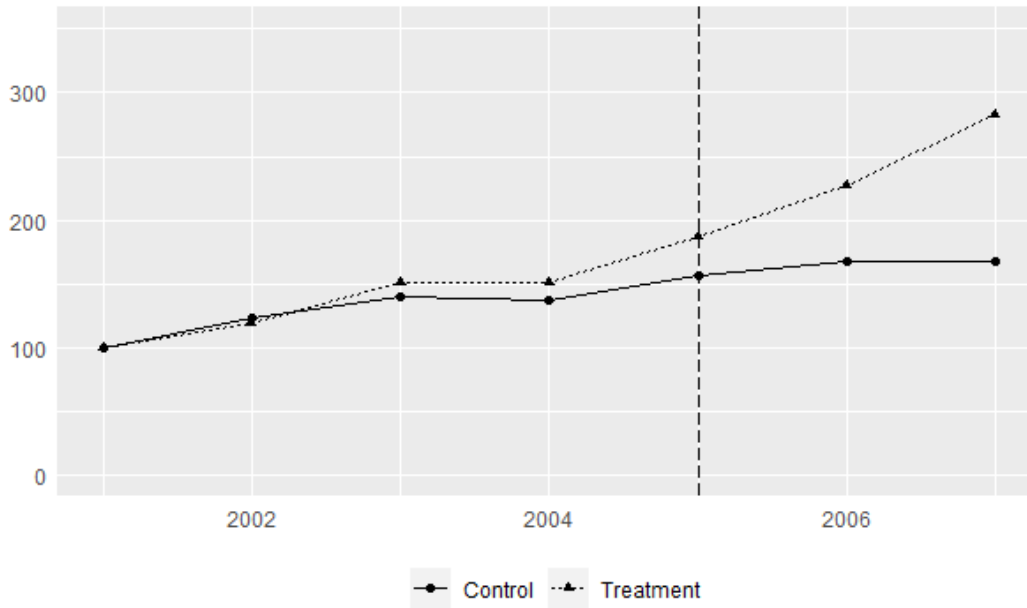
While a simple visual inspection of the figures shows that the pre-treatment trends seem to be relatively similar between groups, I test for this hypothesis by estimating the following dynamic event study logistic regression:

$$HE_{it} = c_0 + \phi_t + \lambda D_i^t + \gamma X_{it} + \sum_{s \neq 2004} \beta_s \times \mathbb{1}[t = s] \times D_i^t$$
$$+ \varepsilon_{it} \tag{3.1}$$

Where HE_{it} is a dummy variable that equals one if the individual i is enrolled in higher education in year t, D_i^t is a dummy variable that equals one if they belong to the treatment group; $\hat{\phi}$ and $\hat{\lambda}$ measure the time-specific and group-specific fixed effects, respectively; X_{it} includes the individual-specific control variables described in Table 3.2; and the coefficients $\{\hat{\beta}\}$ account for the event-study coefficients (which measure the causal effect of the treatment plus the difference in trends between the treatment and control groups), where 2004 is taken as the reference period (whence $\hat{\beta}_{2004}$ is normalized to zero). Therefore, the validity of the parallel trends assumption is tested by examining the significance of the pre-treatment beta coefficients ($\hat{\beta}_{2001}$, $\hat{\beta}_{2002}$ and $\hat{\beta}_{2003}$).

73

After performing the abovementioned examination, I start the analysis by estimating a standard two-periods model (2PDD), in which the model's timespan is divided into a pre-treatment (2001-2004) and post-treatment period (2005-2007). Therefore, in this setting, I evaluate the average effect of the Prouni on the HE enrolment of the treatment group during the entire post-treatment period. For the 2PDD, the following logistic regression is estimated:

$$HE_{it} = c_0 + \Phi W + \lambda D_i^t + \gamma X_{it} + \beta D_i^t W + \varepsilon_{it} \qquad (3.2)$$

where W is introduced, which is a dummy variable that equals one if $t \geq 2005$, that is, if it belongs to the post-treatment period, and the remaining variables are the same from equation 3.1. Moreover, for further reference, I shall refer to the interaction between D_i^t and W ($D_i^t W$) as the treatment dummy, that is, a dummy variable that equals one if the individual belongs to the treatment group and if they are observed after the treatment.

Next, I use the same two-periods difference-in-differences design to investigate if there were any heterogenous effects of the Prouni by gender and race. For this purpose, the interactions by the heterogeneity dimensions of interest are added to equation 3.2. More precisely, the following two equations are estimated:

$$HE_{it} = c_0 + \Phi W + \lambda D_i^t + \gamma X_{it} + \beta D_i^t W + \beta_g D_i^g W + \varepsilon_{it} \qquad (3.3)$$

$$HE_{it} = c_0 + \Phi W + \lambda D_i^t + \gamma X_{it} + \beta D_i^t W + \beta_r D_i^r W + \varepsilon_{it} \qquad (3.4)$$

where D_i^g is a dummy variable that equals one if the individual is female; D_i^r is a dummy variable that equals one if the individual is non-white; and β_g and β_r measure the incremental treatment effect for women and non-whites, respectively.

Lastly, I take a step further and estimate a dynamic event-study regression – for the entire population – in order to explore the effects of the Prouni on HE enrolment at each specific year (from 2005 to 2007). The analysis finishes with a battery of robustness exercises to scaffold the validity of the results obtained - more specifically, a test for checking the existence of anticipatory

effects, a placebo test, and revaluations of the estimations making use of an instrumental variables methodology, a regression discontinuity design and a pre-processes database using Entropy Balancing.

3.5. Results

3.5.1. Partial Prouni Scholarship Model

In this subsection, the results for the partial Prouni scholarship model are presented, that is, the impact of the Prouni on individuals whose monthly per capita family income lies between 1.5 and 3 minimum wages – i.e., eligible for the partial scholarship -. I start this subsection by presenting the results from equation 3.1, so as to assess the validity of the parallel trends' assumption.

The first column in Table 3.4 presents the estimated pre-treatment beta coefficients from equation 3.1 ($\hat{\beta}_{2001}$, $\hat{\beta}_{2002}$ and $\hat{\beta}_{2003}$) in a setting without control variables, while the second column presents the same coefficients in a model with controls. As shown in the table, all coefficients are statistically insignificant (all and each of the p-values are greater than 28%), hence providing further evidence that the treatment and control groups indeed share common trends prior to the Prouni's implementation.

Table 3.4 - Partial Prouni Scholarship Model: Pre-treatment Dynamic Event Study Coefficients

Event Study Coefficients	(1)	(2)
Beta 2001	-0.020 (0.084)	-0.039 (0.086)
Beta 2002	-0.089 (0.083)	-0.081 (0.085)
Beta 2003	-0.067 (0.081)	0.000 (0.085)
Control	No	Yes

Standard errors in parenthesis
* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Next, I present the results of the 2PDD analysis. Column 1 from Table 3.5 contains the results of the estimation for the entire population (equation 3.2), in which a significant treatment effect coefficient of 0.11 is estimated (p-

value of 1.5%), entailing an increase of 11.8% on the odds of attending HE. A detailed regression output with the controls' coefficients is displayed in Table A.3.1 in the appendix. Column 2 from the same table presents the incremental treatment effects by subgroups (equations 3.3 and 3.4). A positive although insignificant coefficient (p-value of 14.9%) is found for the interaction between the gender dummy and the treatment dummy, that is, there is not sufficient evidence to believe that the Prouni partial scholarship exerted different impacts by gender. For the interaction between the race and the treatment dummies, a positive and significant coefficient is estimated, which suggests that the Prouni partial scholarship exerted a greater impact on non-white individuals than it did on white persons. Table A.3.2 and Table A.3.3 in the appendix display the entire set of estimated coefficients from these regressions.

Table 3.5 - Partial Prouni Scholarship Model Two-Periods Regression: Total Population and Incremental Treatment Effect by Gender and Race

Total Population	(1)	Incremental Effect by Subgroup	(2)
Group Fixed Effect	0.067 (0.042)	Gender Coefficient	0.056 (0.038)
Time Fixed Effect	0.453 *** (0.041)	Gender in Odds Ratio	-
Treatment Effect Coefficient	0.112 * (0.046)	Racial Coefficient	0.108 ** (0.041)
Treatment Effect in Odds Ratio	11.80%	Racial in Odds Ratio	11.40%
Control	Yes		
Observations	73,247		
Nagelkerke R²	12.1%		

Column (1) presents the results from equation 2, while column (2) presents the estimated coefficients for the interactions from equations 3 and 4
Standard errors in parenthesis
* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Finally, I estimate a dynamic event-study regression for the entire population. The dynamic treatment effect coefficients are presented in Table 3.6, which indicates that the effect of the partial Prouni scholarship on the HE participation of individuals with per capita family income between 1.5 and 3 minimum wages increased throughout the first three years after its

implementation (indeed, $\hat{\beta}_{2005}$ is not statistically significant, i.e., it was not possible to detect any treatment effect in 2005). The entire set of estimated coefficients from the regression presented in Table 3.6 is displayed in Table A.3.6 in the Appendix.

Table 3.6 - Partial Prouni Scholarship Model: Dynamic Event Study Regression

Independent Variables	
Group Fixed Effect	0.080 * (0.040)
2005 Treatment Effect Coefficient	0.044 (0.068)
2005 Treatment Effect in Odds Ratio	-
2006 Treatment Effect Coefficient	0.134 * (0.066)
2006 Treatment Effect in Odds Ratio	14.3%
2007 Treatment Effect Coefficient	0.179 *** (0.068)
2007 Treatment Effect in Odds Ratio	19.6%
Time Fixed Effect	Yes
Control Variables	Yes
Observations	73,247
Nagelkerke R²	12.8%

Standard errors in parenthesis
* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

3.5.2. Full Prouni Scholarship Model

In this subsection, the results of the full Prouni scholarship model are presented, that is, the impacts of the Prouni on individuals whose monthly per capita family income amounts to at most 1.5 minimum wages – i.e., eligible for the full scholarship. For ease of exposition, the results are presented in an identical structure as in subsection 3.5.1, beginning with the examination of the parallel trends' assumption.

Table 3.7 presents the estimated pre-treatment beta coefficients ($\hat{\beta}_{2001}$, $\hat{\beta}_{2002}$ and $\hat{\beta}_{2003}$) from equation 3.1 in this case. The first column from this table displays these coefficients in a setting without control variables, while the

second column presents them in a model with controls. All coefficients are again not statistically significant (all and each of the p-values above 32%), hence providing further evidence that the treatment and control groups seem to share common trends prior to the Prouni's implementation in the full scholarship model as well.

Table 3.7 - Full Prouni Scholarship Model: Pre-treatment Dynamic Event Study Coefficients

Event Study Coefficients	(1)	(2)
Beta 2001	0.033	0.033
	(0.086)	(0.090)
Beta 2002	-0.083	-0.041
	(0.085)	(0.089)
Beta 2003	-0.038	0.000
	(0.082)	(0.086)
Control	No	Yes

Standard errors in parenthesis
* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Analogously to subsection 3.5.1, I first present the results of the 2PDD design estimations. Column 1 from Table 3.8 contains the results of the estimation for the entire population (Equation 3.2), while Column 2 presents the incremental treatment effects by subgroups. A significant treatment effect coefficient of 0.17 is estimated for the entire population, which is 54% greater than the treatment effect in the partial Prouni scholarship model and which entails an increase of 18.9% on the odds of enrolling in HE. Moreover, in this case, positive and significant coefficients for both the interaction between the gender dummy and the treatment dummy (p-value lower than 0.01) and for the interaction between the race and treatment dummies (p-value lower than 0.01%) are found. Table A.3.5, Table A.3.6 and Table A.3.7 in the appendix display the entire set of estimated coefficients from these regressions.

Table 3.8 - Full Prouni Scholarship Model Two-Periods Regression: Total Population and Incremental Treatment Effect by Gender and Race

Total Population	(1)	Incremental Effect by Subgroup	(2)
Group Fixed Effect	2.203 *** (0.072)	Gender Coefficient	0.166 *** (0.038)
Time Fixed Effect	0.479 *** (0.042)	Gender in Odds Ratio	18.1%
Treatment Effect Coefficient	0.173 *** (0.047)	Race Coefficient	0.195 *** (0.042)
Treatment Effect in Odds Ratio	18.9%	Race in Odds Ratio	21.5%
Control	Yes		
Observations	73,247		
Nagelkerke R²	12.1%		

Column (1) presents the results from equation 2, while column (2) presents the estimated coefficients for the interactions from equations 3 and 4
Standard errors in parenthesis
* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Finally, I also investigate the evolution of the treatment effect for the entire population across the post-treatment period. The dynamic treatment effect coefficients in this model are presented in Table 3.9 which indicates that the effect of the full Prouni scholarship on the HE participation of individuals with a per capita family income of at most 1.5 minimum wages also increased throughout the first three years after its implementation. Equally to the partial Prouni scholarship model, the $\hat{\beta}_{2005}$ coefficient is not statistically significant in this case (i.e., it was not possible to detect any treatment effect in 2005). Additionally, the $\hat{\beta}_{2006}$ coefficient is significant only if a 10% significance level (p-value of 7.2%) is assumed, hence the evidence as for the effects of the full Prouni scholarship on HE enrolment in 2006 are tenuous. The entire set of estimated coefficients is displayed in Table A.3.8 in the appendix.

Table 3.9 - Full Prouni Scholarship Model: Dynamic Event Study Regression

Independent Variables	
Group Fixed Effect	2.236 *** (0.071)
2005 Treatment Effect Coefficient	0.082 (0.069)
2005 Treatment Effect in Odds Ratio	-
2006 Treatment Effect Coefficient	0.119 ' (0.066)
2006 Treatment Effect in Odds Ratio	12.6%
2007Treatment Effect Coefficient	0.314 *** (0.068)
2007 Treatment Effect in Odds Ratio	36.9%
Time Fixed Effect	Yes
Control Variables	Yes
Observations	320,776
Nagelkerke R²	23.8%

Standard errors in parenthesis
 ' Significance at 10% level; * Significance at 5% level; ** Significance at 1% level;
*** Significance at 0.1% level

3.6. Robustness

The results so far suggest that the individuals that were eligible for the Prouni (both for the partial and for the full scholarships) increased their participation in HE after the program's implementation by significantly more than those who were not eligible for the scholarship. In this subsection, four robustness exercises are conducted in order to further qualify these findings. First, a check for anticipatory effects of the Prouni; second, a placebo test, in which both the treatment and control groups are not eligible for the program – and, therefore, should be unaffected by it -; third, a re-estimation of the treatment effect coefficients employing an instrumental variables methodology; fourth, a re-estimation of the treatment effect coefficients employing a regression discontinuity framework; and fifth, an assessment of the robustness of the results to a pre-processed dataset using an Entropy Balancing methodology.

I start by checking for anticipatory effects of the treatment, that is, whether the Prouni had any impact on individuals' HE participation before it was implemented in 2005. First, it should be noted that it is unlikely that there were any anticipatory effects, since the law ("Lei nº 11.096/2005") that instituted the Prouni was published only on January 13th of 2005, hence it seems implausible that an individual would enroll at a higher education institution at least 6 months prior to the creation of the Prouni (in the second semester of 2004) only with the probability of receiving a scholarship out of a program that was still being discussed in the Congress. Nevertheless, I test for this hypothesis by estimating a dynamic difference-in-differences model, in which the years right before and right after the Prouni's implementation (2004 and 2005) are excluded from the regression. Table A.3.9 and Table A.3.10 in the appendix present the results of this design for the partial Prouni scholarship and full Prouni scholarship models, respectively. The estimated treatment effect coefficients remain significant and close to the ones estimated in subsections 3.5.1 (Table 3.5) and 3.5.2 (Table 3.8), which strengthens the hypothesis that there were indeed no anticipatory effects.

In the second robustness exercise, I perform a placebo test using only individuals that were not eligible for the scholarship. The concern here is that the increase in the HE participation of lower-income individuals could be driven by some other factor other than the Prouni, such as noisy data or any unobserved driver. In this exercise, the model's control group is composed of individuals whose monthly per capita family income lies between 5.5 and 7 minimum wages and the treatment group is composed of individuals with monthly per capita family income between 4 and 5.5 minimum wages. I then estimate a difference-in-differences design that is quite similar to the ones presented in subsections 3.5.1 and 3.5.2 – that is, in which the treatment group belongs to a slightly lower income class than the control group – with the crucial difference that both groups are not entitled to any Prouni scholarship. The results of this estimation are presented in Table A.3.11 in the Appendix. The treatment effect coefficient in this case is insignificant, as shown in the table (p-value of 57%), suggesting that the results obtained were not merely a placebo effect and that the Prouni did not exert any impact on higher-income individuals.

