Affirmative Action in Centralized College Admission Systems*

Sebastián Otero[§]
(Job Market Paper)

Nano Barahona[†]

Cauê Dobbin[‡]

July 10, 2021

PRELIMINARY AND INCOMPLETE
PLEASE DO NOT CIRCULATE

[Link to the latest version]

Abstract

This paper empirically studies the distributional consequences of affirmative action in selective college education in the context of a centralized admission system. We examine the effects of a law implemented in Brazil in 2013, mandating all federal public universities to increase the number of seats reserved for public high-school students to half of the total. We find that after the policy was put in place, the student body composition of public institutions became more similar to that of the applicant pool. To study the overall distributional consequences of the policy, we develop and estimate a model of school choice and educational outcomes. We leverage the rules from the centralized admission system to simulate counterfactual allocation of spots under different affirmative action regimes. We find that the policy creates large predicted income benefits for targeted students while imposing a small cost on non-targeted individuals.

1 Introduction

Affirmative action (AA) is widely used across the world to increase the presence of underrepresented groups in more selective colleges and universities. It typically works by placing preferential weight on applications by students from marginalized groups at the expense of displacing non-targeted students. The prioritization of one group over another elicits polarized reactions that makes this one of the most divisive regulations affecting higher education systems (Arcidiacono and Lovenheim, 2016).[1]

*Preliminary draft: Apr, 2021. We are particularly indebted to Nick Bloom, José Ignacio Cuesta, Pascaline Dupas, Liran Einav, Matthew Gentzkow, Caroline Hoxby, Christopher Neilson, and Paulo Somaini for invaluable guidance and support on this project. We would like to thank Claudia Allende, Luis Armona, Arun Chandrasekhar, Michael Dinerstein, Marcel Fafchamps, Diego Jiménez, Ursulla Mello, Melanie Morten, Anna Popova and seminar participants at Stanford University for valuable comments and suggestions. We are especially grateful to Anna Popova for her insights, patience, and dedication in making this article better. We gratefully acknowledge financial support from the Stanford King Center on Global Development, the Stanford Center for Computational Social Sciences, the Stanford Institute for Economic Policy Research (SIEPR), the Haley-Shaw Dissertation Fellowship, and the NAEd/Spencer Dissertation Fellowship Program. All remaining errors are our own. [§]Stanford University. Email: sotero@stanford.edu. [†]University of California, Berkeley. Email: hbaraho@stanford.edu. [‡]Stanford University. Email: dobbin@stanford.edu.

[1]AA in the United States has been at the center of this dispute, with the Supreme Court having rejected a number of university admission procedures explicitly favoring disadvantaged groups (e.g., Regents of the University

The discussion surrounding the value of AA policies centers around a hypothetical tradeoff between equity and efficiency. Equity arguments revolve around the role of AA in reinforcing the equalizing role of colleges and promoting diversity as a pillar of a sustainable and healthy democracy (Singer, 2011; Alon, 2015).[2] While most policy efforts have been motivated by the equity merits of AA, much of the debate has focused on its impacts on the efficiency of the higher education system.

At the heart of the efficiency discussion is the fact that AA pushes targeted students into selective degrees by discriminating against allegedly more suitable and qualified candidates. There is contention around whether this results in net positive efficiency benefits because theory fails to provide unambiguous predictions of AA on winners and losers. On the one hand, AA policies could reduce the efficiency of the system by undermatching high-performing non-targeted students to colleges. In addition, it may even harm targeted students by placing them in schools for which they are ill-prepared (also known as the "mismatch theory") (Sander, 2004). On the other hand, AA could increase the efficiency by improving the matching between students and colleges if it helps to correct for test scores that are statistically biased against targeted students (Chetty et al., 2020). Similar efficiency arguments can be made if displaced individuals have access to alternative colleges that are not affordable to targeted students. In spite of such conflicting predictions, quantitative evidence on the the distributional and welfare consequences of AA is limited.

This paper leverages a large-scale AA regulation in college education implemented in Brazil in 2013 to provide new evidence on the impact of such policies on both beneficiaries and displaced students in the context of a centralized admission system.[3] The regulation mandated all federal public universities to increase the number of reserved seats for students from public high-schools to 50 percent of the total number of incoming students in every program.[4] The regulation was implemented in a staggered fashion over four years. Beginning in 2013, affected institutions were mandated to reserve a minimum of 12.5 percent of their vacancies for eligible students, with the minimum share increasing by 12.5 percentage points per year until reaching 50 percent in 2016.[5]

of California vs. Bakke, 1997). When AA is not banned, its use is voluntary and permitted when narrowly tailored.

[2]The equity arguments for AA revolve around three types of values: equity, diversity, and democracy. Taking each of these in turn, equity values are captured in the power of AA to level the playing field as access to higher education is a strong conduit for the social and economic intergenerational mobility of minorities (Alon, 2015; Chetty et al., 2020). Diversity values are increasingly held by universities - and upheld by the Supreme Court - because "the nation's future depends upon leaders trained through wide exposure to the ideas and mores of students as diverse as this Nation of many peoples" (Bakke 438 U.S. 265 (1978), 320). Democracy values of AA stem from its ability to reengage disenfranchised groups in the democratic process, thus combatting the existing massive representation of one group, which is undesirable in a democracy and likely to lead to political problems (Duflo and Banerjee, 2019)

[3]Law of Social Quotas, no. 12.711/2012.

[4]In the United States, selective universities doing AA implicitly subsidize the application score of students who belong to a given minority. In the rest of the world, many countries use quotas and reserve a share of seats in their college admissions. Some examples include India (cast, gender), France (residence), South Africa (race), Malaysia (ethnic), Sri Lanka (residence), Nigeria (residence), Romania (ethnicity), China (ethnicity), New Zealand (ethnicity), and Chile (class)

[5]Institutions had the freedom to choose whether to adjust to the full regulation (50 percent) immediately or to adjust gradually, as long as they were complying with the mandated minimum share of affirmative action spots in a given year.

