Felipe Ferreira de Barros

Quotas in Higher Education: An Academic Achievement and Labor Market Analysis

Juiz de Fora

September, 2019

Felipe Ferreira de Barros

Quotas in Higher Education: An Academic Achievement and Labor Market Analysis

Federal University of Juiz de Fora

Applied Economics - Graduate Program

Supervisor: Ricardo da Silva Freguglia

Co-supervisor: Fernanda Gonçalves de la Fuente Estevan

Juiz de Fora

September, 2019

Ficha catalográfica elaborada através do programa de geração
automática da Biblioteca Universitária da UFJF,
com os dados fornecidos pelo(a) autor(a)

Ferreira de Barros, Felipe.
Quotas in Higher Education: : An Academic Achievement and Labor Market Analysis / Felipe Ferreira de Barros. -- 2019.
63 p.

Orientador: Ricardo da Silva Freguglia
Coorientadora: Fernanda Gonçalves de la Fuente Estevan
Dissertação (mestrado acadêmico) - Universidade Federal de Juiz de Fora, Faculdade de Economia. Programa de Pós-Graduação em Economia, 2019.

1. Affirmative Action. 2. Quotas. 3. Labor Market. 4. cademic Achievement. 5. Higher Education. I. da Silva Freguglia, Ricardo , orient. II. Gonçalves de la Fuente Estevan , Fernanda, coorient. III. Título.

FELIPE FERREIRA DE BARROS

QUOTAS IN HIGHER EDUCATION: AN ACADEMIC ACHIEVEMENT AND LABOR MARKET ANALYSIS

Dissertação apresentada ao Programa de Pós-graduação em Economia, da Universidade Federal de Juiz de Fora como requisito parcial a obtenção do grau de Mestre em Economia Aplicada. Área de concentração: Economia.

Aprovada em: 07/06/2019

BANCA EXAMINADORA

Prof. Dr. Ricardo da Silva Freguglia – Orientador
Universidade Federal de Juiz de Fora (UFJF)

Proª. Drª. Fernanda Gonçalves de la Fuente Estevan – Coorientadora
Fundação Getulio Vargas (FGV)

Profª. Drª. Laura de Carvalho Schiavon
Universidade Federal de Juiz de Fora (UFJF)

Prof. Dr. Naércio Menezes Filho
Instituto de Ensino e Pesquisa (INSPER)

Abstract

The recent adoption of quota policies by Brazilian Universities, and the public debate on the topic makes it a relevant subject for empirical studies. The Federal University of Juiz de Fora - UFJF - one of the largest public universities in Minas Gerais, adopted the quota program in 2006. Using an individual level database for all incoming students from 2003 to 2010, this study aims to investigate (descriptively) students' academic and labor market performance before and after the quotas. To analyze academic achievement, we control for performance of students relative to their classmates. We also separate by majors that require high entrance exam score (selective), and majors that require low entrance exam score (non-selective). To analyze performance in the labor market, we link the database provided by UFJF to the RAIS database. Preliminary results indicate that quota students do not improve their relative position in selective majors, however in non-selective majors the quota students catch-up with the non-quota students. Besides the catch-up, quota students and non-quota students are close to each other, in terms of grades. In the labor market, the results indicate that quota students tend to receive a lower salary, 18% less, compared to non-quota students. Regarding the probability of finding a job, we found that quota students are more likely to find a job during graduation, 14% higher. However, the likelihood of finding a job after graduation is lower for quota students, 13% less than for non-quota students.

Key-words: Affirmative Action, Quotas, Higher Education, Academic Achievement, Labor Market, Wages.

Resumo

Devido à recente adoção da política de cotas e ao debate público com relação ao tema, fazem deste, um tema relevante para estudos empíricos. A Universidade Federal de Juiz de Fora - UFJF- uma das maiores universidades públicas de Minas Gerais adotou o programa de cotas em 2006. Com dados a nível individual de todos os estudantes ingressantes no período de 2003 a 2010, este trabalho visa investigar (somente de forma descritiva) o desempenho acadêmico e desempenho no mercado de trabalho dos estudantes antes e depois das cotas. Para analisar o desempenho acadêmico, controlamos para o desempenho do aluno com relação aos seus colegas de classe, em somente matérias obrigatórias separando por cursos que exigem alta nota de entrada (seletivos) e cursos que exigem uma menor nota de entrada (não seletivos). No mercado de trabalho, usamos a base de dados da UFJF juntamente com a RAIS. Os resultados indicam que alunos cotistas não ultrapassam, em termos de nota, alunos não cotista em cursos seletivos, no entanto em cursos não seletivos os cotistas se equiparam aos não cotistas. Estatísticas descritivas mostram também que alunos cotistas se aproximam dos alunos não cotistas, em termos de notas. Já no mercado de trabalho os resultados indicam que cotistas tendem a receber um salário menor, 18% menor, em relação à não cotistas. Com relação à probabilidade de encontrar emprego, resultados indicam que cotistas tem maior chance de encontrar emprego durante a graduação, 14% maior. No entanto, a probabilidade de se econtrar emprego depois da graduação é menor para cotistas, 13% menor, em relação à não cotistas.

Palavras-chaves: Ação Afirmativa, Cotas, Educação Superior, Performance Acadêmica, Mercado de Trabalho, Salário.

Contents

Introduction

Affirmative Action policies are used in many countries as a way to compensate historically disadvantaged groups in educational and labor market settings. In Brazil, affirmative action policies in higher education are represented by preferential admission in college. The Quota Law (*Lei de Cotas*)[1]

from 2012 makes it mandatory for public universities to reserve 50% of their vacancies to minority groups. However, as an individual effort, many universities adopted the quota system before the *Lei de Cotas*, in particular the Federal University of Juiz de Fora (*UFJF*), which is the one we analyze in this study.

In this paper we investigate the academic performance, represented by grades, of quota and non-quota students. We follow those students into the labor market and investigate wage differential between quota and non-quota students. We also investigate if there are differences in the probability of finding a job, for quota and non-quota students.

The vigorous debate about affirmative action in higher education becomes more vivid in the academy, especially because data availability. The debate about quotas is varied. Some papers analyze what happens to the students before the university, for example, changes in the effort of the high school students (FERMAN; ASSUNÇÃO, 2005). Others focus on the students' performance after graduation, for example, labor market outcomes (FRANCIS; TANNURI-PIANTO, 2016; BERTRAND; HANNA; MULLAINATHAN, 2010; ALTONJI; ARCIDIACONO; MAUREL, 2016). In this study, we are interested in what happens during college; and after college, analyzing labor market outcomes.

Because of the disadvantaged background of quota applicants some critics claim that quota students would not be able to catch-up with students with more advantaged background. (ARCIDIACONO; AUCEJO; SPENNER, 2012). The mismatch hypothesis is centered in the debate raised by the catch-up issue. It tests whether or not quota students would be better off if they got admitted in less selective universities than those universities that they can get in now because of the quotas. If the quota students are less prepared, they could struggle throughout their major, which can increase the drop

[1] To access the oficial document please refer to http://portal.mec.gov.br/cotas/docs/decreto_7824.pdf

out rate. And when quota students graduate, they probably graduate with lower grades. These two aspects could affect the individual's labor market outcomes (ALON; TIENDA, 2005).

In this study, we are not able to test the mismatch hypothesis directly because the lack of data on those who did not enroll at UFJF. However, we will be able to identify differences in outcomes for quota and non-quota students. In sum, we do here can be seen as mismatch analysis in the spirit of Rothstein and Yoon (2008). In practice, they compare black and white students, from a law school, with similar credentials and they use bar passage as the outcome. Any difference in outcomes can be seen as a mismatch of students with lower credentials attending selective universities. The authors also point that this mismatch estimation is overestimated since many unobserved effects can also explain the differences in outcomes.

Prior research shows that students from disadvantaged background, have substantial gaps in educational achievement. This gap can initiate in kindergarten, and this gap can grow throughout the school years (COLEMAN, 1968; JR; LEVITT, 2006; ARCIDIACONO; AUCEJO; SPENNER, 2012).

This thesis assembles a rich database at the individual level of all *admitted* students at the Federal University of Juiz de Fora, one of the largest universities in the state of Minas Gerais. The database contains information on more than 18,000 students admitted between 2003 and 2010.

The motivation of this work is based on the importance of bringing scientific evidence to the debate, especially for Brazil, where data restriction imposes difficulty on studying affirmative action. Since there is a literature that indicates that quota students fall behind their non-quota counterparts, both in terms of grades and wage, and a literature that points otherwise, we want to contribute to the debate by analyzing what seems to be happening at UFJF.

We have separated the data into two parts that are intimately linked, academic performance, and labor market performance. First, to analyze the academic performance of students, we separated this analysis into two cohorts, before and after the quotas. Before the quotas, we compare students with high achievement in the entrance exam score with students with low achievement in the entrance exam score. We find that students in the lowest percentile in the entrance exam score do not catch-up with students in the highest percentiles of the entrance exam score.

Moreover, we compare quota and non-quota students, and we find that quota students are not able to catch-up with their non-quota classmates when we control for course selection. But, in non-selective majors quota students with high performance, tend to catch-up.

In the second part, we analyze differences in outcomes for quota and non-quota students, we use two different outcomes: wages and probability of finding a job (during and after college). We find that quota students tend to receive a lower salary, 18% less, compared to non-quota students. Regarding the probability of finding a job, we found that quota students are more likely to find a job during graduation, 14% higher. However, the likelihood of finding a job after graduation is lower for quota students, 13% less than for non-quota students.

This study is organized as follow: Section 1 discusses the institutional background, section 2 brings the related literature, section 3 presents the empirical strategy, the sections 4 and 5 present the results, and section 6 brings the conclusion.

1 Institutional Framework

In this section we present the institutional background in what all the analysis provided by this work relies upon.

1.1 A Brief History of Affirmative Action

Affirmative actions are policies that aim to ensure some "compensation" for minority groups, such as women, blacks, disadvantaged socioeconomic groups. The term, affirmative action (AA) was conceived in the 1960s in the United States, more precisely in 1961, through the "Executive Order nº 10925" by the American president at the time, John F. Kennedy. In the beginning, Kennedy's affirmative action aimed the labor market, in which companies that had some contract with the government had to take measures to grant access of the minorities to the labor market. The president's act ended up by institutionalizing the first affirmative actions policies through the *U.S. Equal Employment Opportunity Commission.* In 1965, President Lyndon B. Johnson, expanded the affirmative actions initiative, by creating a government's department attached to the "Labor Division" (KURTULUS, 2015).

Since its adoption in the U.S, affirmative actions had an immediate effect. Goldstein and Smith (1976) and Heckman and Wolpin (1976) found positive effects on productivity for companies that had contracts with the government and had to implement affirmative actions and hire black workers. Smith and Welch (1984) and Leonard (1984) found that these policies contributed to creating a more diverse body of workers in the company's environment.

However, from the 1980 decade, under Ronald Reagan, there was an inflection regarding the government's support towards affirmative action. The result was more autonomy for the states to decide whether or not to adopt AAs. Thus, some states started to abandon these policies, such AA's. Washington in 1998, Michigan in 2006, Nebraska in 2008, Arizona in 2010, and other states are following this same trend (KURTULUS, 2015).

In the educational environment, AA's took place in a court decision with the *Title VI* from *1968 Civil Rights Act*, which stated that it was an individual decision by the Universities, whether or not to adopt admission process with some share of vacancies reserved to minorities.

1.2 Affirmative Actions in Brazil

The Brazilian educational system is composed of a low quality primary and secondary public education, and a high-quality public higher education, which is free of tuition. One consequence of this educational dichotomy is that students from public schools generally have worse performance in the admission exam. Hence, students from good private schools could enter in the university by benefiting from a better score than those counterparts from public schools.

To diminish the inequality in the admission process emerges the firsts steps towards affirmative action policies. In Brazil, the idea to create equal opportunities to students from disadvantaged background, is translated in quota policies (CARDOSO, 2008). The first quota policy begins in 2000 with the University of Brasília (UnB) where 20% of the vacancies were destined to black candidates (QUEIROZ; SANTOS, 2006). In 2002 other three public universities, State University of Rio de Janeiro (UERJ), State University of North Fluminense (UENF) and the State University of Bahia (UENB), adopted the quota system in their admission process.

From 2012, with law No. 12.711, of August 29, 2012, the 50% of the vacancies for the admission system for quotas became effective at federal level, as it is evident in article 1 of the law that provides that of this 50%, half should be reserved for students from families with income equal to or less than 1.5 minimum wages.