Third, I assess robustness of the results to a different estimation strategy. The fact that an individual's grant eligibility is a function of family income could

raise some endogeneity concerns – for instance, there could be an unobserved driver, such as ability or motivation, that impacts higher education attendance and is correlated to family income and hence to the scholarship eligibility dummy D^t-. So as to mitigate this potential bias, I follow a similar strategy to Dearden et al. (2014) and use the percentage of scholarship-eligible individuals by state and degree of ruralization cluster[20] as instrument for actual scholarship-eligibility. More precisely, I allocate the individuals from the database in 216 clusters (27 states times 8 degrees of ruralization) and for each cluster I calculate (i) the percentage of individuals entitled to the partial Prouni scholarship (i.e., individuals with a per capita family income between 1.5 and 3 minimum wages), which is used as an instrument for actual scholarship-eligibility in the partial Prouni model; and (ii) the percentage of individuals entitled to the full Prouni scholarship (i.e., individuals with a per capita family income of at most 1.5 minimum wages), which is used as an instrument for actual scholarship-eligibility in the full Prouni model. The results of the nonlinear two-stage estimations with control function (two-stage residual inclusion) are presented in Table A.3.12 and Table A.3.13 in the Appendix. The tables show that the IV methodology generates a treatment coefficient of 0.11 for the partial Prouni scholarship model and 0.18 for the full Prouni scholarship model, which are very close to the ones estimated in subsections 3.5.1 (Table 3.5) and 3.5.2 (Table 3.8).

Fourth, I evaluate the validity of the inferences using yet another estimation strategy, more precisely a regression discontinuity design (RDD), which also allows for the estimation of unbiased causal effects in the presence of unobserved confounding (Shadish, Cook and Campbell 2002). Since I am working with a binary outcome (whether or not the individual is enrolled in HE), the popular bandwidth procedure by Imbens and Kalyanaraman, (2012), which is developed for the local linear estimator becomes suboptimal (Xu, 2017), and I hence estimate the regression discontinuity treatment effect using a local logistic regression. The results of the estimation for the partial and full scholarship models are displayed in Figure A.3.1, Figure A.3.2 and in Table A.3.14 in the Appendix, which in both cases generate a significant treatment effect (although evidence is tenuous for the partial scholarship

---

[20] We rely on the identifying assumption that the geographical location (interaction between state and degree of ruralization of the individual's census-designated area) does not directly impacts HE enrolment

model, once the treatment effect coefficient is only significant at a 5% level). For further robustness' sake, a placebo RDD test is estimated, in which I use the same placebo treatment and control groups from the difference-in-differences placebo test – individuals with a per capita family income between 4 and 5.5 minimum wages, and 5.5 and 7 minimum wages, respectively. This exercise yields insignificant treatment effects, as shown in Table A.3.14.

Finally, I assess robustness of the models' results to a pre-processed and re-balanced database. Since the Prouni was not randomly assigned (i.e., applicable to individuals with different levels of income) the causal conclusions derived from such observational data might be somewhat polluted by covariate imbalance. It is worth noting, however, that the treatment and control groups in the original model were chosen so that they belong to the closest possible income groups, precisely in order to mitigate this imbalance. Anyway, I reweight the control groups in both the full and partial Prouni scholarship models using Entropy Balancing (Hainmueller 2012), a method which intends to match the covariate moments for the different experimental groups and is double robust with respect to linear outcome regressions (Zhao and Percival 2017). The results of these estimations are displayed in Table A.3.15 in the Appendix, and show that the treatment effect coefficients are again significant and similar, although slightly lower, to the ones presented in Table 3.5 and Table 3.8.

3.7. Discussion

The results obtained suggest that both the partial and the full Prouni scholarships had positive and significant effects on the higher education participation of individuals that were entitled to the program. The dynamic event study coefficients for both models show that the impacts of the Prouni on HE enrolment increased from 2005 to 2007. This is in fact a natural and expected result since there is a cumulative effect of the Prouni on the HE enrolment rate in its initial years of implementation[21]. For this reason, from this point forward I shall focus the discussion on the estimated treatment effects for the year 2007.

---

[21] For instance, in 2007, the HE enrolment rate of low-income individuals might be affected by those who received the scholarship in that year as well as those who had received it in the two previous years and were still attending college.

For the partial Prouni scholarship model, an average treatment coefficient of 0.18 is estimated. There are two main approaches to interpret this result. The first one, already presented in the output tables, is to convert this coefficient into odds ratio, which can be done by simply calculating e^{β}-1, where β stands for the treatment coefficient (in the analyzed case, this would yield $e^{0.18} - 1 =$19.6%). To put into words, by 2007 the partial Prouni scholarship had increased the odds of eligible individuals attending higher education by 19.6%. An alternative mean to interpret this coefficient is through its impact on HE enrolment rates. The HE enrolment rate of the treatment group in this model (population between 17 and 24 years old and with a per capita family income between 1.5 and 3 minimum wages) in 2007 was 26.7%. Using the logistic transformation[22] and some simple algebraic manipulations, it is possible to infer that, had the Prouni not been implemented, this percentage (i.e., the counterfactual) would have been around 23.3%. This implies that the partial Prouni scholarship increased this group's HE enrolment rate in approximately 3.4 percentage points.

A similar analysis can be performed for the full Prouni scholarship model. In this case, a treatment coefficient of 0.31 is estimated for 2007, meaning that, by that year, the Prouni had increased the odds of eligible individuals attending higher education by $e^{0.31} - 1 =$36.9% - as expected, a greater effect than in the partial Prouni model. Furthermore, the HE enrolment rate of this model's treatment group (individuals with age from 17 to 24 and per capita family income of at most 1.5 minimum wages) in 2007 was of 5.4%, whereas, had the Prouni not been implemented, this percentage would have been around 4.0% - an approximate effect of 1.4 percentage points on this group's HE enrolment rate (which is lower than the one estimated for the partial Prouni model due to a lower baseline rate).

Up to 2007, the government abstained from collecting approximately USD 300 million (in 2020 values) due to Prouni's tax exemptions, with an approximate annual cost per student of USD 621 (Ministry of Education and Federal Revenue). Meanwhile, extrapolating the results from the paragraphs above to the entire set of academic age individuals (that is, accounting for all income classes), yields an impact of the program on the HE enrolment rate

---

[22] A logistic regression expressed by $\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ can be rewritten as $p = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2)}}$.

of the overall academic age population of 1.6 percentage points (an average of 0.8 p.p. per year in 2006 and 2007 – recall that no effect was found in 2005). This implies that every USD 100 million spent by the government with tax waiver from the Prouni generates an approximate 0.5 percentage points increase in the HE enrolment rate of these individuals (or, equivalently, that every USD 1,000 spent per student generates an approximate 1.3 percentage points increase in this rate).

The empirical evidence in developed economies (namely, in the US and in the UK) find increases in participation of 3-5 percentage points per $1,000 of student aid (Dearden et al. 2014). Since Brazil has a much lower baseline HE participation rate than these countries, comparing these policies in terms of impacts in percentage points is not the fairest indicator of efficiency. Instead, I compare them in percentage terms. The US and the UK had a HE participation rate of approximately 40% in 2007 (UK Department of Education and US National Center for Education Statistics), hence the 3-5 percentage points impact per US$1,000 of student aid entails a percentage increase in the HE enrolment rate that ranges from 7.5% to 13%. Since the HE participation rate of academic age individuals in Brazil in 2005-2007 was approximately 10-11%, the estimated impact of 1.3 p.p. per year results in a percentage increase in the HE enrolment rate of roughly 11.5% to 12% - i.e., in line with the international cases.

Moreover, besides estimating the effects of the program for the entire population, I have also tested for the presence of heterogenous effects of the Prouni by population subgroups - race and gender -, the main findings being: (i) the program seems to have exerted a greater impact on women than it did on men - although this heterogeneity could only be detected in the full scholarship model; and (ii) the program seems to have exerted a greater impact on non-white individuals than it did on whites, a result that was obtained in both the partial and full scholarships models. There are a few possible explanations for these results. First, non-white and female individuals might respond more strongly to such financial aid policies. In order to verify this hypothesis, I estimate a logit model with the entire pre-treatment population (PNAD data), in which the HE attendance is the dependent variable, and find significant and positive coefficients for the interactions between gender and income, and also race and income (Table A.3.16 in the Appendix). That is, non-white and female individuals' HE

attendance seem to be more sensitive to income restrictions. A second and perhaps complementary explanation for finding (ii) is that this dynamic might stem from the rules of the Prouni, given that, by law, a share of the grants must be designated to blacks, browns and indigenous individuals.

A possible concern that could arise from the estimations from this chapter regards the suitability of the control group. Since private institutions that joined the program agreed to reserve a certain fraction of seats to Prouni students, it would be plausible to assume that these reserved seats could have increased competition for places in private HEIs among higher-income individuals, and hence could have affected their higher education enrolment as well. This scenario, however, is unlikely once the overall number of seats in private HEIs increased, on average, 7.4% per year from 2005 to 2007 (in fact, the number of new seats surpassed the number of granted Prouni scholarships in the period by 55%).

Finally, whilst the estimated effects of the Prouni were indeed sizeable and contributed to narrow the gap in HE enrolment between individuals from different income classes in Brazil, it is worth underlining that this gap remains substantial still at the time of this study –more than 15 years after the first scholarships were awarded-. Additionally, let us recall that the number of granted scholarships throughout the years was not restricted by any supply bottleneck, since the amount of offered Prouni seats by the institutions outnumbered the amount of granted scholarships in each and every year since its conception. Therefore, it seems likely that the persistent inequality in access to HE in Brazil is affected by two additional -and correlated- fundamental factors: (i) credit and income constraints that affect academic performance and educational attainment since pre-schooling; and (ii) long-run family and school environmental factors that shape young students' abilities and motivations. In this sense, policy makers that are willing to reduce the inequality of access to HE should give due weight not only to financial aid policies during HE, but also to programs that could mitigate inequalities since the early stages of the educational system. Notable Brazilian programs that move in this direction are the *Bolsa Família* (descending from the former *Bolsa Escola*), which provides financial assistance to poor families in the country conditional on children and teenagers between six and seventeen years old having a minimum school attendance; and the *Brasil Carinhoso*, a cash transfer program entitled to

families with children up to four years old, which aims to help them finance early childhood education and health care. Osorio and Souza (2012), Soares et al. (2010), and Bourguignon et al. (2003) have provided evidence of the effectiveness of these policies.

3.8. Conclusion

Understanding the extent to which financial aid to college students enhances access to higher education is crucial in order to unravel the effectiveness of such policies on promoting human capital accumulation. Furthermore, although there is a significant amount of evidence pointing towards a positive effect of financial aid on college enrolment, whether or not these policies are actually effective on bolstering access to HE is still a matter of controversy – especially in emerging economies where these sorts of empirical investigations are much more limited. In this research, I contribute to the literature on the effects on non-refundable aids on HE participation in a developing country by exploiting the implementation of Brazil's Prouni.

The Prouni, which was introduced in 2005, grants full and partial scholarships to students from low-income families attending private higher education institutions in the country. I find evidence that the Prouni had a positive and significant effect on the HE participation rate of those who were eligible for the program, increasing the odds of attending HE by 20% and 37% for those entitled to the partial and full scholarships, respectively, by 2007 – which, in turn, entailed an increase in the HE enrolment rate of these individuals of 3.4 and 1.4 percentage points. This study estimates that every USD 100 million spent by the government with tax waiver from the Prouni generated an approximate 0.5 percentage points increase in the HE enrolment rate of academic age individuals (every USD 1,000 per student generated an approximate 1.3 percentage points increase in this rate). Although these impacts seem low when compared to studies from OECD countries, this is largely due to Brazil's low baseline HE participation rate. Put differently, every USD 1,000 per student spent by the Prouni increased the HE participation rate by approximately 11% to 12%, which is in line with the findings from developed economies.

Furthermore, I have also tested for the presence of heterogenous effects of the Prouni across a set of different dimensions – race and gender. Albeit no statistically significant heterogenous effect by gender was found in the partial

scholarship model, the Prouni seemed to have exerted a greater impact on the HE participation of non-white persons (a result found in both models) and women (in the full scholarship model).

Although a battery of robustness exercises strengthens the validity of this research's claims, I acknowledge some limitations in the employed strategy. First, I have controlled for a set of observable individual and socioeconomic characteristics, others remaining as non-observable. Second, since several educational policies were implemented shortly before and after the Prouni, the timespan of the model had to be limited to the first three years of the program. Third, since the PNAD database does not disclose information on the type of high school institution previously attended by HE students (i.e. public or private), the income threshold had to be used as the sole criterium for scholarship eligibility in the Prouni. Nevertheless, several robustness exercises and alternative estimation strategies allow us to provide strong evidence that the Prouni implemented in Brazil indeed had a sizeable effect on the HE enrolment of students from low-income families – a result that is in line with the majority of the other international studies on the impacts of non-refundable aids on HE participation. Thus, this investigation stresses the importance of governments' and policymakers' commitment to financial aid policies that aim to reduce the entry barriers to higher education, especially in emerging economies where such barriers not only amplify educational and social inequalities, but also hampers economic development.

Appendix

Figure A.3.1 - Full Prouni Scholarship RD: Local Logistic Regression
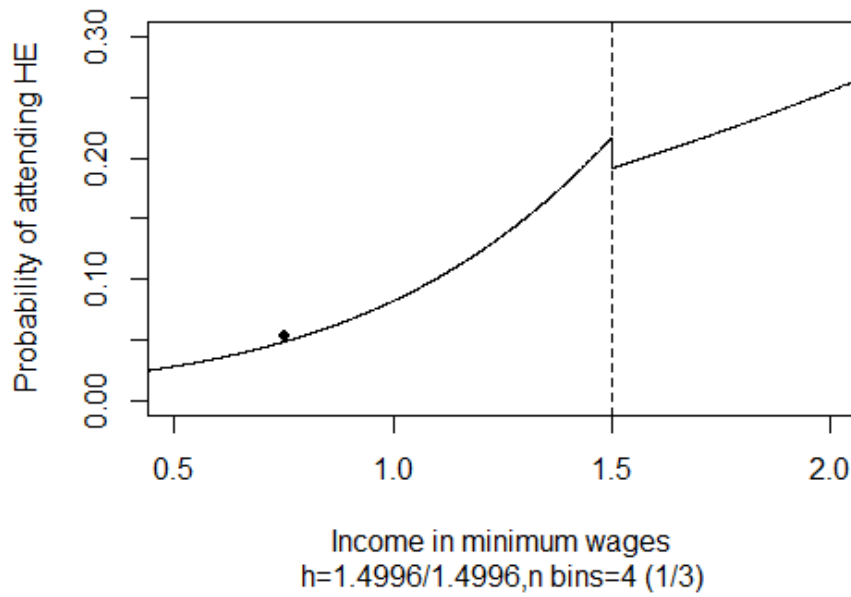
Figure A.3.2 - Partial Prouni Scholarship RD: Local Logistic Regression
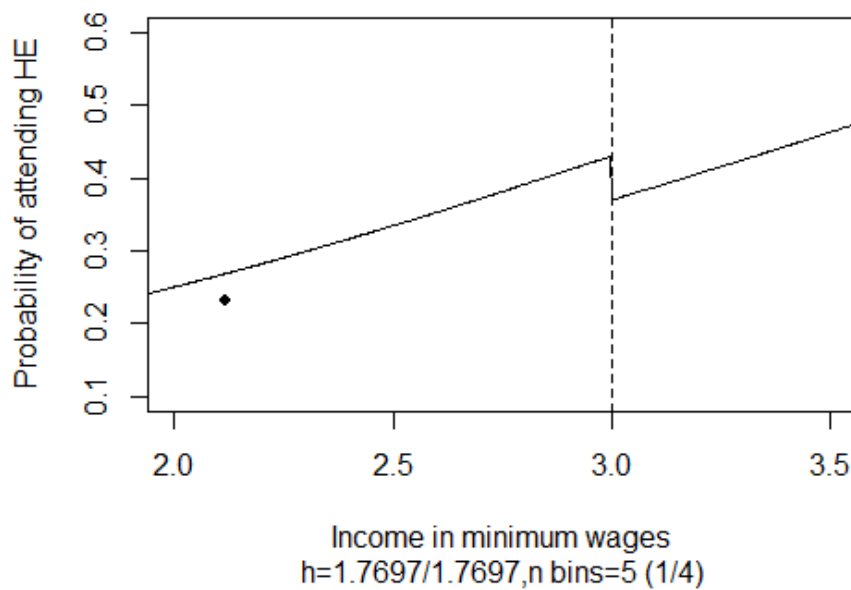
Table A.3.1 - Partial Prouni Scholarship Model: Two-Periods Regression