The policy heavily targeted low-income and marginalized racial groups –who account for most of public high school enrollment– under the rationale that members of disadvantaged groups should not be underrepresented relative to their proportion in the community as a whole.[6]

We combine several datasets to a characterize the population affected by the AA policy and to determine its impact along several margins and outcomes of interest. We have access to detailed administrative education data including students' performance on the national university entrance exam (ENEM), application portfolios to public universities, and admission offers from these institutions. We combine these administrative data with the Brazilian Higher Education Census to track students' progress and academic success in their institution and degree programs over time. Finally, we use matched employer-employee records comprising virtually the universe of formal employment in Brazil, to create a mapping between academic progress and labor market income, and assess the impact of the law on predicted earnings. Overall, these datasets provide a rich and comprehensive characterization of all individuals affected by the AA policy.

Three features make the Brazilian higher education sector an ideal setting in which to study the distributional consequences of AA in selective public colleges. The first is that the AA policy is embedded in a centralized scheme, which produces transparent application and admission data, and clear allocation rules. Admissions into public institutions in Brazil are organized through an online centralized platform (SISU) that assigns students to degrees solely depending on their degree preferences, ENEM test scores, and admission pool. In contrast to the standard assignment mechanism used in centralized college admission systems, in which all students apply through a common admission pool, the Brazilian system allows students to apply through different admission pool based on their AA status. Targeted students are eligible for either a reserved or an open seat, while non-targeted students can only apply for the open slots. After submitting an application, admission to a given degree is feasible for a student if its admission cutoff is below the student's test score. These cutoffs are specific to each degree and admission pool tuple.

The second feature that makes the the Brazilian setting one of special interest, is that public federal institutions are of substantially better quality than their private counterparts, as well as being free of charge, which makes them very attractive for high-performing students from both low and high-SES backgrounds. In practice, federal universities play a similar role to flagship state universities in the United States. They are typically the most prominent, elite, and selective universities in their specific states. In 2011, the average admission threshold across all degrees in public institutions was one standard deviations higher than the score of the average ENEM taker. As a result of their attractiveness and selectivity, public institutions are highly segregated.

[6]Such systems are also known as reservation policies. The design of these reservations has been studied in several real-life applications. These include school choice in Boston (Dur et al., 2018) and Chicago (Dur et al., 2020), college choice in Brazil (Aygun and Bo, 2021), allocation of H-1B visas in the United States (Pathak et al., 2020) and public sector jobs in India (Sonmez and Yenmez, 2021).

The third feature is that Brazil's mixed affirmative action model (i.e. some racial/ethnic criteria plus some socio-economic criteria) is more characteristic of "modern" affirmative action policies than United States policies which i) are opaque because of the holistic admission system, and ii) disproportionately emphasize race for historical reasons.

We exploit this setting to show how AA regulations can change the composition of schools by differentially affecting the probability of admission of targeted and non-targeted individuals. We start by quantifying the extent to which the policy has increased the enrollment of targeted students in federal universities at the degree level. We show that after the policy was put in place, the student body composition of public institutions became more similar to that of the applicant pool. In 2012, one year before the AA regulation was passed, non-targeted students represented 55% and 76% of admissions to public universities in general and to the 10% most selective degree programs, respectively. These proportions stand in stark contrast to the 20% share of overall high-school graduates that these students represent. In 2016, when the AA policy was fully in place, the share of non-targeted admits was substantially closer to their proportion of the high-school population: they represented 35% and 50% of admission to public universities and to the 10% most selective public degree programs, respectively.

Along the same lines, at the individual level, we show that the policy also affected the probability of student admission. We find that, after conditioning for test scores, targeted students in the top quintile of the score distribution are 20 percentage points more likely to attend a public institution than non-targeted students in 2016. This difference represents a striking increase relative to the 5 percentage point difference observed in 2011. Our results suggest that these disparities are largely driven by a growing difference between admission thresholds for targeted and non-targeted students. In 2016, the average difference in admission thresholds for targeted students and non-targeted students was 0.4 standard deviations in ENEM test scores.

To study the overall distributional consequences of the AA policy –that is, how much winners win and how much losers lose– we leverage the rules from the centralized mechanism and simulate a counterfactual allocation of spots in a regime without AA. To do so, we propose a framework that characterizes the allocation of students to spots under different AA schedules, based on an assignment mechanism and its inputs –namely, students preferences for degrees, scores distributions, quota status and degrees capacities. The counterfactual allocation is represented by a new vector of admission cutoffs governing admission into degrees.

Equipped with the quota-specific and counterfactual admission thresholds, we construct observed and counterfactual personalized choice sets of feasible degrees for every student. We use these to identify the set of individuals who were benefited (i.e. winners) and displaced (i.e. losers) by the AA regulation. We define winners as those targeted individuals who in the observed data attend a degree that is not available in their counterfactual choice set. Losers, on the flip side, are defined as the set of non-targeted individuals that in the counterfactual regime attend a degree that is not available in their observed choice-set.

We define the consequences of the policy –our object of interest– as how much benefited individuals gain and how much displaced individuals give away when comparing the observed to the counterfactual assignment. We use a potential outcome framework to show that these gains and losses are characterized by how admission scores, degree attendance, and the outside option of not attending a public school, affect a given outcome of interest (e.g. wages). Heterogeneity in the mapping between these variables and the outcome of interest across targeted and non-targeted students can parsimoniously capture the wide variety of possible distributional consequences of AA.

We then introduce a choice and outcome model to bring the objects from the conceptual framework to the data. We start by fitting a discrete choice model to applicant's allocations using their personalized choice sets (Fack et al., 2019). We use these estimates with two purposes. First, recovering preferences allow us to simulate the counterfactual degree to which students would be assigned in a regime without AA. Second, we use a "control function" approach to construct the average gains on test scores and earnings for attending these counterfactual degrees. The main advantage of this approach is that it allow us to correct these estimates for selection on unobservables. We implement this selection correction by following the multinomial logit control function estimator of Dubin and McFadden (1984) and Abdulkadiroglu et al. (2020).

Identification in our model relies on two exogenous shifters that increase the degree choices that are available for individuals. The first set of shifters consists of the observed degree-specific admission thresholds that create discontinuities for targeted and non-targeted applicants. Two individuals with almost identical observed characteristics and unobserved tastes for degrees, but who fall on different sides of a degree's admission cutoff, will have differences in potential outcomes that can only be explained by differences in degree availability (and attainment). The second shifter exploits exogenous variation in student test scores in the national exam that stems from random assignment to graders who vary in their degree of strictness.[7] Comparing equally skilled students who vary in their personalized choice sets because of their grader assignment allow us to estimate the impact of the AA policy for individuals at different points of the skill distribution

We then use our model to estimate the distributional consequences of AA. We simulate the allocation of students with and without the AA policy. We find that, under AA, targeted students increase their enrollment in selective degrees while displacing non-targeted students to less selective alternatives. We use our estimate of potential outcomes to calculate the gains and losses for beneficiaries and displaced individuals. This exercise suggests that targeted individuals are more likely to switch to the outside option when their preferred degree is removed from their choice set. In addition, the outside option provides less value for targeted individuals than for non-targeted individuals. Accordingly, we find that the AA policy creates large income benefits for the targeted

[7]This instrumental variable approach is also known as the judge fixed-effect instrument. See Doyle (2008) and Dobbie et al. (2018) as examples.

group while imposing a smaller cost on non-targeted individuals.