1.3 Admission Process and Affirmative Action at UFJF

1.3.1 The Admission Process at UFJF

The admission process at UFJF is given by two ways of admission: *Vestibular* and *PISM*, that will be explained in detail later. The applicants must choose one of the two methods at the last year of high school. When deciding which way to apply, the student must also select the major he is applying to. The applicant will be competing for the vacancies destined to the major he is applying to.

The *Vestibular* is a two-phase test that comprehends all subject covered in high school. The applicant does the test at the end of the third year of high school. The first phase is composed of multiple-choice questions, identical to all candidates. The multiple-choice question is composed of the following subjects: Portuguese, Chemistry, Mathematics, Biology, Physics, Geography, History, and Literature. The candidates are ranked by their score, in decreasing order, and the cutoff position is the one that represents the triple of vacancies of the major he applied. The second phase is composed of discursive questions, identical to all students, on the same subjects as the first phase. Again the students are ranked, and the cutoff position is the number of vacancies.

The *PISM* is a three-module test, and the score on each module is cumulative. The first module is a test that the applicant does at the end of the first year of high school. The second module is done at the end of the second year and the third module at the end of the third year of high school. The first module comprehends the subjects' content of the first year of high school. The second comprehends the subjects' content of the second year of high school, and the third comprehends the subjects' content of the third year of high school. The tests are composed of eight multiple-choice questions of each topic (Portuguese, literature, math, physics, biology, chemistry, history and geography) and two discursive questions of each subject. The weight of the modules is 0.25, 0.35, and 0.4, respectively.

Note that the students only choose the major they want to apply to in the third module. Also, the student, in the third module, can choose to continue on *PISM* or choose the *Vestibular*.

1.3.2 Quotas at UFJF

The quotas at UFJF started in 2006. In this year, 30% of the vacancies in each major were destined to quota applicants. In 2007 the percentage increased to 40%, and since 2008 the percentage of vacancies destined to quota applicants increased to 50%.

From 2006 until 2010, there were two different types of quotas: for public school students and other for public school and black students. 25% of the percentage destined to quotas were destined to students who self-declared as black.

To be admitted through the quotas, the student must choose whether he wants to apply to *Vestibular* or *PISM* using the quotas or not. If the student decides to apply for quotas, he needs to determine if he wants to apply for quotas for public school students or public school and black students.

A necessary condition, to apply for quotas, is that the student had to attend at least seven years in a public school, elementary school, or high school. Since 2009, this condition changed; the student had to attend at least seven years in a public school, four years or more in elementary school, and the entire high school.

2 Literature Review

The debate about affirmative action that we are interested in can be divided into two main issues. The first issue focuses on how minority students perform in terms of academic achievement, given that they were admitted through preferential admission. The second issue is whether and how much these programs help those who gain admissions, the so-called mismatch hypothesis. Here we are interested in the literature that covers the catch-up and mismatch, and also, we want to review the literature that links AA (Afirmative Action) to labor market outcomes. [1].

The catch-up issue arises from the fact that we are assuming that students admitted with preferential admission are under prepared compared to non-quota students, since they are admitted with lower entrance exam scores. As stated by Frisancho and Krishna (2016), since quota students are expected to have lower performances than their non-quota counterparts, this gap can be seen as a cost. This cost can be reduced depending on the initial differences in performance between general and minority students, and also depending on the fact that if minority students can catch up (KOCHAR, 2010).

In this sense, Frisancho and Krishna (2016) studying an elite engineering class in India, finds that minority students fall behind their counterparts, especially in selective majors. Alon and Malamud (2014) evaluate the impact of shifting from conventional AA to Top 10% rule, where the top 10% students graduating from public high schools were eligible to preferential admission. The authors argue that the likelihood of graduation rose significantly, especially for black students, which means that top 10% students were better prepared and thus were catching up with non-quota students.

Another work that studies the performance gap of students is Sander (2004). He studies the difference in academic performance of different students at a selective law school in the U.S. He finds that average performance of black students is lower than the average performance of white students, and this gap tends to increase as both groups progress through college. He also finds that giving preferential admission to black students at selective schools, lowers the graduation probability of those students.

Arcidiacono, Aucejo and Spenner (2012) analyze the evolution of racial disparities in college. Using administrative data of Duke University, they conclude that black

[1] For a more detailed discussion about AA see Jr, Loury and Yuret (2007), Holzer and Neumark (2000)

students tend to catch up, and the gap falls by half. However, the authors suggest that this analysis could be naive if not controlling for two significant facts. First, the variance of grades that are given falls across time. In other words, professors tend to give higher grades to students and the variance of the grades among students tend to shrink throughout the college yeas. Second, the grading standards differ across courses in different majors.

As stated by Jr, Loury and Yuret (2007), all affirmative action policies yield lower expected performance among the selected students than it does in the absence of the policy. If minority students are not able to close the gap, AA policies that allow them to enroll in more selective colleges/majors can hurt rather than benefit them (FRISANCHO; KRISHNA, 2016). Assuming that academic credentials are somewhat crucial to labor market outcomes, we can infer that students that are not catching up, can have worse performance in the labor market. This is what mismatch tries to measure.

There are several different ways to measure mismatch, but broadly speaking, mismatch evaluates differences in outcomes due to preferential admission. Thus we can measure mismatch using academic outcomes such as graduation rate and labor market outcomes such as wage.

Sander and Jr (2012) examine the mismatch in a law school. A vital part of the author's argument is that AA has a significant effect on which schools minority students will attend, but it has little impact on whether the admitted students will attend school. The main argument is that eliminating AA would increase the number of black students passing the bar exam. Rothstein and Yoon (2008) also find evidence of mismatch for law school students. They argue that the mismatch is only experienced by black students with the weakest entering academic credentials. They attribute half or more of the black-white gap in law school due to differences in entering academic credential that has not to do with the selectivity of the schools that students attend.

Loury and Garman (1995) find that black students in the U.S gets considerable earning gains by attending selective schools. However, this gain is offset for black students because of lower performance both in terms of grades and in probability of graduation.

On the other hand, Alon and Tienda (2005) examines the effects of college selectivity on the likelihood of graduation for minority students. The authors find no evidence of mismatch, so they claim that minority students thrive in their major. They also find evidence that graduation likelihood increase as school selectivity increases. This evidences is also consistent with Bowen and Bok (1998) and Massey and Mooney (2007).

Bertrand, Hanna and Mullainathan (2010) investigates the effect of affirmative action policy in India that reserves 50% of the vacancies to lower-caste groups. In the analysis, the authors regress earnings on college attendance and several other controls. In sum, they find that lower-caste students obtain a positive return to admission, but income gains of displacing students seem to be smaller than income losses of displaced students.

One of the motivations of this study is to bring this debate to the Brazilian context using empirical evidence. To refer to studies that analyze affirmative action in Brazil, but with a different approach that we do in this study see (FERMAN; ASSUNÇÃO, 2005; ESTEVAN; GALL; MORIN, 2016; FRANCIS; TANNURI-PIANTO, 2012). [2]

The most related works with our study is Francis and Tannuri-Pianto (2016) and Arabage and Souza (2016). Francis and Tannuri-Pianto (2016) make use of a sharp regression discontinuity design to evaluate the labor market outcomes of high-performing students in 2004-2005. The authors suggest that the policy of racial quotas mostly improved outcomes for the targeted group. Relative to quota applicants below the cutoff, quota applicants above the cutoff enjoyed an increase in years of education, college completion, and labor earnings. Those gains were concentrated among men and applicants in more selective majors. For the large part, the mismatch was not prevalent. However, there was evidence of mismatch among those quota students in less selective majors. More broadly, the results for both quota and non-quota applicants confirm the importance of college quality. But the fact that economic returns to admission varied widely by area of study may suggest that major is more relevant than the institution.

Arabage and Souza (2016) employ a difference-in-difference regression on data of two large universities in Rio de Janeiro. They use one university as control and the other as treatment. The authors use two variables as labor market outcomes; the probability of being employed and the log of hourly wages. They analyze those outcomes for eligible students in the presence of affirmative action policy and eligible students in the absence of affirmative action policy. They find that there is no difference between these two groups when it comes to the probability of being employed. However, they find evidence of hourly wage differential among these two groups.

Since we are interested in what happens to students during and after college, our main references from this section are Francis and Tannuri-Pianto (2016), Arcidiacono,

[2] Using data from UFJF there are two studies that analyze the quota policy. Please refer to (GAGO, 2016) and Beraldo and Magrone (2013)

Aucejo and Spenner (2012), Bertrand, Hanna and Mullainathan (2010), Frisancho and Krishna (2016).

3 Data and Methodology

3.1 Data

The data we use in this study is an administrative database provided by UFJF. The analysis focuses on students who took the UFJF vestibular exam between 2003 and 2010. The university granted us complete access to admission records for this period, only for admitted students. We have comprehensive information on the national identification number (CPF)[1], race, gender, age, student's name, name of the parents and school attended before college. The institutional records includes entrance exam score, grade on each course taken by the student (which is used to calculate the student's GPA[2]), number of credits (*on hold*, approved, dismissed and reproved), major, type of admission (quotas or regular) and exit status (cancelled, completed, active, *on hold*) on a semester by semester basis.

The CPF is a key variable that makes it possible to follow the students in the labor market. The CPF is a unique identification number. With this identification number, we can merge our database with the database with information on the labor market.

Note that, since we have information on major, admission year and semester we can group the students in their classes, which is interesting, since we can see how the students fare, in terms of grades, respect to their classmates.

We have constructed a panel data where students are observed in each year. Throughout the years, we have students entering and leaving the database. This database design gives us the possibility of following the student's grade along the years.

We have 18,726 students in total, which 19.5% (3,661 students) are quota students, followed throughout their college life. Table 1 shows how many students were admitted at UFJF by year of admission. Note that in 2007 UFJF joined a federal program (REUNI) that helped federal universities to raise the number of vacancies.

---

1 Cadastro de Pessoa Física - CPF is a unique personal identification number in Brazil.
2 GPA is the grade point average

Table 1 – Students Enrolling per year

Year of Enrollment	Students Enrolled	Percent
2003	2,249	12.01
2004	2,207	11.79
2005	2,149	11.48
2006	2,195	11.72
2007	2,161	11.54
2008	2,223	11.87
2009	2,578	13.77
2010	2,964	15.83

Source: Data provided by UFJF, 2018

A significant limitation is that we do not have follow-up data on all UFJF applicants, but only on admitted students. Bertrand, Hanna and Mullainathan (2010) also mention that the lack of information on follow-up data on those who were not accepted is what harms the most the causal analysis of affirmative action policies. Taking account of this limitation, we do not intend to make any causal analysis.

3.2 Outcome Related to Academic Achievement

To analyze academic performance and catch-up, we use the relative CGPA.[3] To get to this variable, we needed to make some adjustments to our data. Following Frisancho and Krishna (2016) and Arcidiacono, Aucejo and Spenner (2012), we are concerned with several issues that can hurt our analysis of the student's academic performance.

We need to take into account that students in different majors enroll in various courses and face different professors, which can affect the distribution of grades. Also, even in the same major, classes can be mixed and also have a different distribution of grades. Moreover, students have some room to choose in which non-mandatory courses they want to enroll. This fact can also impact our analysis if, for example, students with lower grades tend to enroll in "easy" courses which could inflate their grades. Finally, another critical issue is to compare students with their classmates, and not with students from other majors.

To avoid the issues mentioned, we first divide the courses taken by the students into mandatory and non-mandatory courses. Note that each major has its own manda-

3 CGPA is the cumulative grade point average

tory courses, and all students in the major have to be approved in the mandatory courses to graduate.

We segregate students in selective and non-selective majors. This strategy is used in Frisancho and Krishna (2016) and we implement using the method applied by Velloso (2005). First we took the average entrance score for each major. We normalize those scores such as that they are distributed - N(0,1). Based on the normalized entrance score selective major are those majors in which the average entrance score is above half standard deviation from the mean, and non-selective majors are the majors that did not classify as selective. We also separate each student according to their classes, which means, students that were admitted in the same year and semester and the same major.

With the grades in mandatory courses and the students separated into their classes, we take the percentile distribution of grades in mandatory courses in each class. Then, we take the position of the students in the percentile distribution; we call this variable as the relative CGPA.

3.3 Labor Market Outcome

In this study, we link the academic outcomes from the previous section to the labor market outcomes. We use an administrative dataset called RAIS which is collected by the Brazilian Ministry of Labor (Ministério do Trabalho e Emprego, 2003-2013). Employers annually provide information on employees to the Ministry of Labor, which utilizes the data to determine certain labor benefits. RAIS covers formal workers in the private and public sector. RAIS does not cover informal workers, i.e., persons who are employed without a labor card or are self-employed but have not formally established a company.