Independent Variables			
Ruralization	Yes ***	Average Income	0.712 *** (0.024)
State	Yes ***	Work Factor	0.456 *** (0.021)
Race	Yes ***	Group Fixed Effect	0.067 (0.042)
Age	0.125 *** (0.004)	Time Fixed Effect	0.453 *** (0.041)
Gender (F = 1)	0.393 *** (0.019)	Treatment Effect	0.112 * (0.046)
Observations	73,247		
Nagelkerke R²	12.1%		

Standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.3.2 - Partial Prouni Scholarship Model: Two-Periods Regression with Gender Interaction

Independent Variables			
Ruralization	Yes ***	Work Factor	0.456 *** (0.021)
State	Yes ***	Group Fixed Effect	0.068 (0.042)
Race	Yes ***	Time Fixed Effect	0.453 *** (0.041)
Age	0.126 *** (0.004)	Treatment Effect	0.083 (0.051)
Gender (F = 1)	0.372 *** (0.024)	Treatment * Gender	0.056 (0.038)
Average Income	0.712 *** (0.024)		
Observations	73,247		
Nagelkerke R²	12.1%		

Standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.3.3 - Partial Prouni Scholarship Model: Two-Periods Regression with Race Interaction

Independent Variables			
Ruralization	Yes ***	Work Factor	0.456 *** (0.021)
State	Yes ***	Group Fixed Effect	0.0703 ' (0.042)
Race	Yes ***	Time Fixed Effect	0.455 *** (0.041)
Age	0.125 *** (0.004)	Treatment Effect	0.073 ' (0.041)
Gender (F = 1)	0.372 *** (0.022)	Treatment * Race	0.108 ** (0.041)
Average Income	0.713 *** (0.024)		
Observations	73,247		
Nagelkerke R²	12.1%		

Standard errors in parenthesis

' Significance at 10% level; * Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.3.4 - Partial Prouni Scholarship Model: Dynamic Event Study Regression

Independent Variables			
Ruralization	Yes ***	Year Dummy 2001	-0.492 *** (0.033)
State	Yes ***	Year Dummy 2002	-0.255 *** (0.031)
Race	Yes ***	Year Dummy 2003	-0.047 (0.035)
Age	0.125 *** (0.004)	Year Dummy 2005	0.194 *** (0.030)
Gender (F = 1)	0.397 *** (0.019)	Year Dummy 2006	0.323 *** (0.061)
Average Income	0.725 *** (0.024)	Year Dummy 2007	0.331 *** (0.063)
Work Factor	0.458 *** (0.021)	Beta 2005	0.044 (0.068)
Group Fixed Effect	0.080 * (0.042)	Beta 2006	0.134 * (0.066)
		Beta 2007	0.179 *** (0.068)
Observations	73,247		
Nagelkerke R²	12.8%		

Standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.3.5 - Full Prouni Scholarship Model: Two-Periods Regression

Independent Variables			
Ruralization	Yes ***	Average Income	1.905 *** (0.027)
State	Yes ***	Work Factor	0.414 *** (0.020)
Race	Yes ***	Group Fixed Effect	2.203 *** (0.072)
Age	0.137 *** (0.004)	Time Fixed Effect	0.479 *** (0.042)
Gender (F = 1)	0.437 *** (0.019)	Treatment Effect	0.173 *** (0.047)
Observations	320,776		
Nagelkerke R²	23.4%		

Standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.3.6 - Full Prouni Scholarship Model: Two-Periods Regression with Gender Interaction

Independent Variables			
Ruralization	Yes ***	Work Factor	0.412 *** (0.020)
State	Yes ***	Group Fixed Effect	2.214 *** (0.072)
Race	Yes ***	Time Fixed Effect	0.477 *** (0.042)
Age	0.136 *** (0.004)	Treatment Effect	0.074 (0.052)
Gender (F = 1)	0.362 *** (0.026)	Treatment * Gender	0.166 *** (0.038)
Average Income	1.907 *** (0.027)		
Observations	320,776		
Nagelkerke R²	23.4%		

Standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.3.7 - Full Prouni Scholarship Model: Two-Periods Regression with Race Interaction

Independent Variables			
Ruralization	Yes ***	Work Factor	0.412 *** (0.020)
State	Yes ***	Group Fixed Effect	2.130 *** (0.073)
Race	Yes ***	Time Fixed Effect	0.468 *** (0.042)
Age	0.136 *** (0.004)	Treatment Effect	0.106 * (0.050)
Gender (F = 1)	0.435 *** (0.019)	Treatment * Race	0.195 *** (0.042)
Average Income	1.824 *** (0.029)		
Observations	320,776		
Nagelkerke R²	23.4%		

Standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.3.8 - Partial Prouni Scholarship Model: Dynamic Event Study Regression

Independent Variables			
Ruralization	Yes ***	Year Dummy 2001	-0.494 *** (0.035)
State	Yes ***	Year Dummy 2002	-0.303 *** (0.033)
Race	Yes ***	Year Dummy 2003	-0.022 (0.036)
Age	0.136 *** (0.004)	Year Dummy 2005	0.224 *** (0.030)
Gender (F = 1)	0.438 *** (0.019)	Year Dummy 2006	0.340 *** (0.062)
Average Income	1.920 *** (0.027)	Year Dummy 2007	0.369 *** (0.065)
Work Factor	0.422 *** (0.020)	Beta 2005	0.082 (0.069)
Group Fixed Effect	2.236 *** (0.071)	Beta 2006	0.119 ' (0.066)
		Beta 2007	0.314 *** (0.068)
Observations	320,776		
Nagelkerke R²	23.8%		

Standard errors in parenthesis

' Significance at 10% level; * Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.3.9 - Partial Prouni Scholarship Model: Check for Anticipatory Effects

Independent Variables			
Ruralization	Yes ***	Average Income	1.032 * (0.498)
State	Yes ***	Work Factor	0.422 *** (0.025)
Race	Yes ***	Group Fixed Effect	0.057 (0.049)
Age	0.126 *** (0.005)	Beta 2006	0.154 * (0.070)
Gender (F=1)	0.401 *** (0.023)	Beta 2007	0.199 ** (0.072)

Standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.3.10 - Full Prouni Scholarship Model: Check for Anticipatory Effects

Independent Variables			
Ruralization	Yes ***	Average Income	1.869 *** (0.032)
State	Yes ***	Work Factor	0.397 *** (0.024)
Race	Yes ***	Group Fixed Effect	2.097 *** (0.085)
Age	0.138 *** (0.005)	Beta 2006	0.137 * (0.070)
Gender (F=1)	0.442 *** (0.023)	Beta 2007	0.331 *** (0.072)

Standard errors in parenthesis
* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.3.11 - Placebo Exercise

Independent Variables			
Ruralization	Yes ***	Average Income	0.187 *** (0.038)
State	Yes ***	Work Factor	0.306 *** (0.037)
Race	Yes ***	Group Fixed Effect	0.091 (0.073)
Age	0.112 *** (0.008)	Time Fixed Effect	0.217 *** (0.059)
Gender (F=1)	0.224 *** (0.033)	Treatment Effect	0.041 (0.072)

Standard errors in parenthesis
* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.3.12 - Partial Prouni Scholarship IV Regression

Independent Variables	
First stage	
Percentage of scholarship-eligible individuals[a]	0.240 *** (0.040)
Second stage	
Treatment effect	0.111 * (0.047)

[a] Mean percentage by state*ruralization class

Standard errors in parenthesis

* Significance at 5% level

** Significance at 1% level

 *** Significance at 0.1% level

Table A.3.13 - Full Prouni Scholarship IV Regression

Independent Variables	
First stage	
Percentage of scholarship-eligible individuals[a]	0.537 *** (0.006)
Second stage	
Treatment effect	0.184 *** (0.049)

[a] Mean percentage by state*ruralization class

Standard errors in parenthesis

* Significance at 5% level

** Significance at 1% level

 *** Significance at 0.1% level

Table A.3.14 - RDD Treatment Effect Estimations

Model	Treatment effect estimate
Full Scholarship	0.258 *** (0.072)
Partial Scholarship	0.253 * (0.125)
Placebo Exercise	0.002 (0.289)

Standard errors in parenthesis

* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.3.15 - Partial and Full Prouni Scholarship Model Two-Periods Regression with Entropy Balancing

Independent Variables	Partial	Full
Group Fixed Effect	0.076 (0.042)	2.210 *** (0.075)
Time Fixed Effect	0.468 *** (0.041)	0.492 *** (0.044)
Treatment Effect Coefficient	0.097 * (0.046)	0.161 ** (0.049)
Treatment Effect in Odds Ratio	10.20%	17.5%
Control	Yes	Yes

Standard errors in parenthesis
* Significance at 5% level; ** Significance at 1% level; *** Significance at 0.1% level

Table A.3.16 - HE attendance regression

Independent Variables			
Ruralization	Yes ***	Average Income	0.092 *** (0.001)
State	Yes ***	Work Factor	-0.236 *** (0.012)
Race	Yes ***	Gender (F = 1) * Avg. Income	0.012 *** (0.002)
Age	-0.096 *** (0.001)	Race (non-white = 1) * Avg. Income	0.134 *** (0.003)
Gender (F = 1)	0.316 *** (0.014)	-	-
Observations	936,372		
Nagelkerke R²	18.9%		

Standard errors in parenthesis

4. Welfare and Labor Supply Effects of Student Financing Schemes in Higher Education

4.1. Introduction

The rapid expansion of higher education (HE) in the presence of imperfect credit markets has led governments to rethink student funding systems worldwide. Hence, especially in the last three decades, many countries have introduced student loan programs in an attempt to facilitate access to HE (Larraín and Zurita 2008; Atuahene 2008; Chapman and Sinning 2014). The two most frequently adopted designs have been mortgage (ML) and income-contingent loans (ICL). The latter, pioneered by Australia and the United Kingdom (UK), has become increasingly popular (Britton, et al., 2019), in order to mitigate the adverse effects of observed uncertainty in graduates' real earnings (Chapman, et al., 2014), and have been implemented also in non-OECD nations, as it is the case of Brazil and Thailand (Dearden and Nascimento 2019; Chapman and Lounkaew 2010; Chapman et al. 2010). While the existing literature on the topic has focused mainly on public finance outcomes and the effects on graduates' repayment burdens (for instance, Barr et al. 2019, Chapman et al. 2014, Belfield et al. 2018), there are still dimensions of ICL schemes which have remained largely unexplored. Among these are the graduates' welfare effects of shifts in loan schemes and the labor supply responses to these shifts. This chapter assesses these two dimensions.

This research contributes to the understanding of the effects of ICL schemes in at least three ways. First, I build a model which allows us to analyze the effects of shifting from a ML to an ICL scheme on welfare[23] once uncertainty and risk aversion are factored in - an issue which, to the best of my knowledge, has only been addressed by Migali (2012).

In second place, I explore the labor supply responses of student loan schemes by introducing elastic labor supply into this theoretical framework. Since an ICL acts as an incremental marginal tax on graduates, this could reduce labor

---

[23] I work with a partial equilibrium model, in which individuals are higher education students and graduates. Therefore, what I call welfare throughout this chapter is simply the sum of the utility of these individuals.

supply and potentially loan repayments and tax receipts – as indicated by the well-established literature on the labor supply effects of taxation (Blomqvist 1983; Blomqvist and Hansell-Brusewitz 1990; Blundell, Duncan and Meghir 1998; Ziliak and Kniesner 2005; and Keane 2011). Therefore, understanding the extent to which a move from a ML to an ICL reduces labor supply is crucial in order to unravel the cost-effects of such shift. This is the first research to simulate labor supply responses of different student financing schemes in such fashion.

Finally, I calibrate the former model parameters with data from a real experience of a recent shift from a ML to an ICL in Brazil in order to simulate its' labor supply and welfare effects. Additionally, this research also studies how this shift in loan schemes affects a set of outputs of interest by gender and race. In this sense, I not only simulate the effects of a recently-implemented loan design in an emerging economy, but also assess its' heterogonous effects on welfare and labor supply across different population subgroups. This flexible approach allows me to contribute to the literature by simulating how different income and policy parameters affect welfare and labor supply responses to changes in student loan designs, drawing clear policy implications.

This research's results indicate that changing from a ML to an ICL scheme (i) decreases labor supply; (ii) increases graduates' welfare; (iii) reduces repayment burdens, and (iv) increases the number of years until the debt is fully repaid. Moreover, shifting to an ICL scheme is especially welfare-enhancing for women and non-white people, two population groups who have lower initial earnings, flatter income growth curves throughout their working lifetimes and also face a greater unemployment risk.

The chapter is organized as follows. Section 4.2 provides a brief literature review on the welfare and labor supply effects of student financing schemes. In section 4.3, I present the theoretical model, and derive the individuals' optimal leisure choices under both a ML and an ICL as well as its main implications when earnings are static and determined by a single lifetime shock. In section 4.4, the results of the dynamic simulations are presented. Finally, section 4.5 concludes the article.

4.2. Literature

Most of the literature on the impacts of different loan schemes to fund higher education has focused on public finance outcomes and its' effects on graduates' repayment burdens, especially in the UK, US and Australia (Barr et al. 2019; Chapman et al. 2014; Belfield et al. 2018). Research for non-OECD countries is scarcer, Dearden and Nascimento (2019) being among the few to investigate the recent change in the Brazilian student funding system.

In this research, I focus on the analysis of the welfare effects of shifts in loan schemes and the labor supply responses to these shifts, two largely unexplored issues. While there have been several studies that have investigated the link between educational outcomes and wage uncertainty, such as Padula and Pistaferri (2001), Acemoglu (2002) and Chen (2008), and also the role of ICLs in consumption smoothing (Chapman, et al. 2014), the first to explicitly compute willingness to pay to switch between different student loan schemes was Migali (2012). Migali (2012) compared graduates' discounted expected utilities under a ML and an ICL and under two scenarios: (i) in which earnings are static and determined by a single lifetime normally distributed shock; and (ii) in which earnings are dynamic and wages' growth follows a Geometric Brownian Motion. Migali then calibrated the models' parameters using UK's data and confirmed – in both scenarios - the important insurance benefits and welfare gains of an ICL when compared to a ML, especially among graduates from poor families.

Migali's paper, however, abstracts from labor supply effects, which, according to Turnovsky (2000) and Leal and Turchick (2020) can give rise to misleading policy implications – or, in this case to biased welfare effects of different loan designs. On top of that, it is still unsure whether or not the results obtained would generalize to the case of emerging economies. Both of these points are addressed in the present study.

The second dimension studied in this chapter regards the labor supply effects of shifting to an ICL scheme. If indeed an ICL induces graduates to reduce their labor supply, as has been suggested by the literature on the marginal effects of income taxation (Blomqvist 1983, Blomqvist and Hansell-Brusewitz 1990, Blundell, Duncan and Meghir 1998, Ziliak and Kniesner

2005, and Keane 2011), the implicit reduction in loan repayments and tax receipts could in fact increase the cost of funding higher education.

This topic has been studied by three main articles and results are mixed. The first paper to investigate this issue was Chapman and Leigh (2009), in which they explored a sharp discontinuity in Australia's taxable income of graduates that took out college loans. They find a significant degree of bunching below this threshold, although the effect is economically small (only around 0.3% of those with college debts bunch, and the degree to which they do so is likely to be very small).

The second example is Herbst (2019), which estimates the causal impacts of moving onto an ICL on student debt repayment, financial health, and employment proxies using two identification strategies: an instrumental variable design exploiting variation in the ability of loan service agents to enroll borrowers in an ICL scheme and a difference-in-differences between ICL enrollees and non-enrollees. Although the analysis of labor supply responsiveness of ICL is not his primary objective, Herbst' findings suggest that there are no labor supply responses associated with moving to an ICL.

Finally, Britton and Gruber (2020) investigate this issue using a similar strategy to that of Chapman and Leigh (2009), by exploring bunching at various loan repayment thresholds in the UK between 2002 and 2014. Their findings suggest that the UK's income contingent repayment plan does not cause borrowers to reduce labor supply, at least for those with earnings near to the threshold.

The bottom line is that evidence on the labor supply responsiveness of shifting to an ICL scheme is still tenuous, once the few empirical investigations that have studied this issue have all its' set of caveats and furthermore achieve mixed results. Therefore, the current study proposes an alternative approach to examine this under-studied topic, which consists of building a partial equilibrium model and calibrating its' parameters with real data so as to simulate the labor supply (and welfare) effects of shifting from a ML to an ICL scheme.