Related Literature: Our paper is related to several strands of the literature. First, it contributes to the literature studying the impact of AA policies in higher education. Several studies have focused on the enrollment margin, coming to an overall conclusion that such policies substantially increase the probability of admission to elite colleges for racial minorities. A first set of studies in the United States evaluate state-wide laws mandating universities to switch from race-sensitive to race-neutral admission policies. For instance, Card and Krueger (2005) find that the admission rates of black and Hispanic students fell in Texas and California when these states eliminated the use of AA. Interestingly, Long (2004) finds that eliminating AA and moving to a top $x\%$ programs model does not compensate for the decrease in minority attendance at top-tier institutions. Another set of papers focus on conducting structural policy analysis of college admissions in the U.S. (Arcidiacono, 2005; Epple et al., 2008; Howell, 2010; Bodoh-Creed and Hickman, 2019). These, similarly, consistently find that race-neutral admission policies would reduce the representation of minority students in more selective colleges. [8]

The main drawback of AA studies in the U.S. is that they are unable to identify the AA status of students and must therefore use race or ethnicity as an imperfect proxy for AA eligibility. This limitation has been overcome by studies outside of the United States, where researchers have access to AA implemented in fully transparent admission systems. In India, Bagde et al. (2016), finds that an AA quota policy that reserved 20% of seats across more than 200 engineering colleges for low-caste or female students, increased college attendance by targeted students, particularly at higher-quality institutions. Other studies in Brazil and Israel have focused on how AA implementation shapes student body demographics in a given university, reaching similar conclusions (Francis and Tannuri-Pianto, 2012; Estevan et al., 2018; Alon and Malamud, 2014).

Most recently, Mello (2020) exploited progressive rollout of the Lei de Cotas - our AA policy of interest - to show that it increased the share of black and low-income students enrolled in public institutions. We are able to build on this finding in two main ways. First, we have the unique benefit of having access to detailed application data from the centralized admission system. This allows us to leverage admission rules to estimate counterfactuals in the absence of AA. Thus, we are able to study the impact of AA on both targeted and displaced students. Second, in addition to studying the immediate effects of AA on higher education access, we analyze its metium-term consequences, in terms of academic trajectories and graduation rates.

The structure of the article is as follows. The next section introduces the setting and provides extensive details of the policy and its implementation. Section 3 discusses the data and provides descriptive facts about the Brazilian college admissions. Section 4 analyses the effect of of the AA policy on public colleges admissions. Section 5 presents a model that characterizes de impacts

[8] See Alon (2015) and Arcidiacono and Lovenheim (2016) for an extensive review of this literature.

of AA in a centralized admission system. We estimate this model in Section 6 and present the parameters in Section 7. Section 8 examines the effect of alternative policy counterfactuals and Section 9 concludes.

2 Institutional background and regulation

In this section, we describe the higher education sector in Brazil, and the admission process to public institutions. We also document a strong correlation between the college entrance exam and the traits targeted by the affirmative action policy. Finally, we describe the details of the affirmative action regulation and its implementation over time.

2.1 Higher education in Brazil

The Brazilian post-secondary education system is a mixed system composed of public and private institutions. The system is highly liberalized and market-oriented; approximately 76% of students are enrolled in private fee-charging colleges.[9] The remaining 24% attend public institutions. These institutions charge virtually no tuition fees and are associated with the federal, state, or municipal government, depending on the source of funding. In 2016, there were a total of 107 federal institutions, which served nearly 1.25 million students, and offered more than 6,000 different degrees. Federal institutions comprise 63 universities and 44 vocational institutions, with the former accounting for the vast majority of federal enrollment. Table A.1 presents the number and share of students, institutions, and degrees, by sector.

Federal public universities in Brazil are, in most cases, elite and highly selective. These institutions play a similar role to flagship state universities in the United States.[10] They are typically the most prominent and important universities in their specific states and are, on average, of substantially higher quality than their private counterparts, as measured by student learning, infrastructure, and the quality of peers and faculty.[11] The median state in Brazil has just one selective federal university. As a result of their differential quality and tuition-free policy, public institutions are highly over-subscribed.

Prospective students in Brazil, as is common in most countries, apply in advance to a particular degree program: a specific major at a specific institution (e.g. Law at the University of Sao Paulo). Obtaining a postsecondary degree normally requires 3-6 years for bachelor's degrees, 4-5 years for teaching degrees, and 2-3 years for vocational degrees. In addition to selecting a degree, students must choose the teaching shift in which to study their degree: morning, evening, night, or full-

[9] As a benchmark, in the United States, 3/4 of students attend public colleges, while only 1/4 attend a private institution.

[10] More than 85% of enrollment is by in-state students.

[11] Figure B.1 shows the distribution of quality as measured by an index ranging from 1 to 5 prepared by the Ministry of Education to evaluate degrees.

time.[12]

2.2 College admissions: ENEM and SISU

Several million students take the ENEM at the end of the academic year, with the aim of gaining access to higher education.[13] The ENEM exam is an extremely competitive standardized exam, that consists of four multiple choice tests–on Math, Language, History, and Science–and one written essay. ENEM test scores can be used to gain admission to most public and some private institutions, and are the only merit-based criterion used by the Ministry of Education to assign financial aid to students attending private institutions.

Until 2010, college admissions worked in a fully decentralized way and each institution administered its own specific entrance exam (known as *vestibular*). Since then, federal and state institutions can opt to participate in SISU, a centralized digital platform that matches students to degree programs according to their ENEM test scores, and degree preferences. By 2016, over 93% of the 204,633 incoming students in federal institutions entered through this centralized admission system and 103 out of the 107 federal institutions participated in it. In the same year, about 57% of the 4.8 million ENEM takers applied to a degree program using the SISU platform. Note that students taking ENEM in a given year can only use their score to participate in the SISU process in the following year.