Using the CPF, we were able to find 9,822 different students with a formal job between 2003 until 2013. Note that not all students can be found in RAIS for many reasons. Some professions such as doctors, dentists, lawyers, and physical educators are commonly self-employed, or the employer does not formally register their employees.

Summary statistics on all variables used, and what do they mean can be found in Appendices B and C.

3.4 Methodology

The empirical analysis consists of two parts, one using academic outcomes and the other using labor market outcomes. In the first part, we separate the data into two cohorts, before quotas (2003-2005) and after quotas (2006-2010). In the first cohort, we consider two groups of students, those who scored in the top two deciles in terms of the entrance exam (Top 20) and those who scored in the last two deciles (Bottom 20). The top 20 and bottom 20 are used for comparison purposes. For the second cohort, we have quota and non-quota students. The reason why we segregate the analysis into four groups is to make comparisons between groups before and after quotas were implemented.

The main goals of the first part are to analyze whether or not a quota and bottom 20 students catch-up with non-quota and top 20 students respectively, and analyze which variables are essential to explain academic performance.

To analyze catch-up, we use the same procedure adopted by Frisancho and Krishna (2016); we will look at the student's relative grade in the first year and compare to that of the last year. We want to verify if there is an improvement in relative grades of quota and bottom 20 students. As in Arcidiacono, Aucejo and Spenner (2012), we want to investigate which students improve their relative grade. We normalize the relative GPA of the 1st year of each student, and also the relative CGPA of the last year. Both relative GPA's will be distributed following N(0,1) - we refer to them as transformed relative GPA. We subtract the transformed relative grades of the last year relative CGPA to the 1st year relative GPA. We use this difference as the dependent variable regressed on a series of characteristic. The following equation is estimated:

$$\Delta Y_{it} = \beta_1 X_{it} + \beta_2 Z_{it} + \gamma_{mc} + \epsilon_{it} \tag{3.1}$$

Where ΔY_{it} is the transformed relative grade, X_{it} is an indicator for quota admission, Z_{it} is the normalized entrance exam score and γ_{mc} are a set of controls of major and individual characteristics.

To identify which variables are important to explain the student's grade, we regress the relative CGPA onto the same variables as equation 3.1. The following equation is estimated:

$$Y_{it} = \beta_1 X_{it} + \beta_2 Z_{it} + \gamma_{mc} + \epsilon_{it} \tag{3.2}$$

Where Y_{it} is the relative CGPA. Note that in both equations, our parameters of interest are those related to the quota dummy and that related to the entrance exam score. We want to verify if quota students and those students who scored in the lower part of the distribution of the entrance exam score, improve their grades, and how do they perform in college.

The second part of the analysis focuses on labor market outcomes for the same students we analyzed in the previous part. The main goals of this part are to identify how the indicator of quota students is correlated with wage and the probability of finding a job.

First, we deflated wage using the Price Index (IPCA) with 2013 as the base year. We took the logarithm of this variable to reduce the variance of the outcome. We use the log of wage as our dependent variable in the following equation:

$$Y_{it} = \beta_1 Q_i + \beta_2 S_i + \gamma_{im} + \epsilon_{it} \tag{3.3}$$

Where Y_{igt} is the log of wage for a person i in time t, Q_i is an indicator of admission through the quota system, S_i is the normalized entrance exam score, and γ_{im} are individual characteristics and major characteristics.

In equation 3.3 the sign of the parameter β_1 will give us differences in outcomes for quota and non-quota students. This difference can arise due to many factors. They can come from the lack of preparation in early stages of the life of quota students, or perhaps the employer makes some distinction to hire between quota and non-quota students and other factors that can contribute to this difference in outcomes. Despite all the unobserved effects that can explain the difference, this can be seen as a test of the mismatch hypothesis in the spirit of Rothstein and Yoon (2008). The authors suggests that any difference in outcomes for students with preferential admission compared to regular students, can be characterized as a mismatch between student's credentials and university's academic requirement. Nonetheless, the authors also highlight the fact that this is an overestimation of the effect, because there are many unobserved effects that are not being considered.

To analyze the probability of finding a job, we will use a logit to estimate the marginal effects and get the probability of finding a job. Note that we will use three different dependent variables. The first dependent variable is *Employed*. This variable indicates if we were able to find the student in RAIS. The second dependent variable is

Employed during college that indicates if the student worked while he was in college. The last dependent variable is *Employed after college* that indicates if we find the student in RAIS after his graduation in college. We do these three exercises to test if less favored students tend to get jobs earlier (during college).

4 Results on Academic Performance

4.1 Descriptive Statistics

We first present some summary statistics to characterize our database.[1] The statistics are stratified by the group of students that we prior defined.[2]

Table 2 shows the drop-out rate, the graduation rate, and ‘àctive” status shows the proportion of students that still enrolled at UFJF’s administrative system in 2018 (when data was collected). Quota students and bottom 20 students have similar drop-out rate, and for all groups, the drop-out rate is around 30%. Note that the column labeled “difference” is the difference between groups, and all differences are statistically significant.

Table 2 – Student Status, by group

	Non quota	Quota	Difference	Top 20	Bottom 20	Difference
Drop out Rate	0.334	0.295	0.043***	0.261	0.312	-0.050***
	(0.471)	(-0.455)		(0.439)	(0.463)	
Graduation Rate	0.616	0.631	-0.015***	0.677	0.698	0.010*
	(0.486)	(0.482)		(0.467)	(0.471)	
Active	0.031	0.058	-0.027***	0.002	0.006	-0.003***
	(0.172)	(0.233)		(0.049)	(0.075)	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Standard Errors in parentheses
Source: Data provided by UFJF, 2018

Table 3 presents the mean GPA, by year and by group. We can see that all groups improve their average GPA throughout the years. As expected, the GPA of the non-quota students and the Top 20 students are higher than their respective counterparts. The difference in both groups tend to increase over time, it could be an indication that both quota and bottom 20 students cannot catch-up. It is worth to mention the fact that the average time for a student to graduate in his major is five years, that is why the GPA falls dramatically after the 5th year.

[1] The correlation matrix of all variables used here is attached in the appendix

[2] For tables 2 and 3 consider the number of students such as: Non-quota: 8,185; Quota: 3,936; Top20: 2,730; Bottom 20: 1,043.

Table 3 – Mean GPA, by group and year

	Mean Grade			Mean Grade		
	Non Quota	Quota	Difference	Top 20	Bottom 20	Difference
1st Year	67.6	64.9	2.7***	72.3	68.6	3.8***
	(20.0)	(20.4)		(18.0)	(19.3)	
2nd Year	68.4	65.8	2.6***	73.4	69.1	4.2***
	(19.1)	(19.6)		(17.2)	(18.5)	
3rd Year	70.0	67.1	2.9***	74.5	69.9	4.6***
	(17.9)	(18.5)		(16.1)	(17.7)	
4th Year	71.1	68.1	3.0***	75.6	70.9	4.7***
	(17.0)	(17.5)		(14.9)	(16.5)	
5th Year	69.9	66.8	3.1***	74.5	69.5	5.1***
	(16.7)	(17.1)		(15.0)	(17.6)	
6th Year	64.4	61.3	3.0***	69.5	64.0	5.5***
	(17.5)	(17.5)		(16.2)	(19.0)	
7th Year	55.7	53.4	2.3***	60.7	55.1	5.7***
	(16.7)	(16.5)		(16.8)	(18.0)	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
GPA - Grade Point Average
Standard Errors in parentheses
Source: Data provided by UFJF, 2018

As we want to control for selective major and by area of study, table 4 show the average GPA by group for each of the three great areas of study (Humanities, Health, and Science) and also by major selectivity. Table 4 shows, that science majors have a lower average GPA than humanities and health majors. Selective majors also have a higher average GPA for all groups, which could mean that students in selective majors are more skillfully, or hard-workers than those in non-selective majors.

Table 4 – Mean GPA, by area

	Non quota	Quota	Difference	Top 20	Bottom 20	Difference
Humanities	70.3	67.5	2.8***	73.5	69.8	3.7***
	(17.6)	(18.0)		(16.8)	(17.3)	
Health	76.2	73.2	3.0***	79.2	77.7	1.5***
	(12.1)	(12.3)		(11.7)	(9.9)	
Science	62.6	58.9	3.7***	69.7	58.8	10.8***
	(19.4)	(20.3)		(17.4)	(19.1)	
Selective Majors	77.7	73.8	3.9***	80.2	74.8	5.4***
	(13.6)	(16.1)		(13.40)	(16.7)	
Non selective Majors	72.4	69.8	2.6***	77.7	73.5	4.2***
	(19.0)	(18.8)		(16.5)	(17.7)	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
GPA - Grade Point Average
Standard Errors in parentheses
Source: Data provided by UFJF, 2018

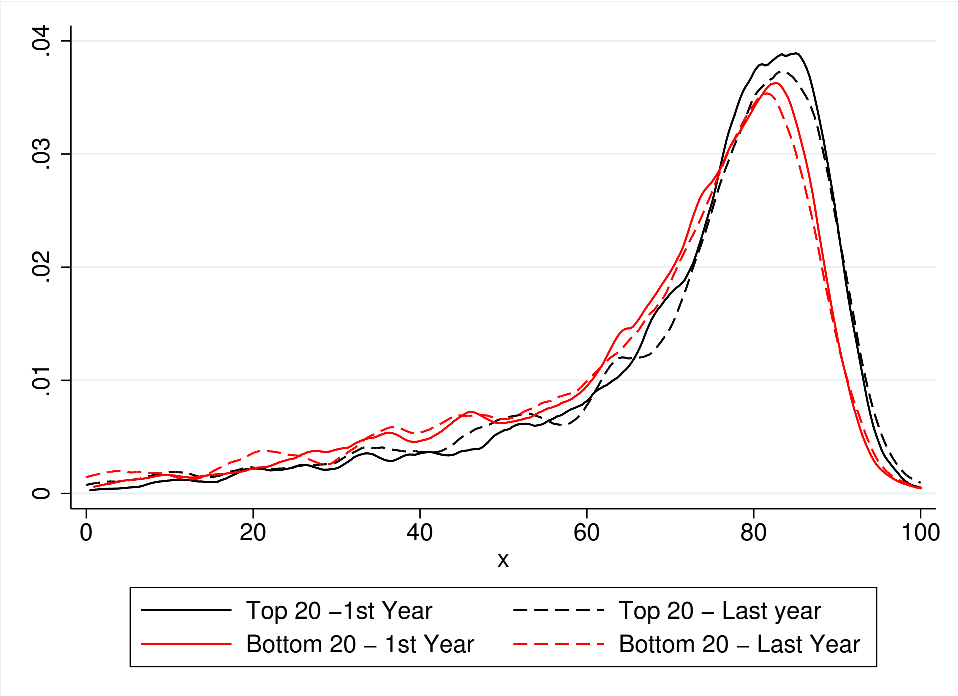
We have showed,that science majors have lower average grades than health and humanities, and that selective majors show higher grades than non-selective majors. Another feature we get is that quota, and bottom 20 students fall behind, in terms of grade, comparing to their respective counterparts, which could lead us to think that there is no catch-up occurring.

However, this is not enough to draw any conclusion about academic performance. In the following subsection, we start analyzing the data using some controls that will be important to avoid the problems cited in section 3. We first begin analyzing the distribution of raw grades. Then we separate the analysis for selective and non-selective majors. Therefore we control for grades in mandatory courses, and finally, we use relative grades as the dependent variable in our regression. Also, we first analyze the period before quotas, and afterward, we analyze the period after the quotas.

4.2 Before Quotas

Figure 1 shows the 1^{st} year GPA (Grade Point Average) and the last year CGPA (Cumulative Grade Point Average) of students from Top 20 and Bottom 20 groups. We can see that, as expected, the GPA of the top 20 students is higher when compared with the Bottom 20 students. The dashed lines for both groups seem to be below the solid lines, indicating that students tend to decrease their last year CGPA, which is normal since we can expect that in the first-year students take introductory courses.