4.3. Theoretical model

In this section, I present the theoretical model and highlight the main implications of introducing elastic labor supply to the existing framework

(Migali 2012). I start (4.3.1) by introducing the agents' instantaneous utility function and the main assumptions behind the model. Then (4.3.2), I compute the individuals' instantaneous utilities under the two student loan schemes of interest – Mortgage Loan (ML) and Income-contingent Loan (ICL) – and compare the optimal leisure choices under each one. Finally (4.3.3), I compare the repayment period and repayment burden under these two schemes assuming static earnings. All the mentioned outputs – in addition to lifetime discounted utilities – in a setting with dynamic earnings will be studied through numerical simulations in subsection 4.3.4.

4.3.1. Main assumptions

All individuals in this economy have identical utility functions that depend on consumption $C \in \mathbb{R}^+$ and leisure $l \in [0, 1]$. As in Turnovsky (2000), the CRRA instantaneous utility at t is given by:

$$U(C_t, l_t) = \begin{cases} \dfrac{(C_t l_t^\eta)^\alpha - 1}{\alpha}, & if \ \alpha \neq 0 \\ \log(C_t l_t^\eta), & if \ \alpha = 0 \end{cases} \tag{4.1}$$

where $\alpha \leq 1$ is a risk tolerance parameter (one minus relative risk aversion), and $\eta \geq 0$ is the elasticity of leisure. It may be noted that the framework of Migali (2012) can be obtained by simply imposing $\eta = 0$, that is, by imposing an inelastic labor supply. The addition of a positive η is key to the analysis, since it not only allows us to comprehend what drives graduates' labor supply and how it is affected by changes in loan repayment designs, but this extension might also impact optimal policy decisions on an extensive margin (Turnovsky 2000, Leal and Turchick 2020).

Finally, I will be working with a discrete-time setting in which each period $t \in \mathbb{N}$ equals one year, and in order to ensure concavity, let us assume $\eta\alpha < 1$.

This exercise focuses on individuals who completed higher education (HE), and initially assumes zero unemployment (this assumption will be relaxed in section 4.4). These individuals go to university for s years full time and education has the same cost $c \in \mathbb{R}^+$ for everybody. Earnings (and consequently consumption) during the schooling are assumed to be zero, while graduates receive an income $y \in \mathbb{R}^+$ for each unit of labor supplied,

whence a person with a leisure of l_t and an income of y_t at t will receive $(1 - l_t)y_t$ in this period. Individuals work for r years before retiring and the economy produces a single homogenous final good to be treated as numeráire.

The only source of student funding available is a government-backed loan, that might follow either a ML or an ICL scheme. Under a ML, individuals take out a loan and repay fixed and mandatory installments at each period, irrespective of the debtor's capacity to pay. The model assumes that all prospective students take out public loans to cover the entire cost of education (therefore, the total ML cost is equal to c). Graduates start repaying their debts right after graduation through fixed installments φ at each period t, where $\varphi < y_t, \forall t$ is assumed, hence, zero default risk. Consequently, the repayment period is equal to $T_{ML} = \dfrac{c}{\varphi}$ years.

Under an ICL, individuals take out a loan and start to pay back soon after graduation according to their earnings. Graduates with higher wages pay back a larger share of their income in each year and finish repaying the loan in less time, whereas graduates with lower wages pay back less and take longer to fully pay off their loans. In this setting, there is no risk of default. The loan covers the entire cost of HE c, and that graduates pay a fixed rate $\gamma \in (0,1)$ of their earnings in each year. Therefore, under an ICL, the repayment period T_{ICL} is not fixed and depend on each person's earnings' path.

4.3.2. Optimal leisure choices under ML and ICL

This section studies the individuals' optimal leisure choices under each one of the proposed loan schemes. Let us start with the ML case. In the absence of a savings mechanism, the agents' problem can be stated in a static form, and it can be subdivided into two different problems: the first one related to the budget constraint during the repayment period (from $s + 1$ to $T_{ML} + s$) and the second during the post-repayment period (from $T_{ML} + s + 1$ to $s + r$). In each year t during the repayment period, individuals choose $\left(C_{ML,t}, l_{ML,t}\right) \in \mathbb{R}^+ \times [0,1]$ so as to maximize $U(C_{ML,t}, l_{ML,t})$ subject to:

$$C_{ML,t} \leq \left(1 - l_{ML,t}\right)y_t - \varphi \tag{4.2}$$

Due to U being strictly increasing in its first argument, (4.2) must hold with equality at the solution, so that it is trivial to find $C_{ML,t}$ once $l_{ML,t}$ has been found. The first-order condition imposes the following solution for $l_{ML,t}$:

$$l_{ML,t} = \frac{\eta(y_t - \varphi)}{(\eta + 1)y_t} \tag{4.3}$$

which belongs to the $[0, 1)$ interval.

Proposition 1. *On the optimal leisure choice during the ML repayment period*

(i) *The leisure of graduates during the ML repayment period increases with η and y, and decreases with φ*

(ii) *The greater the difference between y_t and φ, the less individuals will be willing to work*

Proof. Item (i) stems from simply differentiating (4.3) with respect to each variable, implying $\frac{\partial l_{ML,t}}{\partial \eta} = \frac{y_t - \varphi}{y_t(\eta+1)^2} > 0$; $\frac{\partial l_{ML,t}}{\partial y_t} = \frac{\eta\varphi}{(\eta+1)y_t^2} > 0$; and $\frac{\partial l_{ML,t}}{\partial \varphi} = -\frac{\eta}{(\eta+1)y_t} < 0$. Item (ii) comes straightforwardly form (4.3).

During the post-repayment period, individuals choose $\left(C_{PR,t}, l_{PR,t}\right) \in \mathbb{R}^+ \times [0, 1]$ so as to maximize $U(C_{PR,t}, l_{PR,t})$ subject to:

$$C_{PR,t} \leq \left(1 - l_{PR,t}\right)y_t \tag{4.4}$$

Analogous to (4.2), (4.4) must also hold with equality at the solution. In this case, the first-order condition imposes the well-known solution for $l_{PR,t}$:

$$l_{PR,t} = \frac{\eta}{\eta + 1} \tag{4.5}$$

Let us now move to the ICL case. Under an ICL scheme, the static agent's problem can again be subdivided into two different ones: the first one related to the budget constraint during the repayment period (from $s + 1$ to $T_{ICL} + s$) and the second during the post-repayment period (from $T_{ICL} + s + 1$ to $s + r$). In each year t during the repayment period, individuals choose $\left(C_{ICL,t}, l_{ICL,t}\right) \in \mathbb{R}^+ \times [0, 1]$ so as to maximize $U(C_{ICL,t}, l_{ICL,t})$ subject to:

$$C_{ICL,t} \le (1 - \gamma)(1 - l_{ML,t})y_t \qquad (4.6)$$

Once again, (4.6) must hold with equality at the solution. In this case, the first-order condition imposes the following solution for $l_{ICL,t}$:

$$l_{ICL,t} = \frac{\eta}{\eta + 1} \qquad (4.7)$$

Finally, the agent's problem during the post-repayment period for those who took out an ICL is essentially the same as those who took out a ML. Therefore, (4) and (5) also hold in this case.

Proposition 2. *On the comparison of the optimal leisure choices during the ML repayment period, ICL repayment period and post-repayment period*

   (i)   The leisure of graduates during the ICL repayment period is the same as in the post-repayment period and it only depends on η (it increases with it)

   (ii)   During the repayment period, graduates who took out a ML will work more than those who took out an ICL (that is, $l_{ICL,t} > l_{ML,t}$)

   (iii) The greater the leisure elasticity (η), the greater is the gap between $l_{ICL,t}$ and $l_{ML,t}$

   (iv)   The greater the difference between y_t and φ, the greater is the relative difference between $l_{ICL,t}$ and $l_{ML,t}$. More precisely, $l_{ML,t} = \frac{y_t - \varphi}{y_t} l_{ICL}$

Proof. Item (i) comes straightforwardly from (4.7) and (4.5) in which we have $\frac{\partial l_{ICL,t}}{\partial \eta} = \frac{\partial l_{PR,t}}{\partial \eta} = \frac{1}{(\eta+1)^2} > 0$. For items (ii) and (iii), let us first compute the difference between $l_{ICL,t}$ and $l_{ML,t}$, that is (4.7) minus (4.3), which equals $\frac{\eta\varphi}{(\eta+1)y_t}$. Let us call this difference in leisure choices $\mathbb{D}$. Since $\mathbb{D} > 0$, item (ii) is proved. Item (iii) stems from differentiating $\mathbb{D}$ with respect to η, so that $\frac{\partial \mathbb{D}}{\partial \eta} = \frac{\varphi}{y_t(\eta+1)^2} > 0$. Item (iv) comes simply from dividing (4.3) by (4.7), which equals $\frac{y_t - \varphi}{y_t}$.

4.3.3. Repayment under ML and ICL with static earnings

This subsection investigates and compares the repayment period and repayment burden under the two loan schemes of interest. To that end, I follow Migali (2012) and assume that lifetime earnings are static and determined by a single normally distributed shock (as in Hartog and Serrano 2007). Put differently, graduates obtain an uncertain wage $\tilde{y} \in \mathbb{R}^+$ determined by a random draw by the time of graduation which remains unchanged from then on. Additionally, earnings are assumed to be higher than a minimum level y_{min} (where $y_{min} > \varphi$) with $E(\tilde{y}) = 1$ and $Var(\tilde{y}) = \sigma > 0$. Moreover, in this exercise let us assume zero real interest rates on the loan under both schemes (this assumption is relaxed when evaluating dynamic earnings in section 4.3.4).

Under a ML, the yearly repayment is fixed at φ ($R_{ML} = \varphi$), whence the repayment period T_{ML} remains the same as presented in subsection 4.3.2 (i.e., $T_{ML} = \frac{c}{\varphi}$). Under an ICL, however, the yearly repayment is stochastic and equals $\gamma(1 - l_{ICL})\tilde{y}$, which by using (4.7) – that still stands in this analysis – entails $R_{ICL} = \frac{\gamma\tilde{y}}{(1+\eta)}$ and consequently $\tilde{T}_{ICL} = \frac{c(1+\eta)}{\gamma\tilde{y}}$.

Proposition 3. *If earnings are static and determined by a single lifetime shock, then $E(R_{ICL}) = R_{ML}$ if, and only if, $\gamma = \varphi(1 + \eta)$*

Proof. Since $E(\tilde{y}) = 1$, then $E(R_{ICL}) = \frac{\gamma}{(1+\eta)} \times E(\tilde{y}) = \frac{\gamma}{(1+\eta)}$. Therefore, $E(R_{ICL}) = R_{ML} \leftrightarrow \frac{\gamma}{(1+\eta)} = \varphi \leftrightarrow \gamma = \varphi(1 + \eta)$.

Proposition 4. *If earnings are static and determined by a single lifetime shock and the expected repayment under a ML and ICL are the same, then $E(\tilde{T}_{ICL}) > T_{ML}$.*

Proof. Since $\tilde{T}_{ICL} = \frac{c(1+\eta)}{\gamma\tilde{y}}$ and with $\gamma = \varphi(1 + \eta)$, we have $E(\tilde{T}_{ICL}) = \frac{c}{\varphi} \times E\left(\frac{1}{\tilde{y}}\right)$. By the Jensen's inequality on strictly convex functions, we know that $E\left(\frac{1}{\tilde{y}}\right) > \frac{1}{E(\tilde{y})} = 1$, therefore $\frac{c}{\varphi} \times E\left(\frac{1}{\tilde{y}}\right) > \frac{c}{\varphi}$, and hence $E(\tilde{T}_{ICL}) > T_{ML}$.

Proposition 4 states that, under equal expected annual repayment among schemes, the expected repayment period for an ICL is longer than it is for a ML. Additionally, the repayment burden, which is the ratio between the

annual repayment and the annual income, is also different between schemes. Under an ICL, the repayment burden is fixed at γ ($RB_{ICL} = \gamma$), which by definition is the fixed rate graduates pay on their annual income. For a ML, it is the repayment burden that is stochastic in this case, so that $\widetilde{RB}_{ML} = \frac{\varphi}{(1-l_{ML})\tilde{y}}$, which by using (4.3) yields $\widetilde{RB}_{ML} = \frac{(1+\eta)\varphi}{\tilde{y}+\eta\varphi}$.

Proposition 5. *If earnings are static and determined by a single lifetime shock and the expected repayment under a ML and ICL are the same, then* $E(\widetilde{RB}_{ML}) > RB_{ICL}$ *if* $\gamma < \frac{\sqrt{5\eta^2+8\eta+4}-\eta}{2(1+\eta)}$.

Proof. Since $\widetilde{RB}_{ML} = \frac{(1+\eta)\varphi}{\tilde{y}+\eta\varphi}$ and with $\gamma = \varphi(1+\eta)$, we have $E\left(\widetilde{RB}_{ML}\right) = E\left(\frac{1}{\tilde{y}\gamma + \frac{\eta}{1+\eta}}\right)$. By the Jensen's inequality on strictly convex functions, we know that $E\left(\frac{1}{\tilde{y}\gamma + \eta/(1+\eta)}\right) > \frac{1}{E(\tilde{y})\gamma + \frac{\eta}{1+\eta}} = \frac{1}{\gamma + \frac{\eta}{1+\eta}}$. If $\gamma < \frac{\sqrt{5\eta^2+8\eta+4}-\eta}{2(1+\eta)}$, then $\frac{1}{\gamma+\frac{\eta}{1+\eta}} > \gamma$ and hence $E(\widetilde{RB}_{ML}) > RB_{ICL}$.

Proposition 5 states that the repayment burden under a ML is higher than the one in an ICL for a sufficiently low γ (as long as $\gamma < \frac{\sqrt{5\eta^2+8\eta+4}-\eta}{2(1+\eta)}$). When labor supply is perfectly inelastic (i.e., $\eta = 0$), this condition becomes $\gamma < 1$, which is always true since $\gamma \in (0,1)$. When $\eta > 0$, $E(\widetilde{RB}_{ML}) > RB_{ICL}$ usually, although not always, hold. This is because for more realistic values of η and γ, the inequality $\gamma < \frac{\sqrt{5\eta^2+8\eta+4}-\eta}{2(1+\eta)}$ will indeed be satisfied. Most studies suggest that the individuals' time devoted to leisure lies between 70% and 80%, entailing a leisure elasticity between $\eta = 2$ and $\eta = 4$[24], in which cases γ would only need to be lower than approximately 70% for $E(\widetilde{RB}_{ML}) > RB_{ICL}$ to be true. This threshold rate, in turn, is much higher than the ICL repayment rates observed in the real world - which lies between $\gamma = 4\%$ and $\gamma = 12\%$[25].

---

[24] Turnovsky (2000), Gimenez-Nadal and Sevilla (2012)
[25] Britton et al. (2019)

4.3.4. Dynamic earnings

This subsection extends the model described so far to the case in which graduates' earnings increase with age in a stochastic fashion. I follow Migali (2012) and consider that the earnings growth rate follows a geometric Brownian motion, so that y_τ satisfies:

$$\frac{dy_\tau}{y_\tau} = \lambda d\tau + \sigma dW_\tau \tag{4.8}$$

where W_τ is a Wiener process, λ is the income's deterministic growth rate (the percentage drift), and σ accounts for the earnings' percentage volatility. For a given initial income value y_0, the well-known analytic solution to the stochastic differential equation (4.8) is:

$$y_\tau = y_0 \times exp\left(\left(\lambda - \frac{\sigma^2}{2}\right)\tau + \sigma W_\tau\right) \tag{4.9}$$

Since I am working with a discrete-time setting, I discretize (4.9) over the continuous interval [0, T] following the Euler-Maruyama method as described in Higham (2001). Thus, the following path for annual earnings for an individual working life is generated:

$$y_t = y_{t-1} + y_{t-1}\lambda\Delta\tau + y_{t-1}\sigma\left(W_{\theta,t} - W_{\theta,t-1}\right) \tag{4.10}$$

where $\Delta\tau = T/L$ for some positive integer L and the increments $W_{\theta,t} - W_{\theta,t-1}$ are generated by discretized Brownian paths, in which $\theta, t = \Delta\tau$ [26]. Let us assume individuals start working at 25 and retire at 65, therefore 40 annual earnings, whence $r = 40$, are generated.