Students participating in SISU can submit and rank up to two program choices among the set of available programs in the system. A program is defined as a degree, institution, and shift tuple. Students are then matched following an iterative deferred acceptance mechanism based on ENEM scores.[14] As opposed to the standard deferred acceptance process, students are sequentially asked to submit rank ordered lists over the course of several "trial" days.[15] At the end of each day, the system produces a cutoff score representing the lowest score necessary to be accepted into a specific program, and students are allowed to change their degree preferences given the newly available information. The results of the last day are final and determine the acceptances for every degree.[16]

2.3 Affirmative action regulation

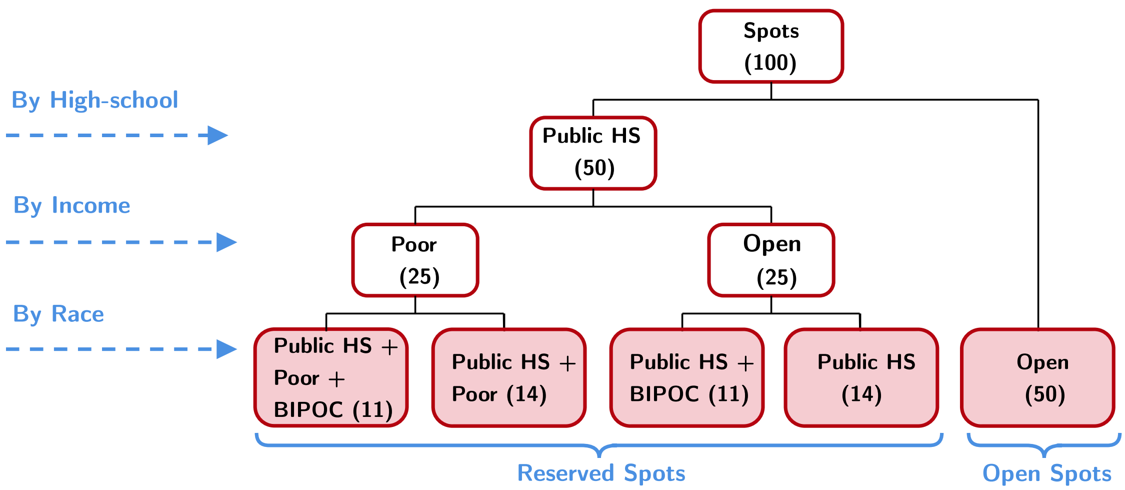
In August 2012, the Brazilian federal government passed the *Law of Social Quotas* (Lei de Cotas, no. 12.711/2012) requiring all public federal institutions to reserve half of their admission spots for students coming from public high schools. The regulation sought to reverse the racial and income inequality in university access. Under the regulation, only students coming from public high schools are eligible to compete for the affirmative action vacancies, while the remaining half of

[12] A student can only attend one shift and is not allowed to switch between them.
[13] It is the second-largest standardized university entrance exam in the world after the Gaokao in China.
[14] Degrees may use different weights for each of the ENEM sub-parts.
[15] This is the same as the mechanism currently used by Inner Mongolia, China university admission systems.
[16] See Bo and Hakimov (2020) for formal properties of this mechanism.

vacancies remain open for broad competition.[17] Reserved seats are then further distributed among students from low-income families, and who are of African or indigenous descent. Throughout the paper we refer to individuals who are eligible for an affirmative action spot as targeted students.

Figure 1 summarizes this distribution by presenting the shares of affirmative action students coming from different demographics. For every 100 university spots, 50 are affirmative action spots for public high-schools graduates. Of those reserved spots, 25 go to poor students with monthly household income per capita below 1.5 minimum wages (about 1,500 BRL, equivalent to 360 USD), and 25 go to non-poor students. Finally, a percentage of the spots in both categories is set aside for black, brown and indigenous students, in accordance with the racial makeup of each of Brazil's 27 states.[18] Overall, this results in five different affirmative action categories to which prospective students can apply.

Figure 1: Affirmative Action Regulation



Notes: This figure describes the affirmative action policy. Under the regulation, for every 100 university spots, 50 are affirmative action spots for students who attended public high schools. Those affirmative actions spots are then divided equally by income, with 25 going to poor students and 25 to non-poor students. Finally, each group of 25 spots is distributed to reflect the proportion of non-white individuals in the population of a given state. This example uses a proportion of 54%, which is the combined share of black, brown, and indigenous people in the state of Minas Gerais, as reported by the Brazilian National Bureau of Statistics (IBGE) in 2012.

The regulation was implemented in a staggered fashion over four years. Beginning in 2013, affected institutions were mandated to reserve a minimum of 12.5% of their vacancies for eligible students, with the minimum share increasing by 12.5 percentage points per year until reaching 50 percent in 2016. Institutions had the freedom to choose whether to adjust to the full regulation (50 percent) immediately or to adjust gradually, as long as they were complying with the mandated minimum share of affirmative action spots in a given year.

[17] Students attending private high schools with full scholarships are also eligible to compete for affirmative action vacancies.

[18] Individuals identified as *pardos* are mixed-race with black ancestry. According to the 2010 Census, the Brazilian population is 47.5% white, 43.4% brown (*pardo*), 7.5% black, 1.1% asian and 0.42% indigenous (IBGE 2011).

Figure B.2 illustrates the staggered implementation of the policy between 2012 and 2017. Here, an observation is a degree program and each panel presents a histogram for the share of students that entered university through affirmative action in a given year. In every year since the regulation was introduced, we observe increased bunching of quota admissions above the 50% threshold. In addition, in 2013, 2014, and 2015 there is substantial bunching above the minimum thresholds corresponding to those years, until the 50% quota is surpassed in 2016. Between 2013 and 2017, virtually all degree programs complied with the minimum quota mandated for the given period. This illustrates that federal institutions complied perfectly with the regulation.

3 Data and descriptive evidence

3.1 Data sources

We combine several datasets to a characterize the population affected by the affirmative action policy and to determine its impact along several margins and outcomes of interest. We have access to detailed administrative education data including students' performance on the national university entrance exam (ENEM), application portfolios to public universities, and admission offers from these institutions. We combine these administrative data with the Brazilian Higher Education Census to track students' progress and academic success in their institutions and degree programs over time. Finally, we link these datasets to matched employer-employee records comprising virtually the universe of formal employment in Brazil.