Figure 1 – Performance Gap (Top 20 x Bottom 20)

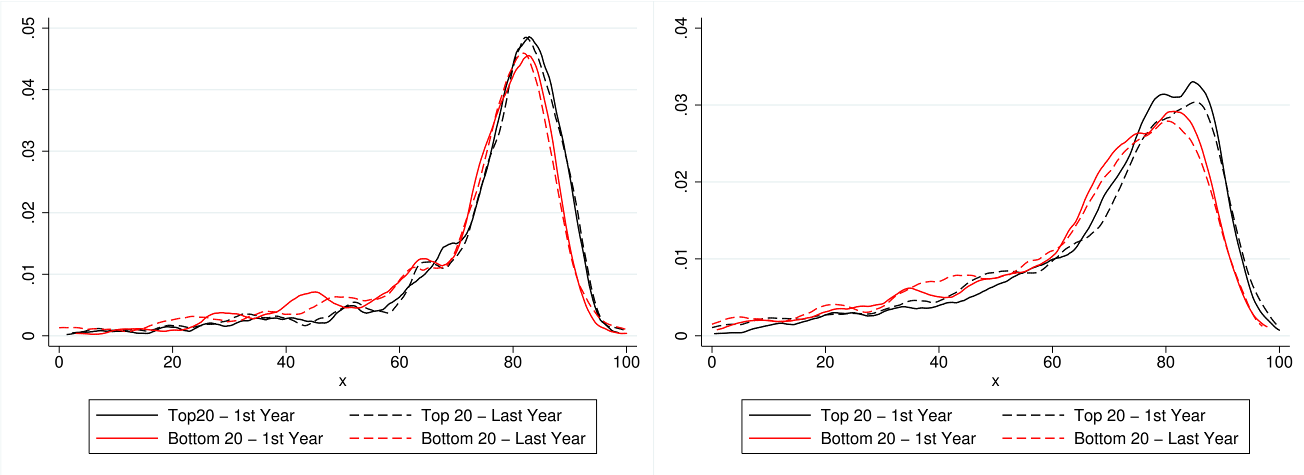


Source: Data provided by UFJF, 2018

Figure 2 shows the distribution of the GPA when we control for major selectivity. We can observe that for selective majors, the students of both groups are closer to each other, which can indicate that regardless of the entrance exam score, students are more homogeneous in terms of grades. Note that for non-selective majors, bottom 20 students seem to fall behind both in the first and last year, which can indicate more heterogeneity in terms of grades.

Figure 2 – Performance Gap (Top 20 x Bottom 20), by major

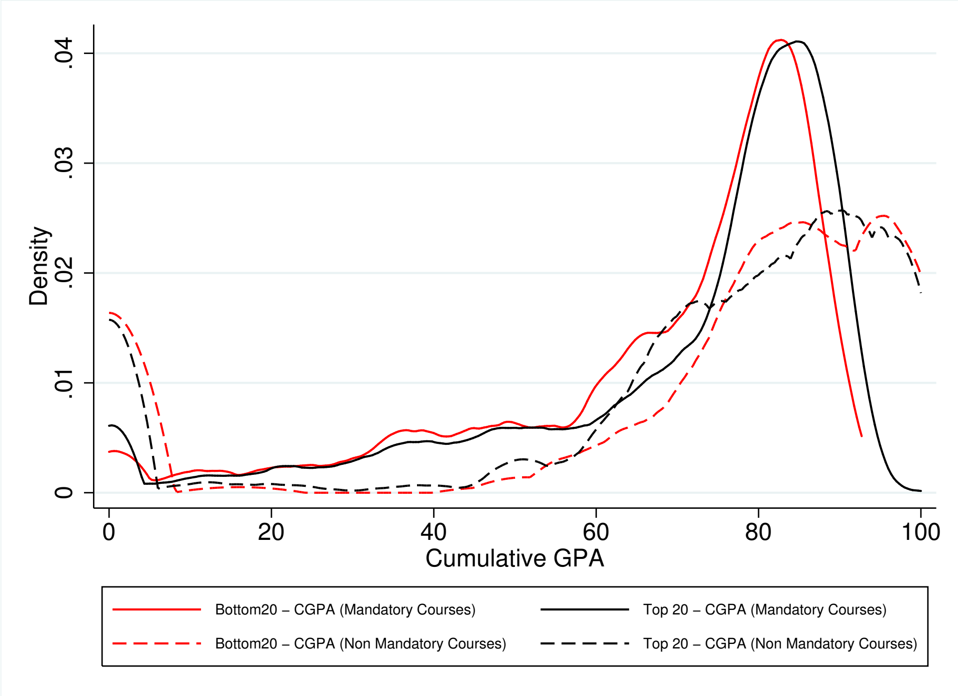
(a) Performance gap, selective (b) Performance gap, non selective



Source: Data provided by UFJF, 2018

Figure 3 shows the distribution of grades for mandatory and non-mandatory courses. We can see that non-mandatory courses give grades higher than 80 to a more significant share of students, which corroborates our assumption that non-mandatory courses can inflate grades.

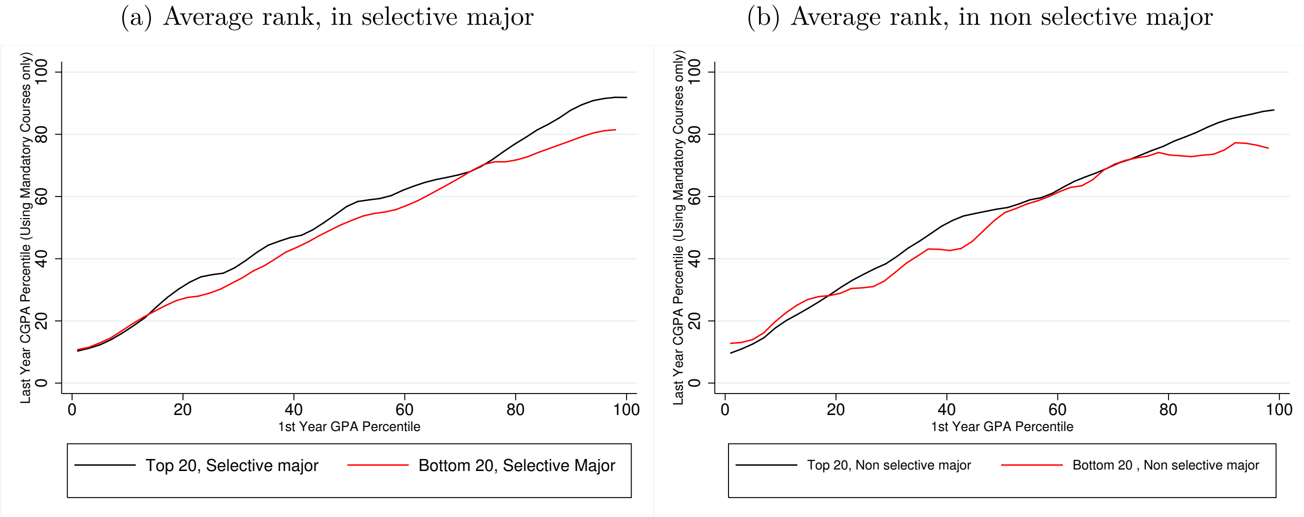
Figure 3 – Performance Gap (Top 20 x Bottom 20), by mandatory and non-mandatory courses



Source: Data provided by UFJF, 2018

Figure 4 show the average percentile, in terms of grades, on the last year against 1st year GPA Percentile. Note that in the first year all students are enrolled just in mandatory courses. The CGPA here is calculated considering only mandatory courses. Students (in both selective and non-selective majors) that are on the lowest percentiles in their classes tend, on average, to raise their relative grades. Students in the highest percentiles tend to decrease their relative grades. This finding is consistent with Frisancho and Krishna (2016) because students in top percentiles do not have a place to go but down, and students in the lowest percentile have no place to go but up.

Figure 4 – 1st year RGPA x Last year RCGPA, by major (using only mandatory courses)



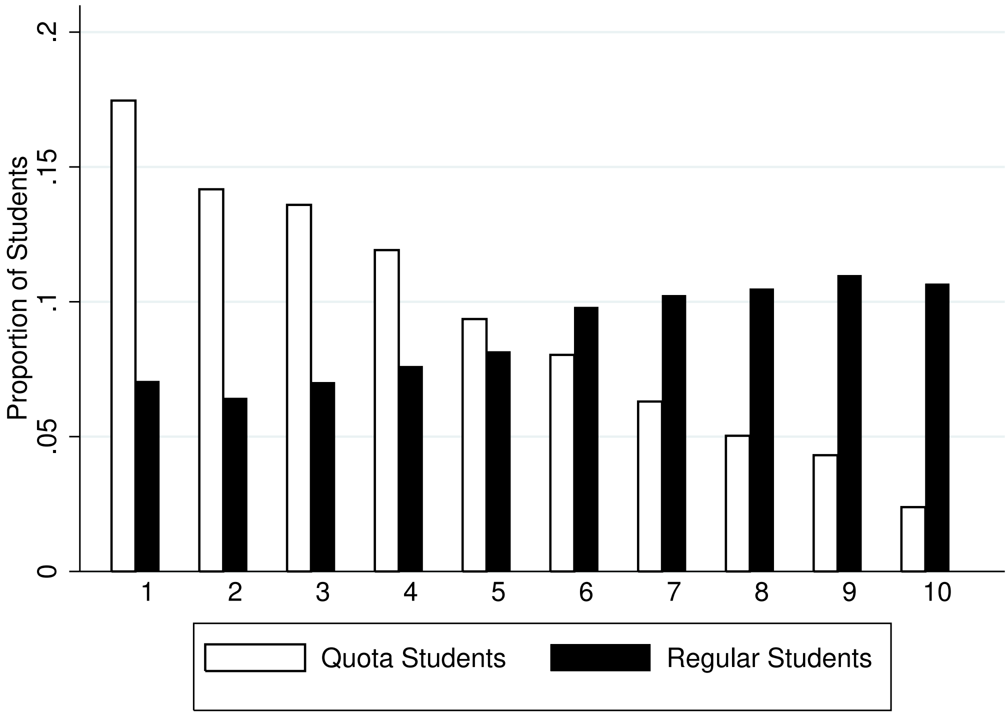
Source: Data provided by UFJF, 2018
Note: The average final rank was constructed using locally linear regression within each group, using a kernel smoother.

Figure 4 shows that bottom 20 students, especially in non-selective majors, do not catch-up with top 20 students. In selective majors, there are regions that the red line (Bottom 20) "touches" the black line (Top 20), but in general bottom 20 students do not catch-up.

4.3 Catch up Analysis After Quotas

Now we turn to the period after the quotas were implemented at UFJF. Figure 5 shows the proportion of students on the y-axis against the decile of the entrance exam score on the x-axis. Quota students perform poorly in the entrance exam, but how the quota and non-quota students fare at their majors is an interesting analysis object. The procedure to analyze the quota and non-quota students will be the same adopted in the previous section.

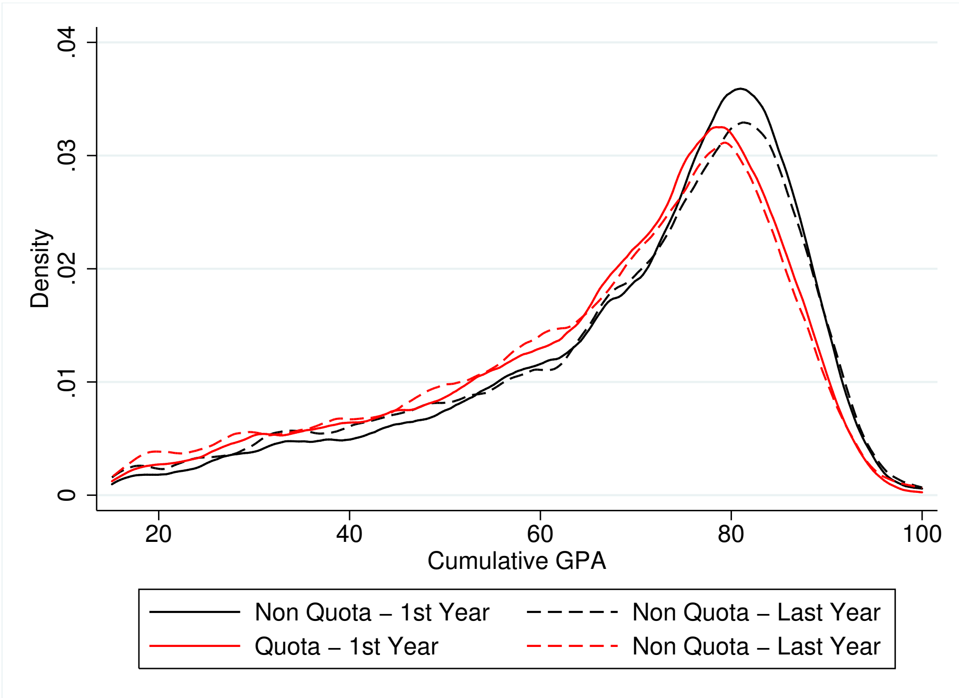
Figure 5 – Proportion of Students Entrance Exam Scores



Source: Data provided by UFJF, 2018

Figure 6 shows the gap between 1^{st} year GPA and the last year CGPA, for quota and non-quota students, irrespective of major selectivity. The interpretation is analogous to that of bottom 20 and top 20 students. Quota students enter with lower entrance exam score, and this seems to be important to predict academic performance because quota students in both 1st year and last year have lower grades than non-quota students.