The outputs with dynamic earnings under a ML and an ICL will be studied through numerical simulations in which the model's parameters will be calibrated with real data from Brazil. These outputs are (i) the dynamic earnings' path; (ii) the dynamic leisure choices; (iii) the dynamic repayment burdens; (iv) the average repayment period; and (v) the lifetime discounted utilities. Outputs (i) to (iv) have already been explained throughout section

---

[26] See Higham (2001) for an explicit derivation of expression (10)

4.3. The lifetime discounted utility (V), on the other hand, has not yet been introduced. For graduates who took out a ML, it is expressed as:

$$V_{ML} = E \left\{ \sum_{t=s+1}^{T_{ML}+s} \rho^t U\big(C_{ML,t}, l_{ML,t}\big) \right.$$
$$\left. + \sum_{t=T_{ML}+s+1}^{s+r} \rho^t U\big(C_{PR,t}, l_{PR,t}\big) \right\} \tag{4.11}$$

While for graduates who took out an ICL, V is expressed as:

$$V_{ICL} = E \left\{ \sum_{t=s+1}^{\tilde{T}_{ICL}+s} \rho^t U\big(C_{ICL,t}, l_{ICL,t}\big) \right.$$
$$\left. + \sum_{t=\tilde{T}_{ICL}+s+1}^{s+r} \rho^t U\big(C_{PR,t}, l_{PR,t}\big) \right\} \tag{4.12}$$

where $\rho \in (0, 1)$ is the subjective discount factor that measures how much the present is taken in consideration with the future.

4.4. Simulation

In this section, I simulate a set of dynamic outputs (earnings path, leisure trajectory, repayment burden), as well as the repayment period and the lifetime discounted utility for different values of the parameters and for different population subgroups (by gender and race). This is done by calibrating the model with real Brazilian data. Brazil is a particularly suitable economy to use as reference, since its' government-backed student loan program (FIES) has recently shifted from a ML to an ICL and due to the substantial gender and racial inequality in Brazil's labor market, which allows us to derive and analyze significantly different results by population subgroup.

I start (4.4.1), by expanding on the parameters' calibration, and then (4.4.2), the results of the simulations are presented, concluding with some comparative statics exercises.

4.4.1. Parameters' Calibration

In order to simulate the dynamic earnings' outputs, I calibrate the model's parameters with data from Brazil. I use as baseline $s = 4$, which is the typical duration of a university degree in Brazil, and $c = 50,000$ BRL[27]. Let us set $\alpha = -2$, entailing a relative risk aversion parameter of 3, and $\rho = 0.9$, which are in line with studies that have estimated these parameters for the Brazilian economy[28]. For the leisure elasticity parameter, I use $\eta = 3.8$, which implies a post-repayment share of time devoted to leisure of 79%, in line with international time-use data. Furthermore, the model from section 4.3 is generalized to accommodate positive real interest rates on the loans (i), a grace period after graduation before which individuals do not need to start the repayments (GP) and an ICL repayment threshold (RT).

In order to calibrate the policy parameters γ and φ, I use as reference the rules of the FIES (*Fundo de Financiamento ao Estudante do Ensino Superior)*, Brazil's higher education student loan program. Created in 1999 by the federal government, the FIES offers financial aid up to 100% of monthly tuition fees to students attending private HE institutions. Since its conception, the program has changed its set of rules multiple times. As of 2021, the following rules apply:

i- Eligibility: Students with up to 5 minimum wages of per capita family income who scored above 450 in the ENEM (*Exame Nacional do Ensino Médio*) and above 0 in the ENEM's dissertation.

ii- Real interest rate on loans: 0% to students with up to 3 minimum wages of per capita family income; 2.5% to 3.5% to students with a per capita family income between 3 and 5 minimum wages from the Central-West, North and Northeast regions; and 6.5% to students with a per capita family income between 3 and 5 minimum wages from the South and Southeast regions.

iii- Until 2017, it followed a fixed-schedule loan repayment, irrespective of the debtor's capacity to pay, with an 18-months grace period after graduation (after which individuals had 3 times the duration of the higher education course to fully repay the FIES

---

[27] According to the National Fund for Education and Development (FNDE), the average tuition fee paid by the federal government for FIES contracts in 2017 was BRL 45,840.
[28] Antunes et al. (2015), Fajardo et al. (2012), Araújo (2005), Issler and Piqueira (2000)

loan). From 2018 onwards, the FIES shifted to an income-contingent scheme, with a maximum repayment rate of 10% and no grace period.

As baseline parameters in the model, let us then set $\gamma = 10\%$, $\phi = c/3s = 4{,}166.7$ BRL, $i = 0\%$, $GP = 0$, and $RT = 0$. Nevertheless, in section 4.4.2, the model's outputs will also be evaluated considering different values for these policy parameters.

Finally, data from the Continuous PNAD, Brazil's national household sample survey, from 2019 is used to calibrate the dynamic earnings' path parameters: the initial income y_0, (that is, the graduates' initial wage at 25), the deterministic growth rate λ, and the percentage volatility σ. Created in 2012, the Continuous PNAD follows a rotating panel structure and provides information on sociodemographic characteristics, education, labor and income at individual level. The year 2019 was selected for it is not polluted by Brazil's economic crisis from 2015/16 neither by the Covid-19 pandemic in 2020. Table 4.1 displays the average value, initial value, average annual growth rate and annual percentage volatility of Brazilian graduates' earnings for the entire population and for subgroups.

Table 4.1 - Graduates annual income information

	Mean	Initial Income	Growth rate	Volatility
Total	53,509	27,144	2.7%	6.2%
By gender				
Males	69,156	32,509	2.8%	9.4%
Females	42,264	23,481	2.1%	8.4%
By race				
White	61,381	29,107	2.8%	8.5%
Blacks/Browns	42,316	23,655	2.3%	8.4%

Income values in 2019 BRLs

Table 4.1 shows that, for the entire population of Brazilian graduates, the average earning in 2019 was of 53,509 BRL, with an initial value (at 25 years old) of 27,144 BRL, an average annual growth rate of 2.7% and an annual volatility of 6.2%. Moreover, males and white persons have a slightly higher income volatility than females and black/brown individuals, but with a

significantly greater initial income value and with a higher wage growth throughout their working lifetimes.

4.4.2. Results

This section presents the results of the dynamic earnings simulations. I start with the evaluation of the outputs for the entire Brazilian population and under the baseline parameters, that is: $\eta = 3.8$, $\alpha = -2$, $\rho = 0.9$, $c = 50{,}000$, $s = 4$, $\gamma = 10\%$, $\phi = c/3s = 4{,}166.7$, $i = 0\%$, $GP = 0$, and $RT = 0$. From Table 4.1, let us set $y_0 = 27{,}144$, $\lambda = 2.7\%$ and $\sigma = 6.2\%$. In order to build the simulations, I run 1,000 random earnings' paths following equation 4.10. For illustrative purposes, Figure 4.1 displays a random subsample of 10 out of these 1,000 simulated curves.

Figure 4.1 - Earnings (in BRLs) per age simulations for the total Brazilian population
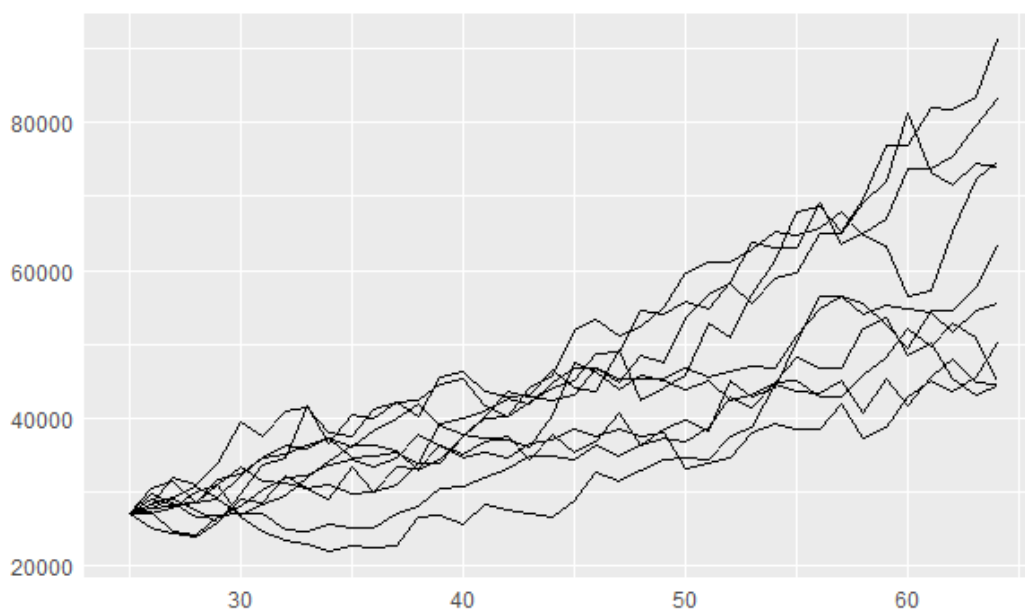
Next, Figure 4.2 and Figure 4.3 compute the trajectory for the average repayment burden and average leisure choices among the 1,000 simulations under both an ICL and a ML scheme.

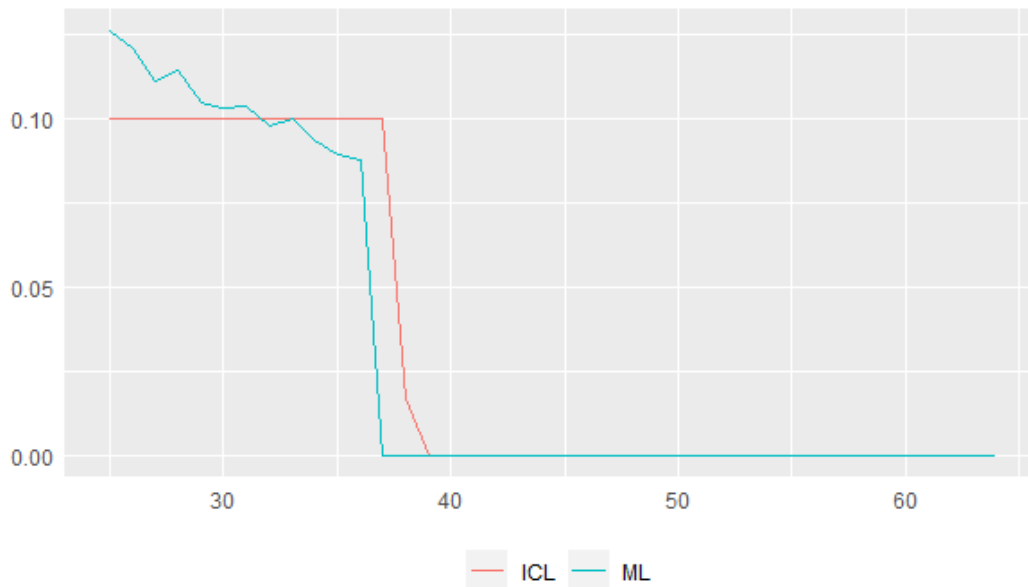*Figure 4.2 - Average repayment burden per age*

Figure 4.3 - Average share of time devoted to leisure per age
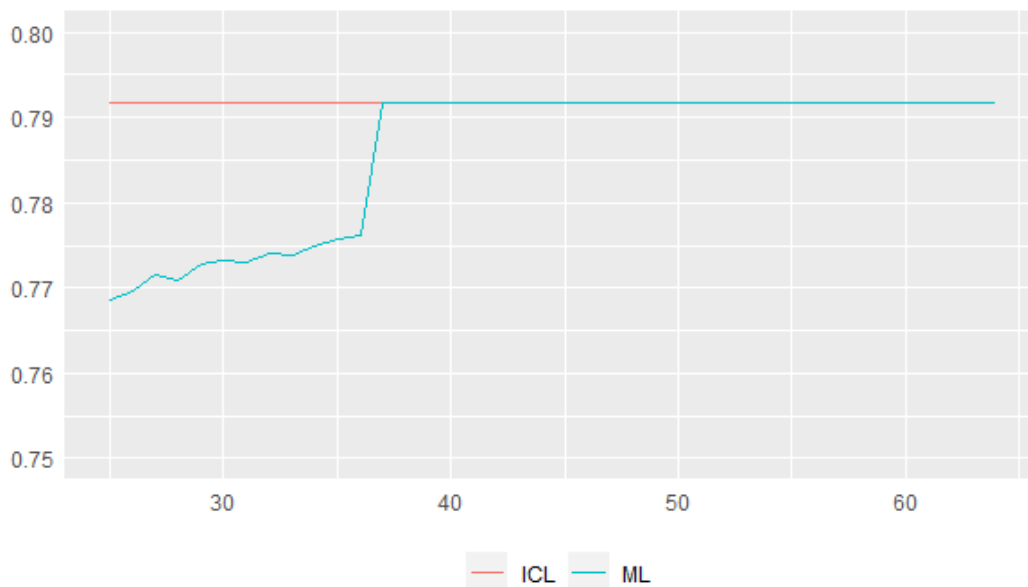
Figure 4.2 shows that the average repayment burden under a ML is greater than it is under an ICL for recent graduates. However, as earnings' increase, this dynamic changes and the ML's repayment burden becomes lower than γ. These simulations generate an average repayment period of approximately 14 years for an ICL, which is greater than the fixed 12 years from the ML. Moreover, as shown in Figure 4.3, a ML scheme induces graduates to work

a greater number of hours during the repayment period, a result that had already been anticipated by Proposition 2.ii (note that Propositions 1 and 2 refer to the individuals' leisure choices at each period, hence they are valid under both a static and a dynamic earnings setting). These simulations indicate that, on average, individuals who took out an ICL will work 3% less than those who took out a ML during the repayment years.

Finally, I now compare the lifetime discounted utilities V between schemes. For each of the 1,000 simulations, I obtain a single value for V_{ML} and V_{ICL}, work out the average of the simulations for each financing scheme, and present it in terms of willingness to pay to switch from a ML to an ICL (WTP) – that is, how much are graduates willing to pay soon after graduating to switch from a ML to an ICL repayment scheme so as to equalize their expected lifetime discounted utilities, presented in terms of percentage of their total cost of education c. For the entire Brazilian population and under the set of baseline parameters, a WTP of 11.8% is obtained, meaning that shifting from a ML to an ICL – under the baseline parameters - induces an increase in graduates' welfare.

Following, I evaluate the same outputs by gender and race. For this purpose, the same baseline parameters are used, and I run 1,000 earnings' paths simulations for each subgroup. The statistics from Table 4.1 are again used to calibrate the income trajectory parameters y_0, λ and σ by subgroup. Let us start with the outputs by gender. Figure 4.4 displays 10 out of the 1,000 earnings simulations for both males and females, while the repayment burdens and leisure choices for each gender are displayed in Figure 4.5 and Figure 4.6, respectively.