Test scores: The first dataset we employe contains test score data for the universe of students taking the university entrance, ENEM. We observe these data for the period 2009-2015, which correspond to the 2010-2016 admission period. The number of exam takers has increased from 2.4 million individuals in 2009 to 4.8 million in 2015. These data include individual-level test scores on each of the components of the test, as well as answers to a detailed survey including questions on socioeconomic background and student perceptions.

Centralized admissions process: We complement the test score data with records from the centralized admission system, SISU. These data only covers applications to federal and state institutions. We focus our attention only on applications to federal institutions, as state institutions were not mandated to comply with the affirmative action regulation. The dataset is at the application level, and only includes records from the final rank-ordered list submitted by each student. For every application, we observe the ranking of the degree in the student preference list, the affirmative action group of the applicant, and the student ranking among all applicants from a given degree and affirmative action group tuple. We also observe the students who were offered an admission and the admission cutoff for every degree and affirmative action group tuple. We observe these data for the 2016 admissions period.

Higher education census: The third dataset we use for our analysis is the Brazilian Higher Education Census. This database contains information on every student enrolled at any higher education institution in Brazil. This allows us to observe the educational path of every student in any degree program between 2009 and 2019. In a given year, the unit of observation is at the degree-student level, and includes a variable indicating if the individual graduated, dropped out, or is successfully enrolled at the end of the academic year. This data is of very high-quality as most institutions have their own systems integrated with the census in real time. In addition to the student data, this database also includes administrative information at the level of degrees and institutions, allowing us to characterize both over time.

Matched employer-employee data: Finally, we combine the Higher Education Census data with matched employer-employee annual administrative records (also known as RAIS) from the Ministry of Labor. This is considered to be a high quality Census of the Brazilian formal labor market. This dataset includes variables at the firm and worker levels, such as payroll, contracted hours, hiring and firing dates, and occupation. We use these data to produce earning profiles for every degree and student type. We observe these data for the 2017 period.

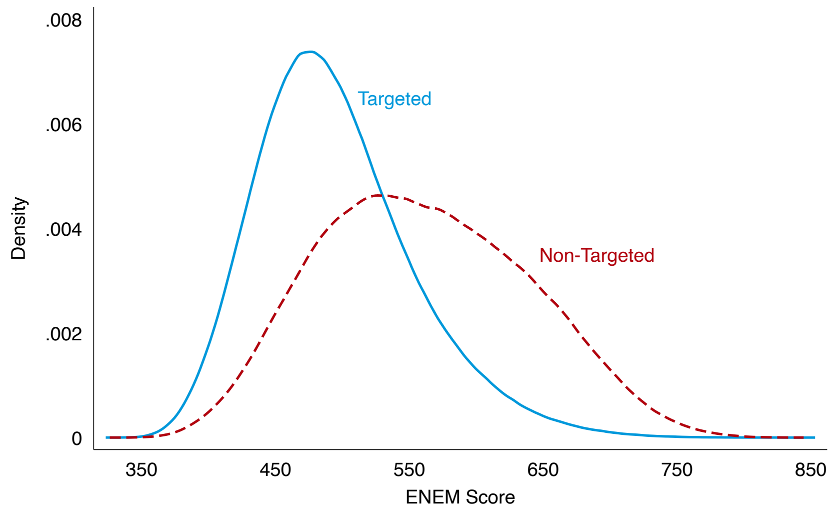
3.2 Descriptive facts about Brazilian college admissions

One of the main motivations behind the implementation of the AA regulation is that ENEM test scores are strongly correlated with socioeconomic status and high school demographics. As such, selective admissions to public institutions are highly segregated in favor of students from wealthier socioeconomic backgrounds. We provide descriptive evidence for the 2016 admission year to show large differences in test performance by targeted status. In addition, we show that the effect of the difference in performance on access to federal institution is partly mitigated by lower admission thresholds for targeted individuals.

Test participation and performance by targeted status: In 2016, targeted individuals represented nearly 85% of the 4.8 million ENEM takers. Figure 2 shows the distribution of the average ENEM score by targeted status. We observe that the average targeted student scored 494 points, while the average non-targeted student scored 560 points. This 66 points difference represents a 1 standard deviation (SD) difference in performance across targeted and non-targeted individuals.

Applications and spots by targeted status: In the first semester of 2016, the centralized admission system offered 200,877 spots across 4,900 degrees in 101 federal institutions. Table 1 presents summary statistics on the number of spots and student applicants to open and reserved spots. Nearly 46% of spots at federal institutions were open for anyone to apply, 48% were reserved for targeted individuals as mandated by the regulation, and the remaining 6% were reserved for institution-specific quotas (e.g. place-based affirmative action). Notably, across all admission pools, selectivity–as reflected in the ratio of spots to applications–was approximately 5%, comparable to many Ivy League colleges in the U.S.

Figure 2: ENEM score distributions by targeted status



Notes: This figure shows kernel density plots of the average ENEM score distribution of targeted and non-targeted individuals. The sample is the universe of ENEM takers who had positive test scores in the 2015 ENEM test. The average test score include math, language, natural science, and social science. The average score is 504 and the standard deviation is 66 points. Targeted students are defined as those who are eligible for any of the affirmative action vacancies.

By normalizing test score admission cutoffs in terms of the average admission cutoff for open seats, we observe that admission cutoffs for targeted students range between 0.43 and 0.88 SDs below that of non-targeted students, depending on the specific admission pool. These differences in admission cutoffs somewhat compensate for the difference in ENEM test scores between targeted and non-targeted individuals.