Figure 6 – Performance Gap (Quota vs. Non Quota)

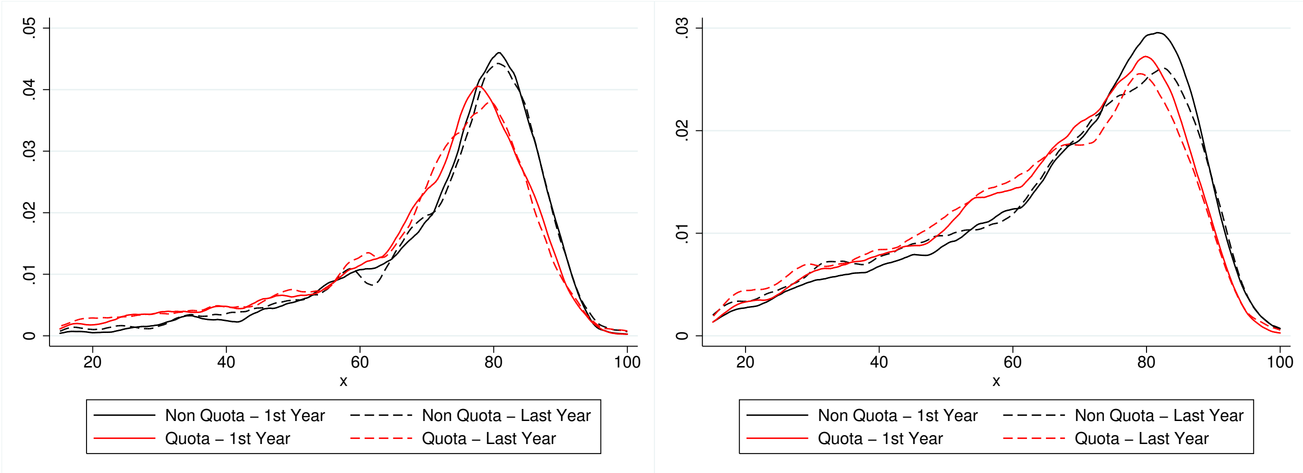


Source: Data provided by UFJF, 2018

When controlling for major selectivity, figure 7. We can see that in selective majors, quota students fall behind their non-quota counterparts. This is consistent to Frisancho and Krishna (2016) that perhaps selective majors demand more effort from quota student than it requires from non-quota students. In non-selective majors quota, students still have lower grades, but red lines seem to be closer to the black lines.

Figure 7 – Performance Gap (Quota X Non Quota), by major

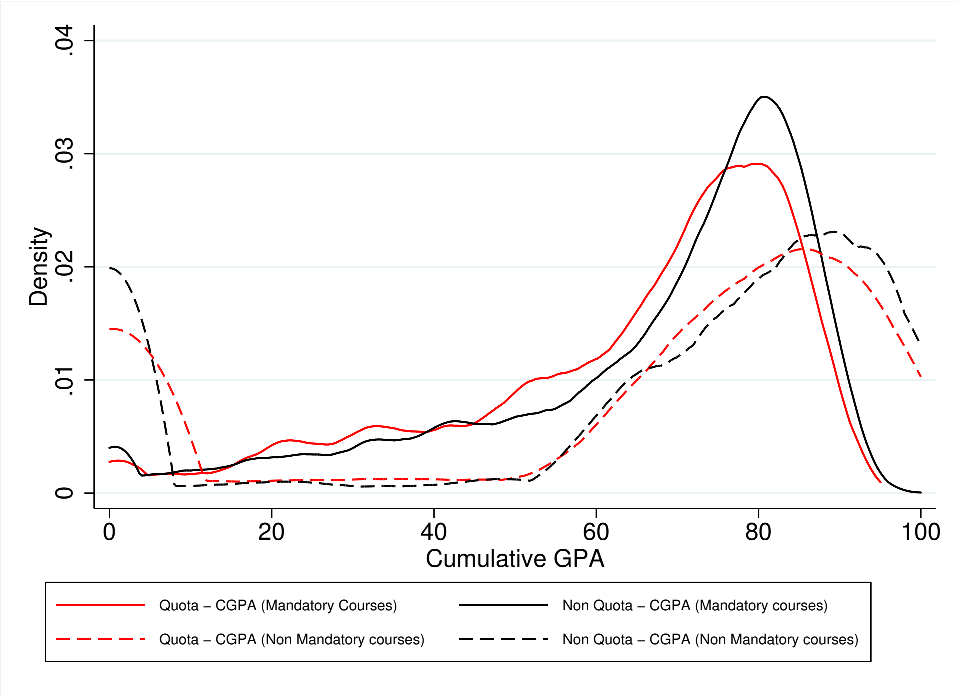
(a) Performance gap, selective (b) Performance gap, non selective



Source: Data provided by UFJF, 2018

However, we still need to account for the grades in mandatory courses. Figure 9 stratify the distribution of grades in mandatory and non-mandatory courses. Similar to what happened with top 20 and bottom 20 students, the share of students who score above 90 in non-mandatory courses is higher. Note also, that quota students seem to be much closer to non-quota students in non-mandatory courses than in mandatory courses.

Figure 8 – Performance Gap (Quota vs. Non Quota)



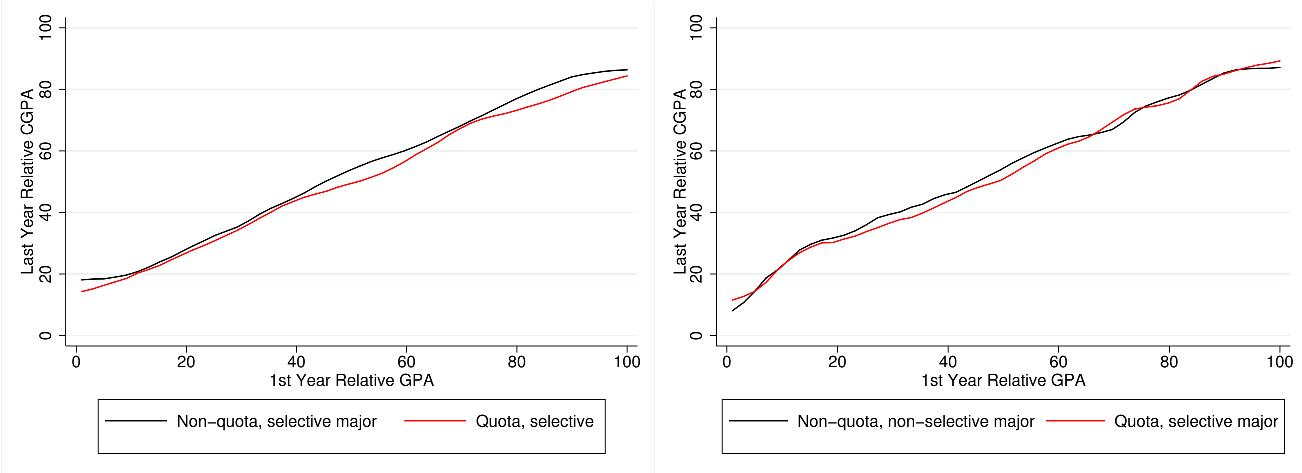
Source: Data provided by UFJF, 2018

To understand how quota students relatively fare with non-quota students, we control for relative grades using only mandatory courses. We add this control because mandatory courses are common to all students in the same class. Figure 9 shows that in selective majors, quota students seem to catch-up in lower percentiles. However, for higher percentiles, the red line (quota) tend to be below the black line (non-quota) indicating that quota students might not be catching-up. When we look at non-selective majors, we can see that quota students are catching up with non-quota students only in higher percentiles.

As an exercise of comparison figure 10, show information but now using both mandatory and non-mandatory courses to calculate the relative grade. We can see that quota and non-quota students seem to be very close.

Figure 9 – Average Relative CGPA, by major (using only mandatory courses)

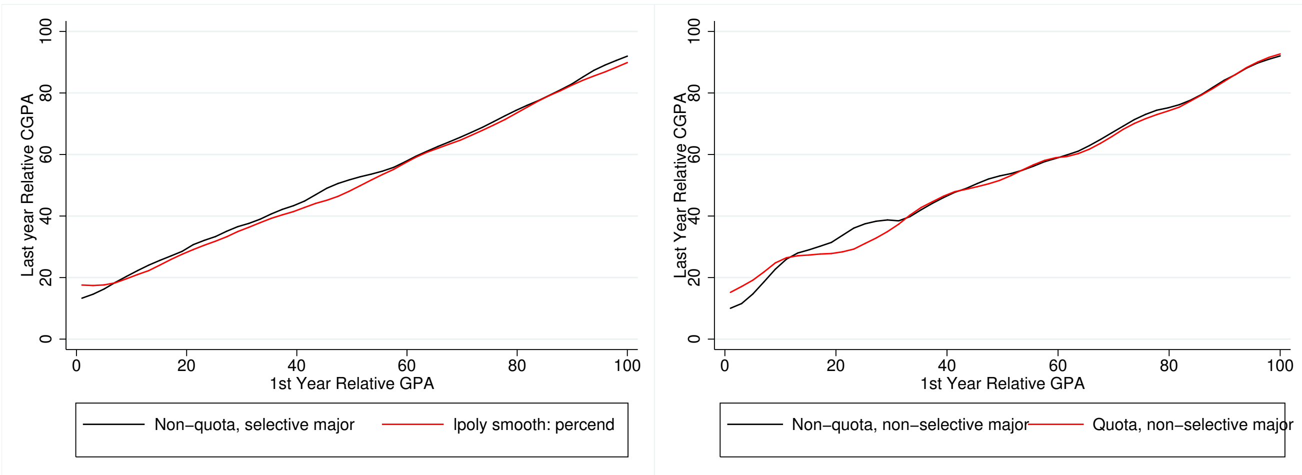
(a) Average Relative CGPA, in selective majors (b) Average Relative CGPA, in non selective majors



Source: Data provided by UFJF, 2018
Note: The average final rank was constructed using locally linear regression within each group, using a kernel smoother.

Figure 10 – Average Relative CGPA, by major (using mandatory and non-mandatory courses)

(a) Average Relative GPA, in selective majors (b) Average Relative CGPA, in non selective majors



Source: Data provided by UFJF, 2018
Note: The average Relative CGPA was constructed using locally linear regression within each group, using a kernel smoother.

Analyzing our transformed outcome of interest, namely relative GPA, we can find little evidence that quota students are catching-up with non-quota students in selective majors. But for non-selective majors, quota students in higher percentiles tend to catch-up.

Now we analyze how the observable variables impact the relative CGPA in the last year of college. The main idea is to understand which variables are significant to explain academic performance measured by relative CGPA. Note that we are most interested in how preferential admission and entrance exam score impacts our outcome.

Table 5 shows the OLS using last year relative CGPA on the observable variables. We first regress using 1st year relative GPA. Then we start adding the other variables. Columns 3 and 4 account for information regarding major and individual characteristics, respectively. Our parameters of interest are those related to the quota system and related to entrance exam score.

On average, quota students are, in lower positions in their classes, when comparing to non-quota students. This was expected, and the results corroborate what both descriptive statistics and graphical intuition showed in the previous sections. The entrance exam score also has the expected effect on relative CGPA since students with higher scores tend to thrive in their classes.

Table 5 – OLS Using Last Year Relative CGPA (Mandatory courses only)

	(1)	(2)	(3)	(4)
1st Year Relative GPA	0.806***	0.807***	0.808***	0.801***
	(699.48)	(633.11)	(637.00)	(613.34)
Quotas		-1.653***	-1.533***	-1.346***
		(-19.02)	(-17.73)	(-15.05)
Entrance Exam Score		-0.424***	0.0654	0.142**
		(-11.17)	(1.51)	(3.28)
Selective Major			-1.634***	-1.537***
			(-18.96)	(-17.90)
Health			-0.301**	-0.497***
			(-3.13)	(-5.19)
Science			3.630***	4.439***
			(42.03)	(49.92)
Female				2.676***
				(34.01)
Nonwhite				-0.845***
				(-8.39)
Age				-0.00652*
				(-2.24)

Source: Data provided by UFJF, 2018

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In the same spirit of Arcidiacono, Aucejo and Spenner (2012) we want to know which students improve their relative position, which means that we want to look at the gains in relative grade. Here we transform the relative GPA of the first year and the relative CGPA of the last year, such as they are distributed N(0,1). Then, we subtract the transformed relative grade of the last year from the transformed relative grade of the first year. Finally, we regress the difference on a series of characteristics.