Figure 4.4 - Earnings (in BRLs) per age simulations by gender

Figure 4.5 - Average repayment burden per age by gender

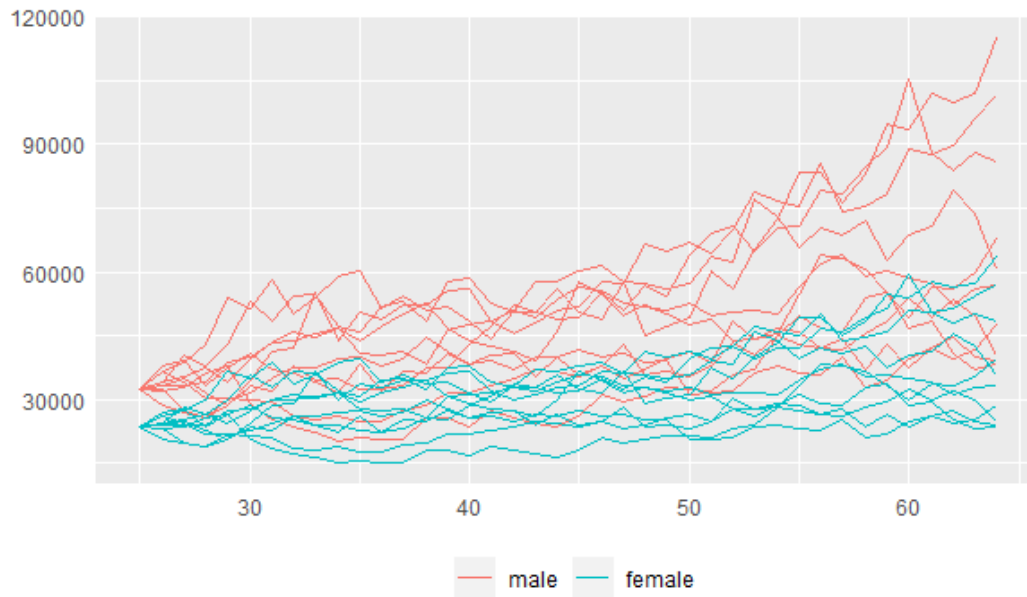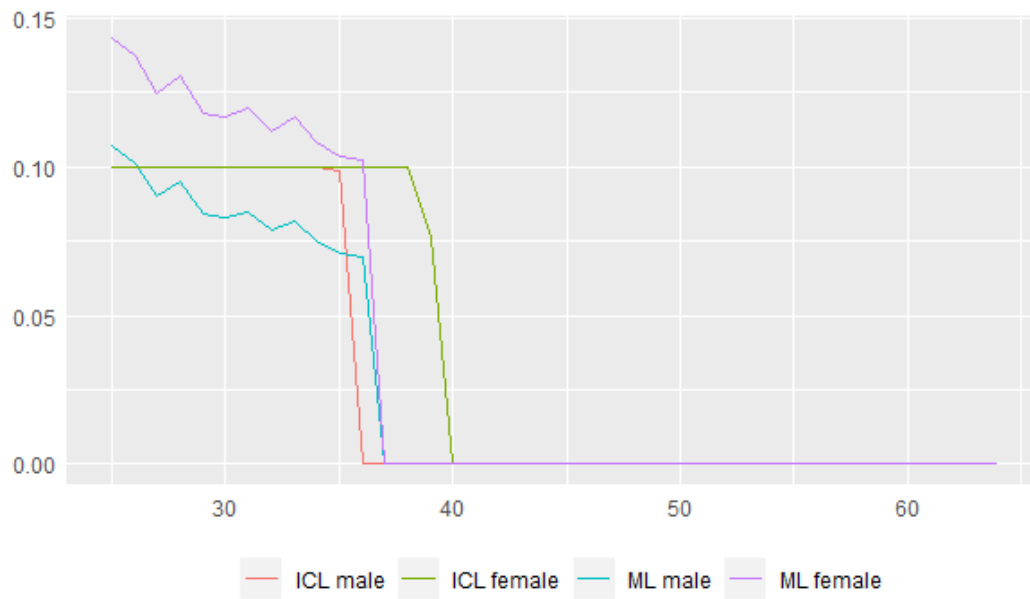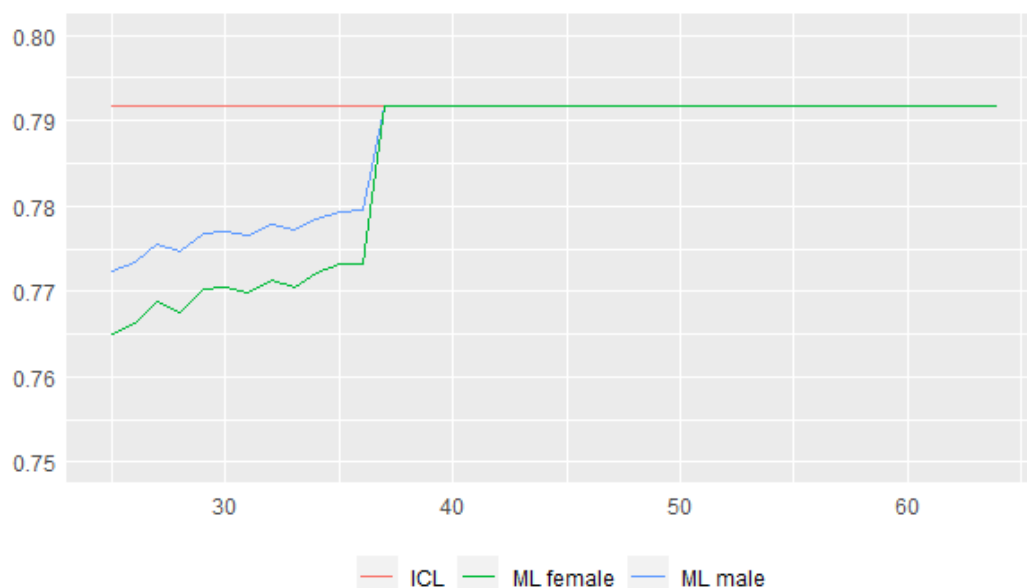*Figure 4.6 - Average share of time devoted to leisure per age by gender*

Because females have lower initial earnings and with flatter growth trajectories (Figure 4.4), they either have a higher average repayment burden under a ML, or a longer average repayment period under an ICL (Figure 4.5). Besides, under a ML, women must drastically reduce their leisure time in order to cope with the higher burdens (Figure 4.6), so that the average leisure choice for females during the repayment period is 1% lower than it is for males. Under an ICL, the share of time devoted to leisure does not change between genders once it only depends on η. As a result, shifting to an ICL scheme is more welfare-enhancing to women (WTP of 15%) than it is to men (WTP of 6.5%).

As for the results by race, the dynamics are very similar to the ones discussed by gender. Figure 4.7 displays 10 out of the 1,000 earnings simulations for white and black/brown graduates, while the repayment burdens and leisure choices for each race are displayed in Figure 4.8 and Figure 4.9, respectively.

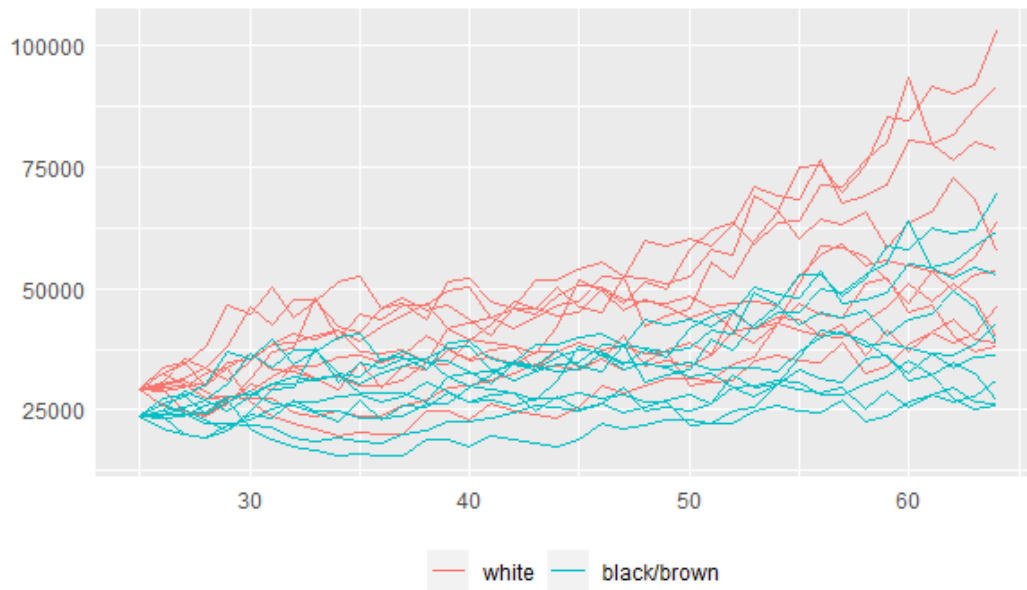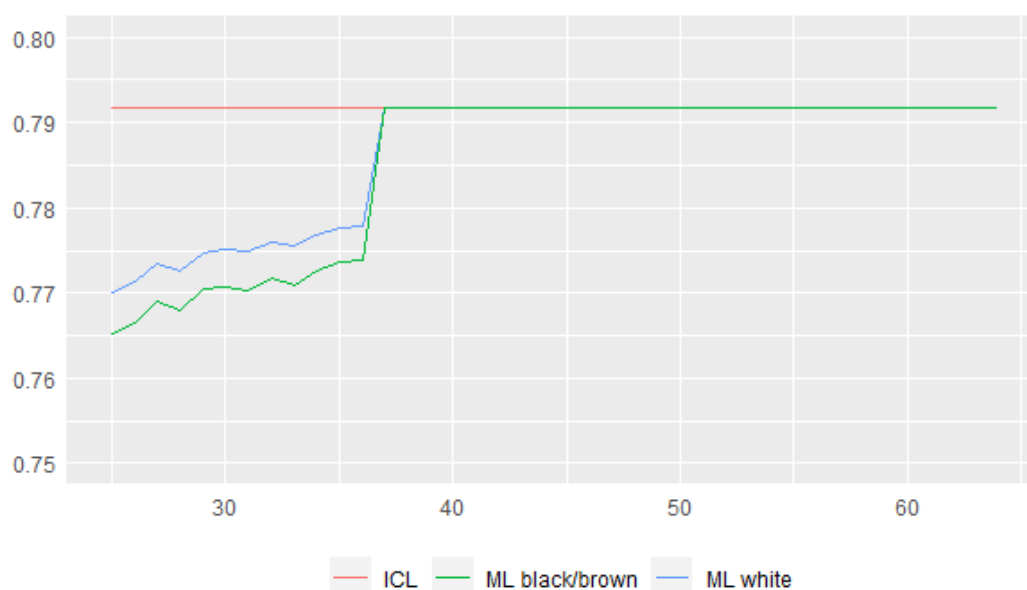*Figure 4.7 - Earnings (in BRLs) per age simulations by race*

*Figure 4.8 - Average repayment burden per age by race*

Figure 4.9 - Average share of time devoted to leisure per age by race

Similarly to the gender analysis, black/brown persons have lower initial earnings, which present flatter growth trajectories (Figure 4.7). As a result, black/brown persons have a higher average repayment burden under a ML and a longer average repayment period under a ICL (Figure 4.8), as well as a lower share of time devoted to leisure (Figure 4.9 – on average, 1% lower than for white persons). Consequently, shifting to an ICL scheme is more welfare-enhancing to black/brown graduates (WTP of 15%) than it is to white graduates (WTP of 10%).

So far, I have simulated the WTP and the average percentage difference in leisure between loan schemes assuming the baseline FIES parameters, as well as higher education parameters that reflect the average Brazilian higher education institution. However, so as to provide subsidies for policy discussion, let us now evaluate how the outputs of interest react to changes in the FIES' policy parameters. Besides, given the great heterogeneity in Brazil's college degrees, let us also assess the results of the simulations for different higher education parameters. In this analysis, whenever I change one of these parameters, the remaining ones remain constant at their baseline levels. For the policy parameters, I now introduce: (i) a positive real interest rate on the loans; (ii) a grace period; and (iii) a repayment threshold on the income-contingent loans. Figure 4.10 summarizes the results of this exercise.

Figure 4.10 - WTP and percentage difference in leisure between schemes during the repayment period $(l_{ICL}/l_{ML} - 1)$ for different values of the policy parameters
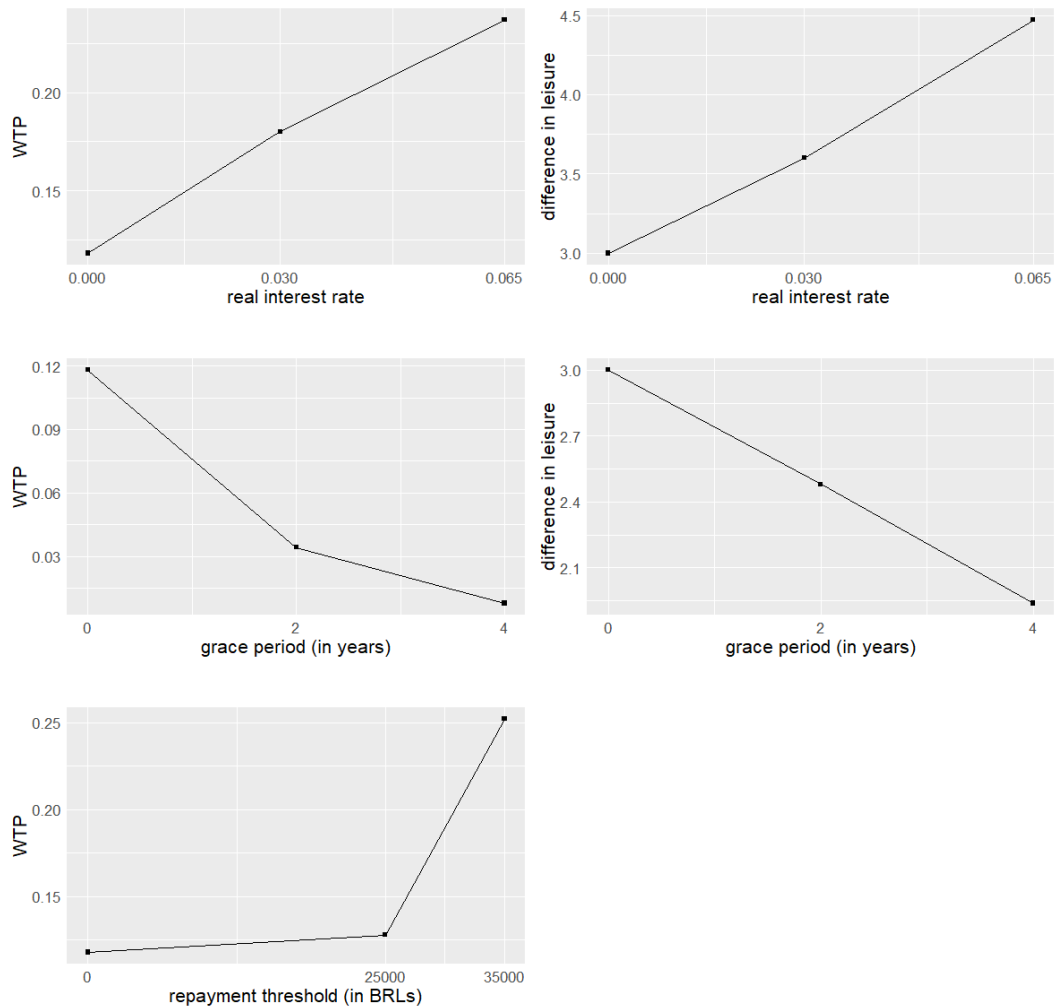
A greater interest rate increases the WTP because it generates greater repayment burdens for graduates under a ML, while for those under an ICL, these effects are smoothed across time. For this reason, the greater the interest rate, the greater is the difference in leisure time between schemes, once individuals who took out a ML will have to increase their working hours in order to bear these interest expenses. On the other hand, introducing a grace period to the loans reduces the WTP, since it provides the graduates with more time to attain higher wages and hence reduce the burdens stemming from the ML repayment. Therefore, in this case, the difference in leisure decreases as the grace period increases. Lastly, as expected, increasing the repayment threshold for the ICL increases the WTP, and, since this measure does not affect any ML parameters, it also does not affect the difference in

leisure (recall that the leisure choice is constant and only depends on η under an ICL).

Finally, I now reevaluate these outputs using different values for the higher education parameters, namely, for the total cost of education c and for the length of the college degree s. Previously, I had selected $c = 50,000$ and $s = 4$ as the baseline case since these are the average values for these parameters in Brazil. However, it is important to investigate how the WTP and difference in leisure vary when we move away from the baseline, since there is a large heterogeneity between college degrees in the country, both in terms of tuition fees and in length (in general, the latest ranges from 2 to 6 years). Figure 4.11 summarizes these results.

Figure 4.11 - WTP and percentage difference in leisure between schemes during the repayment period ($l_{ICL}/l_{ML} - 1$) for different values of the higher education parameters
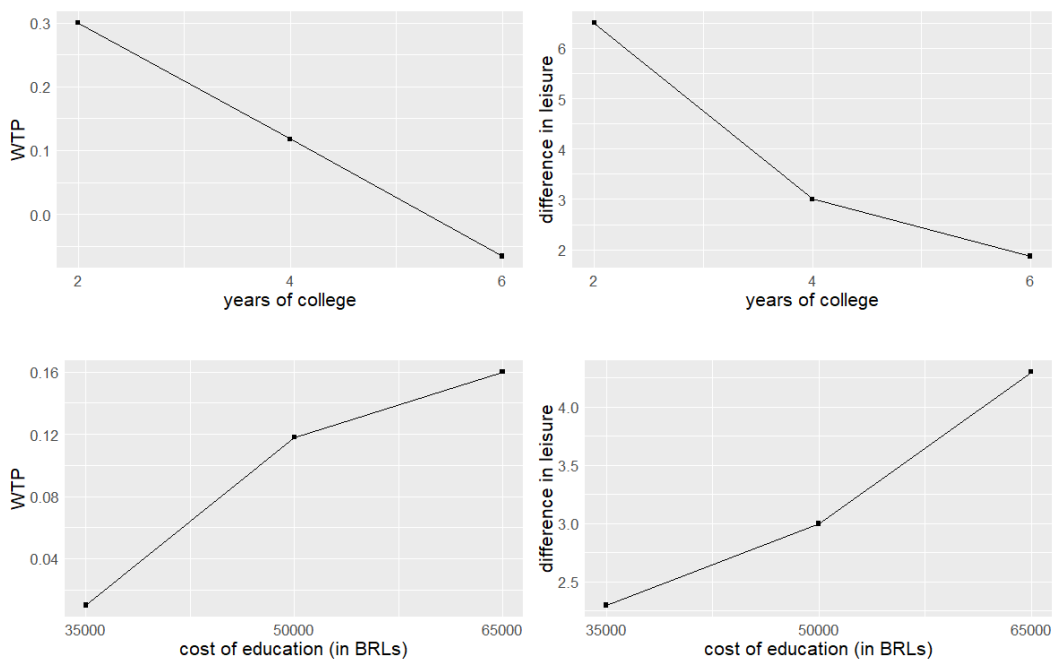

For a shorter college degree, the WTP is higher. This stems from the rules of the FIES, which states that, under a ML, students have four times the length of their degree to fully repay the loan, hence a shorter degree entails a shorter repayment period, greater repayment burdens and a greater difference in leisure time. Analogously, a longer degree extends the repayment period, reduces burdens and decreases the difference in leisure time – indeed a degree of 6 years generates a negative WTP, meaning that the lifetime discounted utility under a ML is actually greater than it is under an ICL in this case. As

for the cost of education, the greater the c, the greater are the repayment burdens that graduates under a ML must face, hence the greater are the WTP and the difference in leisure time between schemes.