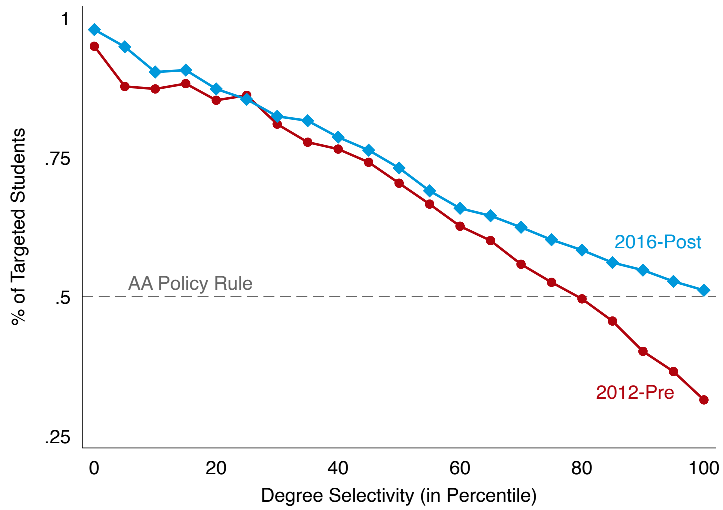
Table 1: SISU Applicants

	Spots		Applications		# Spots to	Cutoff
Admission Pool	#	%	#	%	Apps (%)	in SD
Open	92,409	46.0	1,978,077	44.0	4.7	0.00
Reserved, by Law						
Public HS	18,850	9.4	412,892	9.2	4.6	-0.43
Public HS & Income	19,768	9.8	482,761	10.7	4.1	-0.69
Public HS & Race	28,475	14.2	522,994	11.6	5.4	-0.76
Public HS & Race & Income	29,597	14.7	964,483	21.5	3.1	-0.88
Reserved, Other	11,778	5.9	135,491	3.0	8.7	-0.75
Total	200,877	100	4,496,698	100		

Notes: Own elaboration based on SISU microdata from 2016. This table shows the breakdown of the number of applicants and spots for each of the different admission pools. The "Other" admission channel includes other affirmative action initiatives that are not mandated by the federal regulation.

Composition of the student body: Next, we show that the student body composition in federal institutions became more diverse after the implementation of the AA policy, driven by an increase in the representation of targeted students. Figure 3 presents the average share of targeted students in federal institutions by degree selectivity. We observe a large increase in the share of targeted students for degrees above the 50th percentile of selectivity.

Figure 3: Impact on student body composition



Notes: This Figure presents the average share of incoming targeted students in federal institutions by degree selectivity. An observation is a degree and shift tuple. We keep degree programs that exist between 2010-2017. Selectivity is defined as the average score of incoming students. The dashed line depicts the 50% AA policy rule mandated by the regulation.

3.3 Educational outcomes by targeted status

We explore the educational landscape faced both by targeted and non-targeted students in 2016, the year when the regulation was fully implemented. We present a series of educational outcomes conditioning on student ENEM scores. In Figure B.3 we present college attendance outcomes. We find that targeted individuals are substantially more likely to enroll at a federal institution relative to non-targeted with similar test scores. This result is expected as a direct consequence of the AA policy lowering the admission threshold for targeted students. The differences are larger for students in the upper part of the score distribution where the regulation is more effective. In turn, we observe that non-targeted individuals are more likely to enroll at a private institution.

In Figure B.4 we focus on what happens to those who enroll in federal institutions. We observe that conditioning on test scores, targeted students are more i) likely to enroll in better degrees, ii) more likely to be enrolled at a better degrees 4 years after, iii) and less likely to dropout or switch to another degree. In Figure B.5 we show that these patterns are reversed for targeted students enrolling in the private sector. In addition, targeted students attend cheaper degrees than their non-targeted counterpart. Remarkably, targeted students are substantially more likely to receive financial aid, either through discounted prices, student loans or grants. This last result helps mitigate the differences in access to more selective degrees in the private section, thus improving the value of the outside option for targeted students.

3.4 Predicted income as academic progress aggregator

Next, we ask how to aggregate educational outcomes. The ideal outcome would aggregate both the extensive and intensive margin. The extensive margin captures whether the student attended and/or graduated from a higher education institution. The intensive margin, on the other hand, would capture the academic trajectory of individuals together with the quality and the field of study of the degree attended. The ideal variable to wrap together all these margins is labor market income. Unfortunately, the law was fully implemented in 2016 and it is too early to observe labor market outcomes.

Instead, we use predicted income as currency to aggregate academic progress. For the 2016 sample, we observe college enrollment together with detailed academic progress over four years, from 2016 to 2019. For older cohorts, we observe academic progress and income. We use individual-level microdata of cohorts entering college between 2010 and 2012, and their observed labor market income in 2017, and estimate a model linking academic progress to income. Then we use these estimates to create predicted income measures for the 2016 sample. In Appendix C we discuss the econometric implementation of this procedure.

4 Motivating evidence [SKIP THIS SECTION]

In this section we study the causal impacts of increasing the number of reserved seats using shift-share instrumental variable regressions. We leverage the staggered implementation of the regulation across universities and over time to study the effect of the AA regulation on a host of outcomes.

4.1 Sample for the main analysis

Our master sample of municipalities is the universe of 5,571 municipalities in Brazil. All of them report to have ENEM takers in every year.

4.2 A measure of exposure to the affirmative action policy

We create a measure of exposure to the AA policy based on the exposure to reserved seats in federal universities. Specifically, let r_{jt} denote the share of reserved seats at institution j in year t, and let s_{mt} be the share of ENEM takers that reside in municipality m and enroll at public institution j. We use these variables to create an exposure measure at the municipality level, which is our treatment variable:

$$x_{mt} = \sum_{j \in \mathcal{J}} s_{mt} \cdot r_{jt}$$

Of interest is a causal effect or structural parameter β, relating treatment to outcomes by

$$x_{mt} = \gamma z_{mt} + \delta_m + \delta_t + \nu_{mt}$$
$$y_{mt} = \beta x_{mt} + \delta_m + \delta_t + \epsilon_{mt}$$

where ϵ_{mt} is an unobserved residual. We create an instrument for x_{mt} leveraging the difference between the share of pre-policy reserved spots, and the 50% share mandated by the regulation. The source of this variation is very salient in Figure B.2. In 2012, the year before the policy was implemented, 37% of degrees in federal institutions did not offered reserved spots, while another sizable 28% of the degrees offered more than 50% of their spots through reserved seats. Let g_j denote the difference between pre-policy and mandated reserved shares at institution j:

$$g_j = 2 \cdot \max\left(0, \left[0.5 - \frac{\text{reserved spots}_{j,2012}}{\text{total spots}_{j,2012}}\right]\right)$$

Institutions with $g_j = 0$ had 50% or more reserved seats in the pre-policy period and thus had little exposure to the AA regulation. In contrast, institutions with $g_j = 1$ did not offer any reserved spots and were highly exposed to the policy. We combine this variation with municipality-level exposure share weights based on the share of students in that municipality that enroll at institution j in 2012. We then interact those averages with year-specific dummies. Accordingly, our instrument z_{mt} is given by:

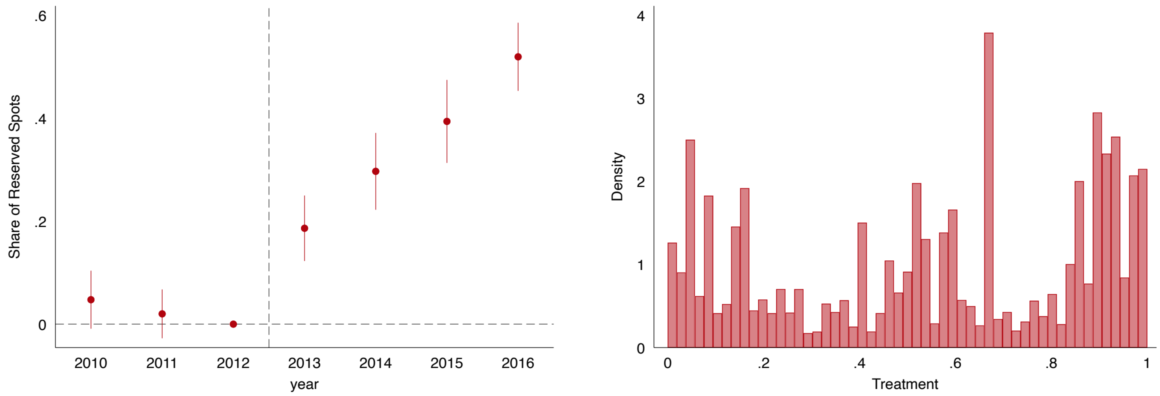
$$z_{mt} = \sum_k \mathbb{1}\{k = t\} \cdot \sum_{j \in \mathcal{J}} s_{m,2012} \cdot g_j$$

In Figure 5(a) we show the coefficients of a first-stage regression between the institution level affirmative action shares r_j and the variation arising due to the policy g_j

5 A model of affirmative action in centralized college admission systems

We propose a framework that characterizes the effects of affirmative action in selective public institutions in the context of a centralized college admission system. We consider a system in which students rank their preferred degree programs, and institutions rank applicants using a priority index comprised of standardized test scores. Our framework embeds an affirmative action regulation within such admission systems by reserving admission spots for the group targeted with preferential treatment. Because the number of spots is fixed, the affirmative action policy operates by pulling in applicants from the targeted group at the cost of pushing out non-targeted students with higher academic scores.

Figure 4: Exposure to the AA regulation



(a) Institution-level First Stage (b) Treatment distribution

Notes: Figure in Panel (a) plots the coefficients of a regression of the share of reserved seats reported by institutions, r_{jt}, on the exposure measure g_j, interacted with year dummies. The regression also includes year and institution fixed effects, and standard errors are clustered at the institution level. Panel (b) shows the distribution of the treatment measure created at the municipality level (weighted by the number of ENEM takers in 2012)

5.1 Environment

We consider a set of individuals indexed by $i \in \mathcal{I} = \{1, ..., n\}$, applying to a finite set of selective college degree programs in public institutions through a centralized platform. Let $\mathcal{J} = \{0, 1, ...J\}$ denote the set of degrees indexed by j offered across all public institutions, where $j = 0$ represents the outside option of either attending a private institution or not attending college.

Students have preferences over degrees based on a strict ordering, $\succ_i$. Degree programs also have preferences over applicants based on students priority scores $s_i = \{s_{i1}, ..., s_{iJ}\}$. Scores are very fine so that no tie-breaker is needed. We allow degree-specific scores which may arise due to degree-specific exams or to degrees assigning different weights for different components of a single entrance exam. For example, ENEM, the entry exam in Brazil, have different examinations for math, language, social science, natural science and writing.[19]

There are two subgroups within the student population, defined by their affirmative action status $t_i \in \{0, 1\}$. Let $t_i = 0$ represent the non-targeted students and $t_i = 1$ targeted students. The status t_i of each student is observable. A student is fully characterized by their type $\theta_i = (\succ_i, s_i, t_i)$. That is, the combination of an applicant's preferences, priority scores across all degrees, and affirmative action status.[20] We denote the set of all student types by $\Theta = \bigcup_{i \in \mathcal{I}} \theta_i$.

Spots at degrees are constrained by a strictly positive capacity vector, $q = \{q_0, ..., q_J\}$. An

[19]While STEM degrees tend to place higher weight on math and natural science, more humanities oriented degrees give more importance to language and social science.

[20]The definition of a student type as a combination of preferences, priorities, and status is similar to that of Abdulkadiroglu et al. (2017).

affirmative action regulation is in place such that for every degree, a share $\omega \in [0, 1]$ of the spots is reserved for applicants from the targeted group. The remaining share is open to all individuals. As such, for any given degree j, ωq_j spots are reserved for targeted students and $(1-\omega)q_j$ spots are open to all individuals. We assume $q_0 = \infty$ since the outside option of not attending a public institution is available to all students.

The centralized mechanism applies a student-proposing deferred acceptance algorithm to generate degree assignments. The inputs to the mechanism are student types Θ, school capacities q, and reservation shares ω. When reserved slots are processed, targeted students with the highest priority score receive them. When open slots are processed, members of any group (targeted or otherwise) are admitted in order of their priority score. When a student qualifies for both a reserved and an open spot, the set of admission rules must specify the relative precedence of different admissions channels. In our empirical application, we assume that spots are horizontally reserved. That is, reserved spots are allocated first to targeted students based on priority scores, and next all open spots are allocated to the remaining individuals with the highest priority scores.[21]

Let $\varphi(\Theta, q, \omega) = \mu$ denote the matching produced by mechanism φ for the problem (Θ, q, ω). The matching is a function $\mu : \Theta \to \mathcal{J}$, such that (i) $\mu(\theta_i) = j$ if student i is assigned to degree j, and (ii) no degree is assigned more students than its capacity. Because the mechanism implements a deferred acceptance algorithm, assignments μ between students and degrees are unique and stable (Gale and Shapley, 1962; Abdulkadiroglu, 2005).[22] The stability property of the mechanism implies that each student enrolls in their most preferred program for which they are eligible.