Table 6 show the OLS when we regress gains in relative grades. In column 2 and 3, we add controls relative to major choice and individual characteristics, respectively.

White females and non-quota students seem to improve their position. On the other hand, black males and quota students do not improve their position. If students do not improve their position in class we cannot say that they are catching-up in terms of relative grades. However we should mention that quota students, although they are not catching-up, they end-up really close to non-quota students.

Table 6 – Gains in relative GPA

	(1)	(2)	(3)
Quotas	-0.0341***	-0.0302***	-0.0248***
	(-10.74)	(-9.58)	(-7.57)
Entrance Exam Score	-0.0355***	-0.0177***	-0.0175***
	(-25.90)	(-11.26)	(-11.09)
Selective Major		-0.0619***	-0.0606***
		(-19.59)	(-19.18)
Health		-0.00302	-0.00494
		(-0.86)	(-1.40)
Science		0.140***	0.150***
		(44.56)	(46.38)
Female			0.0310***
			(11.08)
Nonwhite			-0.0262***
			(-7.14)
Age			0.000375***
			(5.53)

Source: Data provided by UFJF, 2018
t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The results we have found so far, are somewhat consistent with the literature. Quota students on average receive lower grades than non-quota students and especially in selective majors quota students tend to have more difficulty to catch-up, which is what Frisancho and Krishna (2016) find in their paper for India. Following the strategy of Arcidiacono, Aucejo and Spenner (2012), we find that controlling for relative grades and minimizing the problem of course selection, quota students only catch-up with non-quota students in the highest percentiles of non-selective majors. When we consider non-mandatory courses to calculate GPA, catch-up is more likely to occur. This result is also consistent with what Arcidiacono, Aucejo and Spenner (2012) find in their paper.

5 Labor Market Performance

In this section, we follow the students from the previous section, in the labor market. We want to investigate how the students, given their academic credentials, perform in the labor market.

The main goal of this section is to investigate if there are any differences in wages earned by quota and non-quota students. First, we present descriptive statistics, and then we make use of OLS to measure differences in wages.

5.1 Descriptive Statistics

With the analysis of the last chapter, we can link academic performance with labor market outcomes. We expect that academic outcomes are, somewhat, important to explain labor market outcomes, since the literature points out that there is a positive correlation between college and wages (HOEKSTRA, 2009; BREWER; EIDE; EHRENBERG, 1999; BLACK; SMITH, 2006).

Table 7 shows how many students we have in the original database by admission year. Table 8 show the percentage of these students were found in RAIS. WE can see that the percentage of students with a formal job increase after 5 or 6 years after the admission year, which is natural since we expect that after graduation the students tend to find jobs more easily. Also, note that the number of quota students that work while in the college is higher, compared with the number of non-quota students.

Table 7 – Total of students, by quotas and admission year

Admission Year	2003		2004		2005		2006		2007		2008		2009		2010	
Quota	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
N	2,249	-	2,207	-	2,149	-	1,801	394	1,659	502	1,508	715	1,537	1,041	1,680	1,284

Source: Data provided by UFJF, 2018 and RAIS, 2013

Table 8 – Students in RAIS

RAIS	2003		2004		2005		2006		2007		2008		2009		2010		2011		2012		2013	
Quota	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Admission Year																						
2003	10%	-	12%	-	13%	-	14%	-	22%	-	35%	-	46%	-	51%	-	54%	-	55%	-	57%	-
2004	8%	-	10%	-	12%	-	13%	-	15%	-	26%	-	37%	-	47%	-	53%	-	55%	-	57%	-
2005	7%	-	8%	-	12%	-	14%	-	15%	-	19%	-	27%	-	38%	-	50%	-	54%	-	57%	-
2006	4%	8%	6%	12%	7%	15%	10%	18%	11%	17%	14%	20%	16%	20%	24%	27%	38%	39%	47%	53%	52%	58%
2007	3%	10%	4%	12%	6%	12%	7%	15%	10%	17%	13%	20%	13%	21%	16%	22%	23%	30%	33%	38%	43%	50%
2008	3%	4%	4%	8%	4%	10%	6%	13%	8%	16%	12%	19%	13%	17%	12%	18%	14%	20%	20%	26%	27%	38%
2009	2%	4%	3%	6%	4%	8%	5%	11%	7%	15%	8%	18%	9%	18%	10%	14%	11%	15%	13%	16%	15%	19%
2010	2%	4%	3%	5%	3%	6%	4%	9%	5%	10%	7%	13%	8%	14%	8%	14%	8%	12%	9%	13%	10%	13%

Source: Data provided by UFJF, 2018 and RAIS, 2013

Table 9 shows information on labor market outcomes by group. We can see that percentage of quota students that work during college is much higher than the percentage of non-quota students that work during college. Top 20 and non-quota students have around one-fourth of the students working during college. Also, note that Top 20 and non-quota students have a higher rate of students that find a formal job after graduation.

Table 9 – Descriptive Statistics, by group

Variable (Mean)	Top 20	Bottom 20	Quotas	Non-Quotas
Age in Admisison	20.4	22.0	22.5	21.0
Standarized Vestibular Score	0.0178	-0.4864	-0.3119	-0.1036
Health	21%	18%	18%	21%
Humanities	57%	58%	57%	58%
Science	20%	23%	24%	20%
Wage (log)	6.37	6.09	5.92	6.34
Employed after graduation	46%	34%	28%	44%
Employed during college	24%	38%	53%	25%

Source: Data provided by UFJF, 2018 and RAIS, 2013

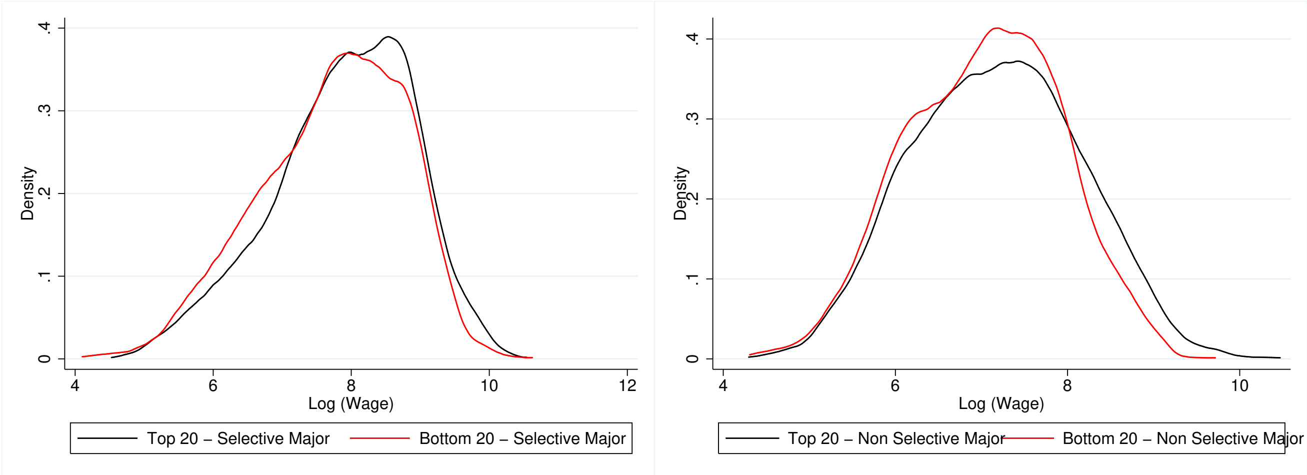
Figure 11 shows how Top 20 students and Bottom 20 students fare, in terms of wage, by selective and non-selective major. Our initial assumption, following Velloso (2005), is that selective majors are well paid, and the students take this fact into account when deciding on which major to apply to, so it becomes more competitive. Figure 11 illustrates that Top 20 students in selective majors receive a higher wage than other Top 20 students in non-selective majors.

We can also see the same information of figure 11, but now considering quota and non-quota students. Figure 12 shows that non-quota students tend to get a higher salary than quota students in both selective and non-selective majors. Again, students enrolled in selective majors tend to get a higher salary than students in non-selective majors.

Figure 11 – Wage gap (Top 20 X Bottom 20), by major

(a) Wage gap , selective

(b) Wage gap, non selective

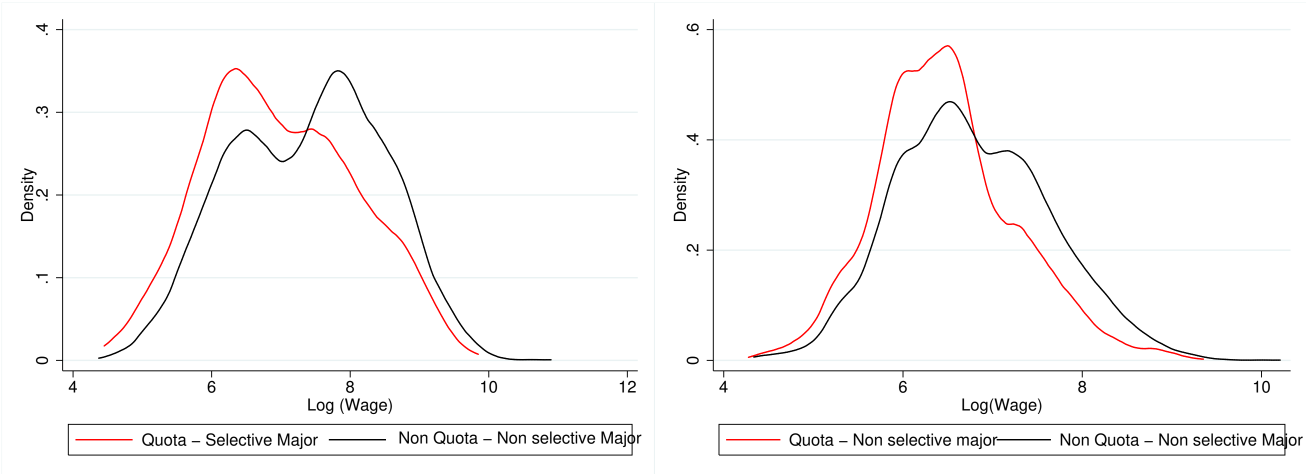


Source: Data provided by UFJF, 2018 and Ministry of Labor

Figure 12 – Wage Gap (Quota X Non Quota), by major

(a) Wage Gap, selective

(b) Wage Gap, non selective



Source: Data provided by UFJF, 2018 and Ministry of Labor

5.2 Labor Market Outcomes

We have seen that Top 20 students tend to get higher monthly wage than their Bottom 20 counterparts, and non-quota students tend to get higher salaries than their quota students counterparts. But how the wages are correlated with the observable variables?

Table 10 shows the coefficients of an OLS when we regress the log of wages on a series of characteristics using only students admitted before 2006. Note that, since quotas were implemented in 2006, and the eligibility criteria were to come from public school, we add a variable *public school*. If the student came from public schools, he would be eligible to apply for quotas if it were available at the time.

Students from public school, in an environment without quotas, receive around 7% less than a student from a private school. And nonwhite students from public schools, which would be other criteria for quotas after 2005, receive around 20% less than white students from private schools. Also, note that students in selective majors have substantial gains than those students in non-selective majors.

Table 11 shows the same information as table 10 but using students admitted after 2005. Quotas have a significant negative correlation with wages. On average, quota students receive around 18% less than non-quota students. And note that nonwhite quota students, on average, receive around 30% less than white, non-quota students.

Thus, table 10 and 11 show that students from public school and/or quota students receive a lower salary. Despite all the unobserved effects that can explain the difference, this can be seen as a test of the mismatch hypothesis (ROTHSTEIN; YOON, 2008).