At last, let us compute the WTP and contrast the leisure choices between the two loan schemes assuming scenarios in which individuals face certain years of unemployment after graduation. Let us assume that graduates who do not have the resources to repay their debts are not charged any fine, however, accrued interest increases the amount to be repaid once the individual starts working (for those who are charged a positive real interest rate – see section 4.4.1). Therefore, the more the student remains unemployed, the greater the debt to be repaid. In this sense, the effect of increasing years of unemployment on graduates' WTP and difference in leisure is analogous to the one of increasing costs of education (as seen in Figure 4.11), that is, it increases the willingness to switch to an ICL and the difference in leisure time between schemes. Figure 4.12 illustrates this dynamic.

Figure 4.12 - WTP and percentage difference in leisure between schemes during the repayment period ($l_{ICL}/l_{ML} - 1$) for increasing years of unemployment after gradiation
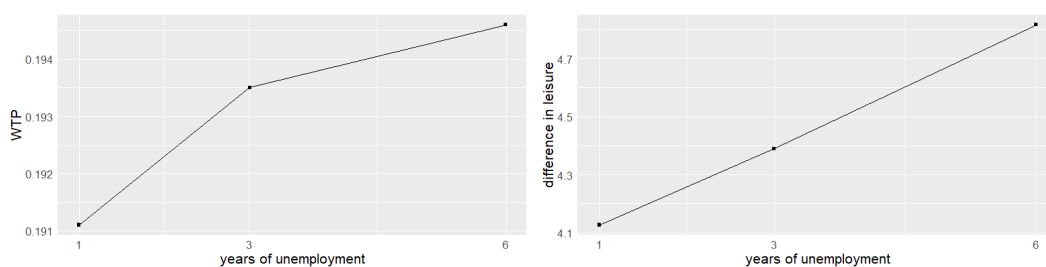
Finally, a note on the heterogenous effects of unemployment on WTP and difference in leisure between schemes will be relevant. Since unemployment rates are higher for females and non-white persons (according to the 2019 PNAD, the unemployment rate for males in that year was 10.1%, while for females it was 13.5%; and for whites it was 9.2%, while for non-whites it stayed roughly at 14%), these groups of individuals would be willing to pay a greater amount to switch from ML to ICL, and would have to work a greater number of hours under a ML once they were to leave unemployment.

The results obtained in the simulations for the specific set of calibrated parameters provide some insights on how to best structure the student funding system in order to make it more efficient and equitable. The first insight is that an ICL, in contrast to a ML scheme, generates more benefits

for graduates at their early career stages. To begin with, a ML design induces graduates to work a greater number of hours during the repayment period to deal with the burdens from their student debts and tend to generate less welfare (i.e., a positive WTP) to those individuals when compared to an ICL scheme. Besides, although this dimension was not analyzed in the present study, many papers have highlighted the role of student debt on graduates' occupational choice (Rothstein and Rouse 2011; Zhang 2013). Hence, it is likely that shifting to an ICL would not only alleviate graduates' working hours and repayment burdens, but could also allow graduates' to freely choose their post-graduation paths, with less financial constrains forcing them to choose higher-paying but less desirable jobs.

Moreover, the positive effects of an ICL design are amplified in economies with higher real interest rates and with a higher probability of early unemployment. These two factors of risk are especially present in emerging economies (World Development Indicators, The World Bank Group), hence students from these nations – in which, access to higher education tends to be more restrict than in developed economies (Roser and Ortiz-Ospina, 2013) - should be more affected by such a change on the higher education funding system's design.

Finally, shifting to an ICL scheme also increases the progressivity of the system. This is due to the fact that, when compared to a ML, an ICL design favors individuals who are economically vulnerable, more specifically, those who face greater unemployment risks, who earn less and who have flatter income growth curves throughout their working lifetimes.

4.5. Conclusion

Income-contingent loans have become an increasingly popular method to finance higher education, not only in developed economies, but also, more recently, in non-OECD countries. In this research, I investigate two largely unexplored dimensions of shifting from a ML to an ICL – its' effects on labor supply and on graduates' welfare. This is done by building a partial equilibrium model in which graduates maximize their lifetime expected utilities under wage uncertainty, risk aversion and elastic labor supply. Moreover, I calibrate the model's parameters with real data from Brazil, where the government-backed student loan program (FIES) has recently moved from a ML design onto an ICL.

In the first part of this study, an analytical extension to the existing framework is presented, in which the dimension of elastic labor supply is introduced to the model. Among the implications and conclusions that might be derived from such model, the following are highlighted: (i) under a ML scheme, graduates' labor supply depends on their income (decreases with it), on the repayment installments (increases with it) and on the leisure elasticity parameter (decreases with it), while under an ICL, it only depends on the leisure elasticity parameter (decreases with it); (ii) the labor supply under an ICL is lower than it is under a ML; and (iii) if earnings are static and the expected repayment is the same between the loan schemes, then the time for an ICL to be fully repaid is greater than the time for a ML to be fully repaid, and the repayment burden under a ML is usually greater (at least for realistic values of η and γ).

The second part of the article presents some additional results and complements the former. In this section, I simulate welfare and labor supply responses under the two loan schemes of interest assuming that wages evolve across time following a Geometric Brownian Motion process. The parameters of the model are calibrated using real Brazilian data. This exercise suggests that changing the FIES to an ICL shall (i) decrease labor supply; (ii) increase graduates' welfare; (iii) reduce repayment burdens, and (iv) increase the number of years until the debt is fully repaid. Also, these effects tend to be greater for women and non-white persons, who have lower initial earnings and with lower growth throughout their working lifetimes, and who also face greater unemployment risks. Therefore, shifting to an ICL is especially welfare-enhancing for these groups of individuals. Finally, this study indicates that implementing a grace period after graduation favors a ML scheme (in terms of welfare-enhancement), while introducing a repayment threshold and positive real interest rate on the loans favors an ICL; and also that, under the FIES rules, shifting to an ICL is a more welfare-enhancing move to students from more expensive colleges and with shorter degrees.

The results obtained in the simulations give us valuable information on the benefits of choosing an ICL over a ML design that could be used to enhance the efficiency and equality of the higher education funding system. First, the simulations indicate that shifting from a ML to an ICL scheme leads graduates to obtain greater increases in utility at their early career stages, by reducing repayment burdens and allowing these individuals to work a lesser

number of hours. Also, this shift possibly increases graduates' freedom of choosing their post-graduation paths with fewer constraints arising from their student debts – although this dimension has not been properly analyzed in the present work and could hence be the object of future studies. Second, it is important to emphasize that these benefits tend to be even greater in developing nations, where, on average, real interest rates are higher and early unemployment is a more present threat. Third, shifting to an ICL scheme also increases the progressivity of the system, since it favors individuals who are more economically vulnerable.

Finally, a few caveats are in order. As with most microsimulation models used for prediction and policy analysis, the results of these exercises must be taken with considerable circumspection. The WTP, in particular, which indicates whether an ICL is preferred over an ML, might be different for different economies, depending on the set of parameters employed. In any case, the results of the comparative statics exercises are universal. That is, regardless of the economy under study, an ICL scheme would become more attractive as real interest rates rise, or as the income of recent graduates decline, and so forth. In this regard, the present study provides important contributions to the literature on methods to finance higher education, with a series of directions and guidelines to educators and policymakers, especially on how to structure their student financing system in a way to make it more equitable and less burdensome for the students.

5. Conclusion

The inequality of access in HE has become a persistent problem in many developed and emerging economies. The barriers faced by vulnerable groups to be able to participate in tertiary education perpetuate inequalities in various social dimensions, such as gender, race/ethnicity, and by income, as well as resonate in many layers of society.

In this dissertation, I investigate three different Brazilian policies that aim to tackle precisely this problem: (i) a law stating that a specific share of seats in Brazilian federal higher education institutions must be filled by non-white students, who have historically been left on the sidelines of the educational system (the Law of Quotas); (ii) a program that grants full and partial scholarships to students from low-income families (the Prouni); and (iii) a student funding program targeted to low-income individuals and with special credit conditions for those facing greater socioeconomic vulnerabilities (the FIES).

Although the three chapters contained in this dissertation are independent and deal with different problems, they essentially address the same topic: how to efficiently overcome the challenge of inequality in HE through public policy. Besides, ultimately, the three chapters aim to ask the same question: May these – or similar - policies serve as a guide to educators and policy makers who wish to increase the efficiency and equality of opportunity of their educational systems?

In chapter 2, I show that the Law of Quotas induced non-white students to attain higher scores in the high school exit exam (the ENEM). In this sense, this chapter provides evidence that affirmative action in education – in this case, a reserved number of seats in higher education institutions for specific racial groups – provide positive incentives for ex-ante human capital accumulation. Therefore, not only has the Law of Quotas increased the equality of Brazil's tertiary education directly through an increased number of seats for blacks, browns and indigenous, but it has also generated efficiency gains, encouraging students from these racial groups to close the performance gap with white students by the end of secondary education.

In chapter 3, I provide evidence that the Prouni had a positive and significant effect on the HE participation rate of low-income individuals, increasing the

odds of attending HE by 20% and 37% for those entitled to the partial and full scholarships, respectively, by 2007 – which, in turn, entailed an increase in the HE enrolment rate of these individuals of 3.4 and 1.4 percentage points. Moreover, this study estimates that every USD 1,000 per student spent by the Prouni increased the HE participation rate by approximately 11% to 12%, which is in line with the findings from developed economies, and that these effects were greater for non-white persons and women, whose HE attendance seem to be more sensitive to income restrictions.

Lastly, in chapter 4, among the various implications of changing the higher education funding system from a mortgage loan to an income-contingent loan design, I highlight the following: (i) it induces a decrease in labor supply; (ii) it increases graduates' welfare in their early careers; (iii) it reduces repayment burdens, and (iv) it increases the number of years until the debt is fully repaid. Furthermore, these effects tend to be greater for women and non-white individuals, who have lower initial earnings and with lower growth throughout their working lifetimes, and who also face greater unemployment risks.

This dissertation reinforces the importance of public policy as a crucial tool to generate equality of opportunity in access to HE. The three policies analyzed in this research not only have proven to be important engines of equality in education and to actively contribute to achieve the primary goal of enhancing participation in HE, but have also generated relevant and positive side effects. For instance, the introduction of quotas for vulnerable groups in higher education promptly facilitates the access of these groups in the tertiary system directly through an increased number of reserved seats. However, this type of affirmative action policy also has an indirect contribution to the reduction of inequality in education, once it generates positive incentives for these vulnerable groups to increase their investment in human capital accumulation during primary and secondary school levels. As suggested by the extensive theoretical literature on this subject, these positive effects are most likely due to the mitigation of the discouragement effect, since students who benefit from the quotas are dislocated to the margin of selection. In this sense, this research stresses the importance of quotas in education, which today are being implemented only in a handful of economies, as a motor for equality and efficiency, especially in countries with a high level of social segregation.

In regard to income-based scholarships, the present research corroborates the findings from the existing literature on developed economies, as it finds positive and significant effects of the scholarships on HE participation – which, in turn, are similar to those found in the UK and the US. Moreover, an additional implication of the implementation of these scholarships is that its effects are greater for groups whose HE attendance is more sensitive to income restrictions, possibly because these groups have been historically excluded from the educational system – in Brazil, women and non-white persons. Therefore, the introduction of income-based scholarships not only enhances access of low-income individuals in HE, but it also helps to reduce other social inequalities, such as by race/ethnicity and gender.

Chapter 4 emphasizes the role of student financial schemes targeted to low-income individuals on mitigating educational inequalities, and it also stresses its effects on welfare and on the labor market, as well as the importance of selecting the appropriate method and rules for credit concession. The main difference between the program analyzed in chapter 4 from the ones evaluated in chapters 2 and 3 is that the effects of student funding directly reverberate in the students' early careers, as they face substantial repayment burdens. Consequently, as much as these student loans indeed encourage and help individuals from lower social classes to access tertiary education, policy makers must consider the effects of these loans on graduates' welfare and early career choices when designing the terms of the credit line. The research from this chapter indicates that an income-contingent loan provides much more benefits for early graduates than a mortgage loan scheme (which is used in several countries), inducing them to achieve higher utility levels, and allowing them to work fewer hours.

The three policies and programs analyzed in this dissertation were implemented in Brazil in recent years (including the discussed change in design of the higher education funding system to an income-contingent loan). All of them had positive effects that helped and still help to combat the substantial inequality in the country's higher education system, and, as above-mentioned, also spawned positive by-products on human capital accumulation, gender and racial/ethnical inequality and student's welfare. I hope that this research contributes to the literature on the evaluation of public policy in education and, hence, might serve as reference to educators and

policy makers worldwide, especially from emerging economies with similar problems to Brazil.

Finally, I must highlight that this study contains a number of limitations, and future research can – and should - be developed to overcome them. In terms of methodology, since the strategies employed in chapters 2 and 3 differ substantially from those used in chapter 4, it is necessary to assess their limitations separately. Let us begin with the causal inference technique from chapter 2 and 3. First, I have controlled for sets of observed variables, while others remained as non-observable; and second, in both cases the analyzed timespan had to be limited in order to prevent concomitant policies to pollute the results. Anyway, various robustness exercises and sensitivity analysis have been performed in both studies so as to strengthen the validity of the inferences.

In relation to chapter 4, it is important to recall that, as with most microsimulation exercises, these results must be treated with due caution. Moreover, I have built and simulated policy exercises in a partial equilibrium model, in which the students were at the focus of the investigation. I have not, however, evaluated the effects of changing from a mortgage loan to an income-contingent loan scheme on public finance, neither on economic growth nor on the level of activity (the latest could indeed be harmed by a change to an income-contingent loan design, due to the estimated reduction in labor supply).

Furthermore, an additional issue that must be discussed regards the universality (external validity) of the results. Even though the findings from this thesis can be used as valuable inputs for policy makers from around the world, let us remember that Brazil was used as a laboratory for the investigations and, consequently, the cultural and socioeconomical idiosyncrasies from the country might limit how general these results are.

The findings and limitations from this dissertation leave an avenue for future research. A first idea would be to explore the effects of these – or similar – policies on different educational and labor market variables or even on alternatives outputs that, according to the existing literature, are affected by education, such as crime/violence, civil/political engagement, and democracy. Second, although many positive contributions of these policies were shown and discussed throughout this thesis, there is room for further

investigation on how these policies might be optimized: for instance, would it make sense to integrate racial/ethnic quotas with some sort of merit-based program to further enhance the positive incentive effects provided? Could student loan programs have gender-specific rules, as there is for race/ethnicity, since women's HE attendance tends to be more income-restricted than men's? And so on. Third, many of the analyzed policies have also been implemented in a similar fashion in other countries, making room for comparative education exercises, which could thus answer the question on the universality of the results. Finally, the model from chapter 4 could also be expanded to enable the simulation of (i) additional (even hypothetical) loan designs, and (ii) the effects on different outputs, such as public finance variables and economic growth.

Bibliography

Abraham, K. G., and Clark, M. A. (2006). "Financial aid and students' college decisions: Evidence from the District of Columbia Tuition Assistance Grant program". The Journal of Human Resources, 41(3), 578-610.

Acemoglu, D. (2002). "Technical change, inequality, and the labor market". Journal of Economic Literature, 40, 7-72.

Akhtari, M., Bau, N., and Laliberté, J-W (2020). "Affirmative Action and Pre-College Human Capital", NBER Working Paper 27779.

Alon, S. (2009). "The evolution of class inequality in higher education: Competition, exclusion and adaptation". American Sociological Review, 74, 731-55.

Altonji, J., Elder, T. and Taber, C. (2005). "Selection on observed and unobserved variables: assessing the effectiveness of catholic schools". Journal of Political Economy, 113(1), 151-184.

Antonovics, K., and Backes, B. (2014). "The effect of banning affirmative action on human capital accumulation prior to college entry", IZA Journal of Labor Economics, vol. 3(1), p. 5.

Antunes, A., Cavalcanti, T., and Villamil, A. (2014). "The effects of credit subsidies on development". Economic Theory, 58, 1-30.

Araújo, E. (2005). "Avaliando três especificações para o fator de desconto estocástico através da fronteira de volatilidade de Hansen e Jagannathan: um estudo empírico para o Brasil". Pesquisa e Planejamento Econômico, 35(1).

Arcidiacono, P. (2005). "Affirmative action in higher education: How do admission and financial aid rules affect future earnings?". Econometrica, 73, 1477–1524.

Assunção, J. and Ferman, B. (2015). "Does affirmative action enhance or undercut investment incentives? evidence from quotas in Brazilian public universities". Unpublished Manuscript.

Atuahene, F. (2008). "The challenge of financing higher education and the role of student loan schemes: an analysis of the student loan trust fund (SLTF) in Ghana". Higher Education, 56, 407-21.

Barr, N., Chapman, B., Dearden, L., and Dynarski, S. (2019). "The US college loan system: Lessons from Australia and England". Economics of Education Review, 71, 32-48.

Barros, R. et al. (2010). "Determinantes da queda na desigualdade de renda no Brasil". Brasília: Ipea, Discussion Paper.

Baumgartner, H. and Steiner, V. (2006). "Does more generous student aid increase enrolment rates into higher education? Evaluating the German student aid reform of 2001". IZA Discussion Papers 2034, Institute for the Study of Labor (IZA).

Becker, K. L. and Mendonça, M. J. C. (2019). "Avaliação de impacto do Prouni sobre a performance acadêmica dos estudantes". Brasília: Ipea, Discussion Paper.

Belfield, C., Britton, J., and van de Erve, I. (2017). "Higher education finance reform: Raising the repayment threshold to £25,000 and freezing the fee cap at £9,250". IFS Briefing Note BN217.

Binelli, C., Meghir, C. and Menezes-Filho, N. (2008). "Education and Wages in Brazil". University College London (mimeograph).

Blomqvist, S. (1983). "The effect of income taxation on the labor supply of married men in Sweden". Journal of Human Resources, 22(2), 169-97.

Blomqvist, S. and Hansell-Brusewitz, U. (1990). "The effect of taxes on male and female labor supply in Sweden". Journal of Human Resources, 25(3), 317-57.

Blundell, R., Duncan, A., and Meghir, C. (1998). "Estimating labor supply responses using tax reforms". Econometrica, 66(4), 827-61.

Bourguignon, F. et al. (2003). "Conditional cash transfers, schooling, and child labor: Micro-simulating Brazil's Bolsa Escola program". The World Bank Economic Review, 17(2), 229-54.

Britton, J., Van der Erve, L., and Higgins, T. (2019). "Income-contingent student loan design: Lessons from around the world". Economics of Education Review, 71, 65-82.

Cameron, S., and Heckman, J. J. (2001). "The dynamics of educational attainment for Black, Hispanic, and White males". Journal of Political Economy, 109, 455-99.

Canton, E. and Blom, A. (2004). "Can student loans improve accessibility to higher education and student performance? An impact study of the case of SOFES, Mexico". Working Paper S3425, Washington, DC, The World Bank.

Carneiro, P. and Heckman, J. J. (2002). "The evidence of credit constraints in post-secondary schooling". Economic Journal, 112, 705-34.

Chapman, B., Higgins, T. and Stiglitz, J. (2014). "Income contingent loans: Theory, practice and prospects). Basingstoke and New York: Palgrave Macmillan Ltd.

Chapman, B., and Lounkaew, K. (2010). "Income contingent student loans for Thailand: Alternatives compared". Economics of Education Review, 29, 695-709.

Chapman, B., Lounkaew, K., Polsiri, P., Sarachitti, R., and Sitthipongpanich, T. (2010). "Thailand's student loan fund: Interest rate subsidies and repayment burdens". Economics of Education Review, 29, 685-694.

Chapman, B., and Sinning, M. (2014). "Student loan reforms for German higher education: Financing tuition fees". Education Economics, 22(6), 569-88.

Cheadle, J. (2008). "Educational investment, family context, and children's math and reading growth from kindergarten through third grade". Sociology of Education, 81, 1–31.

Chen, S. H. (2008). "Estimating the variance of wages in the presence of selection and unobservable heterogeneity". Review of Economics and Statistics, 90(2), 275-89.

Chesters, J., and Watson, L. (2013). "Understanding the persistence of inequality in higher education: Evidence from Australia". Journal of Education Policy, 28(2), 198-215.

Coate, S. and Loury, G. (1993). "Will affirmative-action policies eliminate negative stereotypes?". American Economic Review 83 (5), 1220–1240.

Cornwell, C., et al. (2006). "The enrollment effects of merit-based financial aid: Evidence from Georgia's HOPE program". Journal of Labor Economics, 24(4).

Cotton, C. Hickman, B. and Price, J. (2016). "Affirmative action and human capital investment: theory and evidence from a randomized field experiment". Queen's Economics Department Working Paper 1350.

Dale, S. and Krueger, A. (2002) "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables", Quarterly Journal of Economics 117(4), 1491-1527.

Dale, S. and Krueger, A. (2014) "Estimating the Return to College Selectivity over the Career Using Administrative Earning Data", Journal of Human Resources 49(2), 323-358.

Deardon, et al. (2014). "Money for nothing: Estimating the impact of student aid on participation in higher education". Economics of Education Review, 43, 66-78.

Dearden, L., and Nascimento, P. (2019). "Modelling alternative student loan schemes for Brazil". Economics of Education Review, 71, 83-94.

Department of Education and Department of Justice (2011). "Guidance on the voluntary use of race to achieve diversity in postsecondary education". Retrieved September 10, 2020, from https://www2.ed.gov/about/offices/list/ocr/docs/guidance-pse-201111.html

Dynarski, S. M. (2000). "Hope for whom? Financial aid for the middle class and its impacts on college attendance". National Tax Journal, 53(Part 2), 629-61.

Dynarski, S. M. (2003). "Does aid matter? Measuring the effect of student aid on college attendande and completion". The American Economic Review, 93(1), 279-88.

Egerton, M. (2002). "Higher education and civic engagement". British Journal of Sociology, 53, 603-20.

Egwakhe, A. and Osabuohlen, E. (2009). "Educational Background and Youth criminality in Nigeria". International Forum, 12(1).

ENEM's Official Act 2020: "*Edital do ENEM*". Retrieved November 4, 2020, from http://download.inep.gov.br/educacao_basica/enem/edital/2020/edital_enem2020_digital.pdf

ENEM's Syllabus: "*Matriz de Referência ENEM*". Retrieved November 4, 2020, from http://download.inep.gov.br/download/enem/matriz_referencia.pdf

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

Fajardo, J., Ornelas, J., and Farias, A. (2012). "Estimating risk aversion, risk-neutral and real-world densities using Brazilian real currency options". Economia Aplicada, 16(4).

Francis, A.. and Tannuri-Pianto, M. (2012). "Using Brazil's racial continuum to examine the short-term effects of affirmative action in higher education". Journal of Human Resources 47(3), 754–784.

Furstenberg, E. (2003). "Affirmative action, incentives, and the black-white test score gap". College William and Mary Department of Economics Working Paper 3.

García-Peñalosa, C. and Turnosky, S. (2006). "Growth and income inequality: A canonical model". Economic Theory, 28(1), 25-49.

Gimenez-Nadal, J., and Sevilla, A. (2012). "Trends in time allocation: A cross-country analysis". European Economic Review, 56(6), 1338-59.

Gurgand, et al. (2011). "Student loans: Liquidity constraint and higher education in South Africa". Paris School of Economics Working Paper 2011-20.

Hainmueller, J. (2012). "Entropy balancing for causal effects: a multivariate reweighting method to produce balanced samples in observational studies". Political Analysis, 20(1), 25-46.

Higham, D. (2001). "An algorithm introduction to the numerical simulation of stochastic differential equations". SIAM Review, 43, 894-912.

Ibarra, R. (2001). "Beyond affirmative action: reframing the context of Higher Education". Madison: University of Wisconsin Press.

Imbens, G.W., and Kalyanaraman, K. (2012). "Optimal bandwidth choice for the regression discontinuity estimator". Review of Economic Studies, 79, 933-59.

Issler, J., and Piqueira, N. (2000). "Estimating relative risk aversion, the discount rate, and the intertemporal elasticity of substitution in consumption for Brazil using three types of utility function". Brazilian Review of Econometrics, 20(2), 201-39.

Kane, T.J. (2003). "A quase-experimental estimate of the impact of financial aid on college-going". NBER Working Paper 9703.

Kane, T. J. (2007). "Evaluating the impact of the D.C. Tuition Assistance Grant program". The Journal of Human Resources, 42(3), 555-82.

Keane, M. P., Wolpin, K. I. (2001). "The effect of parental transfers and borrowing constraints on educational attainment". International Economic Review, 42, 1051-1103

Keane, M. (2011). "Labor supply and taxes: A survey". Journal of Economic Literature, 49(4), 961-1075.

Kelly, P. (2005). "As America becomes more diverse: The impact of state higher education inequality". Boulder, CO: National Center for Higher Education Management Systems (NCHEMS)

Khanna, G. (2020). "Does affirmative action incentivize schooling? evidence from India". The Review of Economics and Statistics, 102(2), 219-233.

Knight, J. and Hebl, M. R. (2005). "Affirmative reaction: the influence of type of justification on nonbeneficiary attitudes toward affirmative action plans in higher education". Journal of Social Issues, 61(3), 547-568.

Larraín, C., and Zurita, S. (2008). "The new student loan system in Chile's higher education". Higher Education, 55, 683-702.

Leal, G., and Turchick, D. (2020). "Labor supply and the no-growth trap". FEA/USP Department of Economics Working Paper 2019-54.

Lepine, A. (2018). "Financial aid and student performance in college: Evidence from Brazil". Brazilian Review of Econometrics, 38(2), 221-61.

Lisboa, M. and Menezes-Filho, N.A. (2001). "Microeconomia e sociedade no Brasil". Rio de Janeiro, Ed. Contra Capa.

Long, B.T. (2004). "The impact of federal tax credits for higher education expenses". C.M. Hoxby (ed.), College Choices: The Economics of Where to Go, When to Go, and How to Pay for It (pp. 101–168). Chicago: University of Chicago Press.

Loury, G. (1992). "Incentive effects of affirmative action". Annals of the American Association of Political and Social Science, 523, 19–29.

Loury, L. and Garman, D. (1993). "Affirmative action in higher education." American Economic Review, 83(2), 99–103.

Loury, L. and Garman, D. (1995). "College selectivity and earnings." Journal of Labor Economics, 13(2), 289–308.

Lundberg, S. and Startz, R. (1983). "Private discrimination and social intervention in competitive labor markets". American Economic Review, LXXIII, 340-47.

Lynch, K., and O'riordan, C. (1998). "Inequality in Higher Education: A study of class barriers". British Journal of Sociology of Education, 19, 455-78.

Marteleto, L. (2012). "Educational inequality by race in Brazil, 1982-2007: structural changes and shifts in racial classification." Demography, 49, 337-358.

Mello, U. (2019). "Essays in educational and integrational inequality". PhD dissertation presented at Universidad Carlos III de Madrid.

Metcalf, H. (2003) "Increasing inequality n higher education: The role of term-time working". Oxford Review of Education, 29(3), 315-29.

Migali, G. (2012). "Funding higher education and wage uncertainty: Income contingent loan versus mortgage loan". Economics of Education Review, 31, 871-89.

Milgrom, P. and Oster, S. (1987). "Job discrimination, market forces, and the invisibility hypothesis". The Quarterly Journal of Economics, 102(3), 453-476.

Neves, C.B. (2015). "Demand and supply for higher education in Brazil". In: Schwartzman, S., Pillay, P., Pinheiro, R., Higher education in the BRICS countries: investigating the pact between higher education and society. Springer, 2015.

Nguyen, H. (2020). "Free college? Assessing enrollment responses to the Tennessee Promise program". Labour Economics, 66, 101882.

Nielsen, et al. (2010). "Estimating the effect of student aid on college enrollment: Evidence from a government grant policy reform". American Economic Journal: Economic Policy, 2(2), 185-215.

O'Gorman, M. (2010). "Educational disparity and the persistence of the black-white wage gap in the U.S.". Economics of Education Review, 29, 526-542.

Oreopoulos, P., and Petronijevic., U. (2011). "Making college worth it: A review of research on the returns to higher education" NBER Working Paper 19053.

Osorio, R. and Souza, P. (2012). "O Bolsa Família depois do Brasil Carinhoso: uma análise do potencial de redução da pobreza". Brasília: Ipea, Discussion Paper.

Oster, E. (2019). "Unobservable selection and coefficient stability: theory and evidence". Journal of Business & Economic Statistics, 37:2, 187-204.

Padula, M., and Pistaferri, L. (2001). "Education, unemployment and wage risk". Working Paper No. 67. Italy: Centre for Studies in Economics and Finance (CSEF), University of Salerno.

Pereira, J., Bittencourt, M., and Braga, B. (2015). "Affirmative action in higher education: impacts of the national exam in Brazil". 55th Congress of the European Regional Science Association: World Renaissance: Changing roles for people and places.

Perry, L. and McConney, A. (2013). "School socioeconomic status and student outcomes in reading and mathematics: A comparison of Australia and Canada". Australian Journal of Education, 52(2), 124–140.

Queiroz, Z., Miranda, G., Tavares, M. and Freitas, S. (2015). "A Lei de Cotas na perspectiva do desempenho acadêmico na Universidade Federal de Uberlândia". *Revista Brasileira de Estudos Pedagógicos*, 96(24), 299-320.

Rambachan, A. and Roth, J. (2019). "An honest approach to parallel trends." Unpublished Manuscript.

Roser, M., and Ortiz-Ospina, E. (2013). "Tertiary Education". Published Online at ourworldindata.org. Retrieved from: 'https://ourworldindata.org/tertiary-education'.

Rimm-Kaufman, S., Fan, X., Chiu, Y.J. and You, W. (2007). "The contribution of the responsive classroom approach on children's academic achievement: Results from a three year longitudinal study". Journal of School Psychology, 45, 401–421.

Rothstein, J., and Rouse, C.E. (2011). "Constrained after college: Student loans and early-career occupational choices". Journal of Public Economics, 95, 149-63.

Rothstein, J. and Yoon, A. (2008a). "Mismatch in law school". NBER Working Paper 14275.

Rothstein, J. and Yoon, A. (2008b). "Affirmative action in law school admissions: what do racial preferences do?" NBER Working Paper 14276.

Saeme, D. (2014). "Does the implementation of affirmative action in a competitive setting incentivize underrepresented public school applicants performance? Evidence from São Paulo". The Yale Journal of Economics, 2, 93–114.

Sampaio, B. and Guimarães, J. (2009). "Diferenças de eficiência entre ensino público e privado no Brasil". Revista de Economia Aplicada, 13(1), 45-68.

Shadish, W.R., Cook, T.D., and Campbell, D. T. (2002). "Experimental and quasi-experimental designs for generalized causal inference". Neww York: Houghton Mifflin.

Soares, F. et al. (2010). "Evaluating the impact of Brazil's Bolsa Família". Latin American Research Review, 45(2), 173-90.

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

Sui-Chu, E. and Willms, J. (1996). "Effects of parental involvement on eight-grade achievement". Sociology of Education, 69(2), 126–147.

Tangkitvanich, S. and Manasboonphempool, A. (2010). "Evaluating the student loan fund of Thailand". Economics of Education Review, 29(5), 710-721.

Turnovsky, S. (2000). "Fiscal policy, elastic labor supply, and endogenous growth". Journal of Monetary Economics, 45(1), 185-210.

Vidigal, C. (2018). "Racial and low-income quotas in Brazilian universities: impact on academic performance". Journal of Economic Studies, 45(1).

Walker, J., Pearce, C., Boe, K., and Lawson, M. (2019). "The power of education to fight inequality: How increasing educational equality and quality is crucial to fighting economic and gender inequality". Oxfam Briefing paper.

Walker, I. and Zhu, Y. (2018). "University selectivity and the relative returns to higher education: Evidence from the UK." Labour Economics, 53, 230–249.

Xu, K. (2017). "Regression discontinuity with categorical outcomes". Journal of Econometrics, 201(1), 1-18.

Zhang, L. (2013). "Effects of educational debt on graduate school attendance and early career and lifestyle choices". Education Economics, 21:2, 154-75.

Zhao, Q. and Percival, D. (2017). "Entropy Balancing is doubly robust." Journal of Causal Inference, 5(1).

Ziliak, J.P., and Kniesner, T. J. (2005). "The effect of income taxation on consumption and labor supply". Journal of Labor Economics, 23(4), 769-96.