Table 10 – OLS using log (wage) as dependent variable for students admitted before 2006 (2003 - 2005)

	(1)	(2)	(3)	(4)	(5)
Public School	-0.195***	-0.0805***	-0.0729***	-0.0788***	-0.0722***
	(-14.16)	(-5.32)	(-4.78)	(-5.18)	(-4.75)
Vestibular Score		0.0568***	0.0469***	0.0426***	0.0436***
		(7.92)	(6.52)	(5.92)	(6.06)
Selective Major		0.605***	0.565***	0.560***	0.558***
		(38.02)	(33.68)	(33.48)	(33.37)
Health			0.107***	0.165***	0.164***
			(4.88)	(7.32)	(7.25)
Humanities			-0.110***	-0.0647**	-0.0645**
			(-5.37)	(-3.10)	(-3.09)
Female				-0.173***	-0.176***
				(-11.47)	(-11.64)
Nonwhite					-0.135***
					(-5.67)
N	21842	17660	17660	17660	17660

Source: Data provided by UFJF, 2018 and RAIS, 2013

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Data provided by UFJF, 2018 and RAIS, 2013

Table 11 – OLS using log (wage) as dependent variable for students admitted after 2005 (2006 - 2010)

	(1)	(2)	(3)	(4)	(5)
Quotas	-0.238***	-0.175***	-0.196***	-0.191***	-0.185***
	(-17.22)	(-11.84)	(-13.15)	(-12.83)	(-12.32)
Vestibular Score		0.337***	0.287***	0.303***	0.285***
		(35.10)	(25.92)	(26.70)	(25.30)
Selective Major			0.170***	0.112***	0.112***
			(8.69)	(5.55)	(5.63)
Health				-0.163***	-0.0788**
				(-6.16)	(-2.86)
Humanities				-0.180***	-0.122***
				(-9.55)	(-6.28)
Female					-0.184***
					(-12.28)
Nonwhite					-0.118***
					(-7.28)
N	19028	15610	15610	15610	15610

Source: Data provided by UFJF, 2018 and RAIS, 2013

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Data provided by UFJF, 2018 and RAIS, 2013

We now turn to the probability of finding a job. Table 12 and 13 show the logit marginal effects using "Employed" as dependent variable, " Employed during college" as dependent variable and "Employed after college" as dependent variable, respectively.

Employed means that the student was found in RAIS at some point. *Employed during college* means that the student worked during college. *Employed after college* means that the student was working after college.

Table 12 shows the logit marginal effects using the three dependent variables above, for students admitted after 2005. The parameter of interest is the one related to the quota variable. When using *Employed* as dependent variable quota students have 8.4% more chance to be found in RAIS than non-quota students. Now, using *Employed during college* as the dependent variable, we find that quota students have a higher probability, around 14.8% to work during their college time. On the other hand, when we use *Employed after college* as the dependent variable, we find that quota students have a lower probability, around 13%, to find a job.

Table 13 shows the same information we just described above, but using students admitted before 2005. Instead of quotas, we use the variable *Public School*, which is a criterion for quota eligibility. In table 13 we use *Employed* as the dependent variable, we can see that students from public school have a higher probability to be found in RAIS. When using *Employed during college* as the dependent variable, we observe a higher probability, around 4%, that students from public schools to work during college. When we use *Employed after college* as the dependent variable, we see that the probability of finding a student from public school in RAIS, after college, decrease by 5.3%.

It is reasonable to think of quota and public school students as less favored, relative to non-quota and private school students. In both tables we can see that the less favored students have a higher probability to work during college, this fact can have many explanations, but one of these explanations can be attributed to the fact that students with lower income are more eager to increase their family income by working during college. On the other hand, non-quota students and private school students can wait until they graduate.

These results are somewhat converging to what Arabage and Souza (2016) find in their paper. We also find that quota students receive a lower salary. But, in this work, we find that the probability of be employed is also affected by quotas.

Table 12 – Logit using Students admitted after quotas (2006-2010)

	Marginal Effects using Employed as dependent variable					Marginal Effects using Employed During college					Marginal Effects using Employed after college				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Quotas	0.041***	0.057***	0.075***	0.085***	0.084***	0.143***	0.144***	0.155***	0.164***	0.148***	-0.115***	-0.114***	-0.129***	-0.142***	-0.130***
	(8.21)	(10.22)	(13.52)	(15.40)	(14.79)	(23.76)	(21.79)	(23.19)	(24.20)	(21.52)	(-19.83)	(-17.73)	(-19.85)	(-21.58)	(-19.33)
Vestibular Score		-0.024***	0.024***	0.046***	0.044***		-0.085***	-0.058***	-0.046***	-0.043***		0.097***	0.058***	0.036***	0.037***
		(-7.39)	(6.29)	(11.42)	(11.04)		(-22.09)	(-13.10)	(-10.07)	(-9.30)		(25.95)	(13.43)	(8.05)	(8.23)
Selective major			-0.171***	-0.165***	-0.166***			-0.093***	-0.071***	-0.075***			0.131***	0.135***	0.138***
			(-24.28)	(-21.80)	(-21.85)			(-11.86)	(-8.74)	(-9.19)			(16.95)	(16.61)	(16.98)
Health				-0.158***	-0.144***				-0.090***	-0.073***				0.188***	0.150***
				(-17.27)	(-15.17)				(-8.47)	(-6.60)				(18.10)	(13.88)
Humanities				0.035***	0.044***				0.088***	0.102***				-0.003	-0.030***
				(5.03)	(5.96)				(10.88)	(12.10)				(-0.41)	(-3.68)
Female					-0.025***					-0.042***					0.086***
					(-4.39)					(-6.17)					(12.88)
Nonwhite					-0.003					0.078***					-0.038***
					(-0.52)					(10.19)					(-5.06)
N	29601	24677	24677	24677	24677	29601	24677	24677	24677	24677	29601	24677	24677	24677	24677

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Data provided by UFJF, 2018 and RAIS, 2013

Table 13 – Logit using Students admitted before quotas (2003-2005)

	Marginal Effects using Employed as dependent variable					Marginal Effects using Employed During college					Marginal Effects using Employed after college				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
Public School	0.036***	0.030***	0.028***	0.026***	0.026***	0.076***	0.071***	0.065***	0.043***	0.039***	-0.081***	-0.071***	-0.068***	-0.057***	-0.053***
	(14.39)	(9.97)	(9.88)	(9.40)	(9.07)	(17.08)	(15.16)	(13.80)	(10.38)	(9.61)	(-14.41)	(-11.13)	(-10.52)	(-8.64)	(-8.12)
Vestibular Score		-0.009***	-0.006***	-0.006***	-0.006***		0.009***	0.014***	0.020***	0.018***		0.011***	0.008**	-0.002	-0.001
		(-7.25)	(-5.07)	(-4.91)	(-4.95)		(5.28)	(7.65)	(11.31)	(10.22)		(3.99)	(2.77)	(-0.73)	(-0.49)
Selective Major			-0.022***	-0.024***	-0.0.24***			-0.043***	-0.041***	-0.043***			0.029***	0.003	0.02
			(-7.16)	(-7.29)	(-7.21)			(-9.66)	(-10.24)	(-10.76)			(4.41)	(0.47)	(0.40)
Health				-0.009	-0.010				-0.149***	-0.136***				0.151***	0.147***
				(-2.12)	(-2.23)				(-39.62)	(-35.47)				(16.57)	(15.71)
Humanities				-0.010**	-0.011**				0.088***	-0.022***				-0.050***	-0.053***
				(-2.73)	(-2.89)				(-7.87)	(-5.01)				(-5.95)	(-6.18)
Female					0.035					-0.053***					0.086
					(1.23)					(-13.56)					(1.36)
Nonwhite					0.012**					0.035***					-0.065***
					(2.83)					(5.32)					(-6.38)
N	29601	24677	24677	24677	24677	29601	24677	24677	24677	24677	29601	24677	24677	24677	24677

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Data provided by UFJF, 2018 and RAIS, 2013

Conclusion

To start the conclusion of this study, it is necessary to highlight that we have data limitations, and that this is a preliminary study. We do not try to make any causal inference. Another important limitation is that we are analyzing the UFJF's case, we cannot generalize the results to any other case. Nevertheless, we are not making a policy analysis, we cannot say anything about the quota policy itself, we are just analyzing the students over time.

In this study, we investigate the academic performance and labor market performance of students at UFJF for the period of 2003 to 2010. First, we compare students before quotas (2003-2005), we separated students into the top 2 deciles in terms of entrance exam score and students in the bottom two deciles. Separating into selective and non-selective majors, controlling for mandatory courses and using relative grades, we find that students with low entrance exam score do not catch-up, in terms of grades, with their counterparts (top 20 students) in both selective and non-selective majors.

Second, we compare quota students and non-quota students, using the same strategy from above. We find that quota students when we do not control for mandatory courses are very close to their non-quota counterparts, indicating that catch-up is occurring. However, when we control for mandatory courses, quota students seem to fall behind their counterparts, but in non-selective majors quota students in high percentiles seem to catch-up. This fact is consistent with Arcidiacono et al. (2011). Controlling for relative grades and course selection, we find that quota students do not catch-up with their counterparts. However, quota students are close to their non-quota peers, meaning that although they are not catching-up they are not falling behind either.

Moreover, we analyze the same students in the labor market. We find that, in general, students in selective majors receive higher wages than those students in non-selective majors, which is consistent with Velloso (2005).

We also found that quota students, on average, receive a lower salary than non-quota students, around 18% less. The goal of this work was to analyze the probability of finding the students in the labor market. We used three different variables as dependent variable, *Employed*, *Employed during college* and *Employed after college*. We find that quota students have a higher probability of finding a job during their college time, but the probability of finding a job after college is lower for quota students.

Studying affirmative action in Brazil is useful for several reasons. Much of the literature about AA is done in a US context. Few are the works about AA outside the US. Quota policy implemented by public universities, thus public resources are at stake, and they must be allocated efficiently. Therefore studies about AA contributes to this matter.

This work and its findings contribute to the literature for affirmative action in higher education and to the public debate by providing scientific evidence. This is the first analysis of catch-up and labor market performance using this database. Another significant contribution is that we have created a database that can be used in future researches.

It is also important to highlight the limitations of the work. This thesis is an applied study to the case of the UFJF; it does not mean that we can generalize the result to another context. The lack of follow-up data on those who did not enroll at UFJF does not allow us to make any causal analysis.

Future research may be able to extend the analysis. To analyze the causal relationship between observable and outcomes. To follow the students in the labor market and see if quota students tend to get different occupations than their same major counterparts.

Understand more about the implications, the mechanisms that quota policies operates, and also bring scientific evidence to the debate is something that we still have to pursue. Quota policy still a matter of public debate, and will still affect the lives of millions of people worldwide.

Bibliography

ALON, S.; MALAMUD, O. The impact of israel's class-based affirmative action policy on admission and academic outcomes. *Economics of Education Review*, Elsevier, v. 40, p. 123–139, 2014. Citado na página 17.

ALON, S.; TIENDA, M. Assessing the "mismatch" hypothesis: Differences in college graduation rates by institutional selectivity. *Sociology of education*, SAGE Publications Sage CA: Los Angeles, CA, v. 78, n. 4, p. 294–315, 2005. Citado 2 vezes nas páginas 10 and 18.

ALTONJI, J. G.; ARCIDIACONO, P.; MAUREL, A. The analysis of field choice in college and graduate school: Determinants and wage effects. In: *Handbook of the Economics of Education*. [S.l.]: Elsevier, 2016. v. 5, p. 305–396. Citado na página 9.

ARABAGE, A. C.; SOUZA, A. P. Quotas in public universities and labor outcomes: Evidence for rio de janeiro1. 2016. Citado 2 vezes nas páginas 19 and 50.

ARCIDIACONO, P. et al. Does affirmative action lead to mismatch? a new test and evidence. *Quantitative Economics*, Wiley Online Library, v. 2, n. 3, p. 303–333, 2011. Citado na página 53.

ARCIDIACONO, P.; AUCEJO, E. M.; SPENNER, K. What happens after enrollment? an analysis of the time path of racial differences in gpa and major choice. *IZA Journal of Labor Economics*, SpringerOpen, v. 1, n. 1, p. 5, 2012. Citado 10 vezes nas páginas 9, 10, 17, 19, 20, 22, 24, 40, 41, and 42.

BERALDO, A. F.; MAGRONE, E. Política de cotas na universidade federal de juiz de fora: avaliação 2006-2011. *O impacto das cotas nas universidades brasileiras (2004-2012)*, CEAO Salvador, p. 105–134, 2013. Citado na página 19.

BERTRAND, M.; HANNA, R.; MULLAINATHAN, S. Affirmative action in education: Evidence from engineering college admissions in india. *Journal of Public Economics*, Elsevier, v. 94, n. 1-2, p. 16–29, 2010. Citado 4 vezes nas páginas 9, 19, 20, and 22.

BLACK, D. A.; SMITH, J. A. Estimating the returns to college quality with multiple proxies for quality. *Journal of labor Economics*, The University of Chicago Press, v. 24, n. 3, p. 701–728, 2006. Citado na página 43.

BOWEN, W. G.; BOK, D. *The Shape of the River. Long-Term Consequences of Considering Race in College and University Admissions.* [S.l.]: ERIC, 1998. Citado na página 18.

BREWER, D. J.; EIDE, E. R.; EHRENBERG, R. G. Does it pay to attend an elite private college? cross-cohort evidence on the effects of college type on earnings. *Journal of Human resources*, JSTOR, p. 104–123, 1999. Citado na página 43.

CARDOSO, C. B. Efeitos da política de cotas na universidade de brasília: uma análise do rendimento e da evasão. 2008. Citado na página 14.

COLEMAN, J. S. Equality of educational opportunity. *Integrated Education*, Taylor & Francis, v. 6, n. 5, p. 19–28, 1968. Citado na página 10.

ESTEVAN, F.; GALL, T.; MORIN, L.-P. Redistribution without distortion: Evidence from an affirmative action program at a large brazilian university. *The Economic Journal*, Wiley Online Library, 2016. Citado na página 19.

FERMAN, B.; ASSUNÇÃO, J. Affirmative action in university admissions and high school students' proficiency. *XXVii Encontro Brasileiro de Econometria. Anais*, 2005. Citado 2 vezes nas páginas 9 and 19.

FRANCIS, A.; TANNURI-PIANTO, M. Black movement: Using discontinuities in admissions to study the effects of college quality and affirmative action. 2016. Citado 3 vezes nas páginas 9, 19, and 20.

FRANCIS, A. M.; TANNURI-PIANTO, M. Using brazil's racial continuum to examine the short-term effects of affirmative action in higher education. *Journal of Human Resources*, University of Wisconsin Press, v. 47, n. 3, p. 754–784, 2012. Citado na página 19.

FRISANCHO, V.; KRISHNA, K. Affirmative action in higher education in india: targeting, catch up, and mismatch. *Higher Education*, Springer, v. 71, n. 5, p. 611–649, 2016. Citado 10 vezes nas páginas 17, 18, 19, 20, 22, 23, 24, 32, 35, and 41.

GAGO, T. F. T. *AVALIAÇÃO DA POLÍTICA DE COTA DA UNIVERSIDADE FEDERAL DE JUIZ DE FORA (2006, 2012 e 2016)*. Dissertação (Mestrado) — Universidade Federal de Juiz de Fora, 2016. Citado na página 19.

GOLDSTEIN, M.; SMITH, R. S. The estimated impact of the antidiscrimination program aimed at federal contractors. *ILr review*, SAGE Publications Sage CA: Los Angeles, CA, v. 29, n. 4, p. 523–543, 1976. Citado na página 13.

HECKMAN, J. J.; WOLPIN, K. I. Does the contract compliance program work? an analysis of chicago data. *ILr review*, SAGE Publications Sage CA: Los Angeles, CA, v. 29, n. 4, p. 544–564, 1976. Citado na página 13.

HOEKSTRA, M. The effect of attending the flagship state university on earnings: A discontinuity-based approach. *The Review of Economics and Statistics*, MIT Press, v. 91, n. 4, p. 717–724, 2009. Citado na página 43.

HOLZER, H.; NEUMARK, D. Assessing affirmative action. *Journal of Economic literature*, v. 38, n. 3, p. 483–568, 2000. Citado na página 17.

JR, R. G. F.; LEVITT, S. D. The black-white test score gap through third grade. *American Law and Economics Review*, Oxford University Press, v. 8, n. 2, p. 249–281, 2006. Citado na página 10.

JR, R. G. F.; LOURY, G. C.; YURET, T. An economic analysis of color-blind affirmative action. *The Journal of Law, Economics, & Organization*, Oxford University Press, v. 24, n. 2, p. 319–355, 2007. Citado 2 vezes nas páginas 17 and 18.

KOCHAR, A. Affirmative action through quotas: The effect on learning in india. *Stanford Center for International Development, Working Paper*, n. 430, 2010. Citado na página 17.

KURTULUS, F. A. The impact of affirmative action on the employment of minorities and women over three decades: 1973-2003. 2015. Citado na página 13.

LEONARD, J. S. The impact of affirmative action on employment. *Journal of Labor Economics*, University of Chicago Press, v. 2, n. 4, p. 439–463, 1984. Citado na página 13.

LOURY, L. D.; GARMAN, D. College selectivity and earnings. *Journal of labor Economics*, University of Chicago Press, v. 13, n. 2, p. 289–308, 1995. Citado na página 18.

MASSEY, D. S.; MOONEY, M. The effects of america's three affirmative action programs on academic performance. *Social Problems*, Oxford University Press Oxford, UK, v. 54, n. 1, p. 99–117, 2007. Citado na página 18.

QUEIROZ, D. M.; SANTOS, J. T. d. Sistema de cotas: um debate. dos dados à manutenção de privilégios e de poder. SciELO Brasil, 2006. Citado na página 14.

ROTHSTEIN, J.; YOON, A. H. *Affirmative Action in Law School Admissions: What Do Racial Preferences Do?* [S.l.], 2008. Citado 4 vezes nas páginas 10, 18, 25, and 47.

SANDER, R.; JR, S. T. *Mismatch: How Affirmative Action Hurts Students It S Intended to Help, and Why Universities Won T Admit It.* [S.l.]: Basic Books (AZ), 2012. Citado na página 18.

SANDER, R. H. A systemic analysis of affirmative action in american law schools. *Stan. L. Rev.*, HeinOnline, v. 57, p. 367, 2004. Citado na página 17.

SMITH, J. P.; WELCH, F. Affirmative action and labor markets. *Journal of Labor Economics*, University of Chicago Press, v. 2, n. 2, p. 269–301, 1984. Citado na página 13.

VELLOSO, J. *Vestibular com cotas para negros na UnB: Candidatos e aprovados nos exames.* [S.l.]: Brasília: NESUB e Faculdade de Educação, Universidade de Brasília, 2005. Citado 3 vezes nas páginas 23, 45, and 53.

A List of Major and Normalized Entrance Exam Score statistics

Table 14 – Entrance Exam Score by Major

Major	Mean Score	Max. Score	Min. Score	S.d	Median Score
Business Administration	0.365	2.388	-1.660	0.780	0.576
Architecture	0.420	2.055	-1.424	0.759	0.658
Arts	-0.163	1.277	-1.840	0.676	-0.178
Human Sciences	-0.446	1.620	-2.229	0.661	-0.596
Computer science	0.543	2.660	-1.945	0.788	0.701
Biology	0.450	2.131	-3.025	0.819	0.725
Accounting	-0.335	1.022	-1.288	0.618	-0.335
Economics	0.209	2.105	-1.871	0.750	0.365
Science	-0.378	2.238	-2.217	0.697	-0.445
Social Sciences	-0.070	1.882	-1.837	0.671	-0.020
Social Communication[1]	0.502	2.397	-1.548	0.790	0.680
Law	0.831	2.515	-1.258	0.856	1.120
Physical Education	-0.093	1.500	-1.848	0.696	0.012
Nursing	0.214	2.089	-2.185	0.805	0.348
Sanitary Engineering	0.080	1.808	-1.872	0.906	0.161
Civil Engineering	0.573	2.496	-2.131	0.800	0.276
Production Engineering	0.819	2.584	-1.571	1.070	1.274
Electrical Engineering	0.575	2.531	-2.069	0.790	0.730
Mechanical Engineering	-0.448	1.471	-1.346	0.715	-0.544
Statistics	-0.046	1.758	-1.327	0.680	-0.083
Pharmacy	0.865	2.609	-1.437	0.909	1.214
Philosophy	-0.567	1.357	-2.754	0.638	-0.547
Physiotherapy	0.523	2.425	-1.724	0.865	0.728
Physics	-0.180	1.751	-2.473	0.699	-0.192
Geography	-0.255	1.558	-2.090	0.625	-0.222
History	-0.012	2.015	-1.589	0.670	0.074
Arts - Language	-0.133	2.022	-1.835	0.686	-0.072
Mathematics	0.072	1.928	-1.575	0.697	0.113
Medicine	1.514	2.683	-2.688	1.106	2.078
Nutrition	-0.287	1.536	-1.867	0.715	-0.416
Dentistry	0.579	2.241	-1.785	0.832	0.823
Pedagogy	-0.506	2.292	-2.293	0.618	-0.528
Psychology	0.413	2.328	-2.490	0.740	0.536
Chemistry	0.122	1.926	-1.722	0.976	0.309
Social Sciences - Social Service	-0.305	1.523	-2.679	0.616	-0.258
Information System	0.407	1.627	-0.920	0.620	0.530
Tourism	-0.287	1.183	-1.936	0.581	-0.253

[1] in red are the majors considered selective

B Summary Statistics

Table 15 – Summary Statistics

	All Students N= 18,726		Non Quota N=8,185		Quota N=3,936		Top 20 N=2,730		Bottom 20 N=1,043	
Variables	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
Female	0.527	0.499	0.525	0.499	0.525	0.499	0.516	0.500	0.532	0.499
Age in admission	20.303	17.267	21.112	4.322	21.112	4.322	19.069	28.001	20.643	32.137
Nonwhite	0.190	0.392	0.180	0.384	0.349	0.477	0.097	0.296	0.117	0.322
Selective Major	0.407	0.491	0.379	0.485	0.428	0.495	0.441	0.497	0.407	0.492
Area										
Humanities	0.505	0.500	0.502	0.500	0.481	0.500	0.510	0.500	0.525	0.500
Health	0.264	0.441	0.251	0.433	0.276	0.447	0.282	0.450	0.258	0.438
Science	0.222	0.416	0.234	0.424	0.241	0.428	0.191	0.393	0.217	0.412
Enrollment Status										
Active (in 2018)	0.027	0.162	0.031	0.172	0.058	0.233	0.003	0.051	0.006	0.076
On Hold (in 2018)	0.004	0.066	0.616	0.487	0.006	0.078	0.001	0.459	0.003	0.471
Completed	0.646	0.478	0.006	0.076	0.631	0.482	0.700	0.027	0.667	0.054
Cancelled	0.305	0.461	0.334	0.472	0.295	0.456	0.257	0.437	0.312	0.463
CGPA	72.023	21.808	71.582	21.989	69.70595	21.32642	74.649	22.292	72.380	21.133
Normalized Entrance Score	0.000	1.000	0.212	0.911	-0.275	0.931	0.028	1.148	-0.228	0.992

Table 16 – Correlation Matrix

	Log(wage)	Graduated	1st Year RGPA	Last year RCGPA	Quotas	Score	Selective Major	Health	Science	Female	Nonwhite	Age	Cia. Size	LE
Log(wage)	1.000													
Graduated	0.161	1.000												
1st Year Relative GPA	0.106	0.483	1.000											
Last year Relative CGPA	0.106	0.470	0.829	1.000										
Quotas	-0.214	-0.133	-0.092	-0.108	1.000									
Entrance Exam Score	0.179	0.185	0.065	0.030	-0.063	1.000								
Selective Major	0.271	0.187	0.033	-0.005	-0.033	0.357	1.000							
Health	0.112	0.217	0.036	0.007	-0.047	0.206	0.087	1.000						
Science	0.085	-0.238	-0.064	-0.022	0.027	0.024	0.295	-0.257	1.000					
Female	-0.099	0.175	0.156	0.166	-0.060	-0.062	-0.126	0.139	-0.274	1.000				
Nonwhite	-0.121	-0.067	-0.040	-0.044	0.269	-0.053	-0.034	-0.046	0.016	-0.056	1.000			
Age in Enrollment	-0.013	-0.046	-0.035	-0.031	0.050	0.011	-0.016	-0.031	0.006	0.000	0.030	1.000		
Company Size	0.436	0.024	0.028	0.030	-0.066	0.052	0.088	0.075	0.037	-0.078	-0.010	0.014	1.000	
Length of Employment	0.199	-0.123	-0.040	-0.049	0.004	-0.037	-0.006	-0.112	0.037	-0.158	0.039	0.070	0.153	1.000

C Dictionary

Table 17 – Dictionary

Variables	Meaning
Female	Dummy variable - 1 if female 0 if male
Age in admission	Student's age when he enrolled at UFJF
Nonwhite	Dummy Variable - 1 if nonwhite, 0 if white
Selective Major	Majors with a entrance exam score that are above half standard deviation from the mean
Area	
Humanities	Dummy variable - 1 for humanities majors, 0 otherwise
Health	Dummy variable - 1 for health majors, 0 otherwise
Science	Dummy variable - 1 for Science majors, 0 otherwise
Enrollment Status	
Active (in 2018)	Students that still enrolled in 2018
On Hold (in 2018)	Students that put their enrollment on hold
Completed	Students that graduated
Cancelled	Students that cancelled their enrollment at UFJF, before graduate
CGPA	
Normalized Entrance Score	Entrance score exam that were normalized so that they are distributed such as a N(0,1)
Company Size	Number of employees in the company
Length of Employment	How long the student is in the labor market