The matching μ has a unique representation in terms of a vector of market clearing admission cutoffs $c_{tj}(\mu)$ for each degree program and affirmative action status combination (Azevedo and Leshno, 2016).[23] A cutoff $c_{tj}(\mu)$ is a minimal score s_{ij} required for admission at degree j for students with affirmative action status t.[24] Since targeted students can be admitted through the reserved or the open spots, admission cutoffs for targeted students are always lower than for non-targeted individuals, i.e. $c_{0j} \geq c_{1j}$ for all $j \in \mathcal{J}$. Because the outside option is always feasible, we assign it an admission cutoff score of $-\infty$.

As in any strict-priority mechanism, the availability of options will be different for students

[21]The literature has established a difference between horizontal and vertical affirmative action depending on which of the spots (reserved or open, respectively) are processed first. A horizontal reservation is a "minimum guarantee" in the sense that it only binds when there are not enough targeted students who receive a spot on the basis of their test-score alone. A vertical reservation works under a "over-and-above" basis. This means that targeted individuals receiving an spot solely on the basis of their priority score does not count towards a vertically reserved position (Sonmez and Yenmez, 2021). Our framework encompasses both types of affirmative action.

[22]A matching μ is stable if there is no student-degree pair (i, j) where: (i) student i prefers degree j to their assignment, and (ii) student i has higher priority than some other student who is assigned to j

[23]Other mechanisms also have a unique representation in terms of admission thresholds as long as the matches between students and degree are pairwise stable (Agarwal and Somaini, 2018).

[24]For instance, the degree of psychology at the University of Brasilia would have two admission cutoffs: one for reserved spots students and another for open spots.

according to their priority score. Let $\Omega_i(\mu) = \{j \in \mathcal{J} \mid s_{ij} \geq c_{t_i j}(\mu)\}$ represent the *feasible choice set* for individual i, defined as those degrees to which they could have gained access based on their score and affirmative action status under a given matching μ. Let the variable $D_i(\mu) = \{j \in \Omega_i \,|\, j \succeq k \text{ for all } k \in \Omega_i(\mu)\}$ denote the preferred option in the feasible choice set defined by $\Omega_i(\mu)$. In other words, $D_i(\mu)$ represents the highest option ranked by individual i among the degrees to which they could have been admitted. We refer to this option as the *preferred feasible degree*. From the stability condition we know that $D_i(\mu) = \mu(\theta_i)$.

The realized outcome for student i is given by $Y_i(\mu) = \sum_j \mathbb{1}\{D_i(\mu) = j\} \cdot Y_{ij}$, where $D_i(\mu)$ indicates the degree attended, and Y_{ij} denotes the potencial value of some outcome of interest for student i if they attend degree j.

5.2 Counterfactual admission cutoffs, choice sets, and preferred degrees

The main advantage of studying affirmative action in the context of a centralized system is that it follow systematic and transparent admission rules. The idea is that by leveraging the assignment rules from the mechanism, we can characterize student allocations in a regime with a different affirmative action schedule ω'. For ease of exposition, we think of ω' as an increase in the shares reserved to affirmative action students.

The new schedule, together with the students types and degree capacities, result in a counterfactual matching function $\varphi(\Theta, q, \omega') = \mu'$. In turn, this allocation can be represented by a new vector of of admission thresholds $c_{tj}(\mu')$.[25] In this scenario, the feasible choice sets, preferred feasible degree, and ultimately the realized outcomes, would also change as a result of the changing admission cutoffs. These objects are represented by $\Omega_i(\mu')$, $D_i(\mu')$ and $Y_i(\mu')$, respectively. Combining these elements we can characterize the set of individuals who were effectively benefited and displaced by the policy, together with their respective gains and losses.

5.3 Gains and losses of affirmative action

Our goal is to define how much a student would gain or lose if they were pulled into or pushed out of their preferred feasible degree as an increase in the affirmative action policy. We refer to μ as the matching function in the status quo, and to μ' as the counterfactual matching function when the affirmative action policy is increased. The conditional average treatment effect of increasing ω to ω' for individuals of type θ is:

$$\tau(\theta) = \mathbb{E}[Y_i(\mu') - Y_i(\mu) | \, \theta_i = \theta] \tag{1}$$

[25] In the extreme case of $\omega' = 0$, applications from targeted and non-targeted students are processed together and spots are allocated solely based on priority scores, and thus every student would face the same admission threshold irrespective of their affirmative action status.

To define the aggregate effect of increasing affirmative action in degree j, we integrate the conditional gains (or losses) of all individuals affected by the policy in the sense that their observed degree $D_i(\mu)$ is different from their counterfactual preferred option $D_i(\mu')$. To aggregate these treatment effects, we assume an equal weight for all individuals within similar affirmative action groups. Thus, the aggregate effects for each affirmative action group t are given by:

$$\Delta_t(\omega', \omega) \quad = \quad \sum_{i \in \mathcal{I}} \tau(\theta_i) \cdot \mathbb{1}\{t_i = t\} \tag{2}$$

We can use the status-quo and counterfactual admission thresholds, $c_{tj}(\mu)$, and $c_{tj}(\mu')$, to identify the individuals who gain or lose access to any given degree as a result of the policy. For the case of targeted students, individuals with $D_i(\mu') = j$ and priority score $s_{ij} \in [c_{1j}(\mu'), c_{1j}(\mu))$ will benefit and gain admission into degree j when the affirmative action policy is intensified.[26] Conversely, non-targeted individuals with priority scores $s_{ij} \in [c_{0j}(\mu), c_{0j}(\mu'))$, will be displaced from degree j by the regulation.

5.4 A simple example

Under the lenses of this framework, the aggregate consequences of the policy are given by how much pulled-in students gain, and how much pushed-out students lose in terms of a given outcome of interest. For instance, the policy maker could be interested in how much earnings benefited individuals gain, and how much earnings displaced individuals lose as a result of the policy. We show that these gains and losses are characterized by how admission scores, degree attendance, and the outside options affect the outcome of interest.

Based on the dependance between outcomes, degree attendance and priority scores, this stylized framework can accommodate a wide range of distributional consequences of affirmative action. In Figure 5, we describe one of these cases, in which the gains for targeted individuals outweigh the losses experienced by displaced individuals. For expositional clarity, assume that $\mathcal{J} = \{0, 1\}$, such that there is only one selective degree available, or an outside option. For simplicity, we also assume that all students have strict preferences for the selective degree ($j = 1$) over the outside option ($j = 0$). We also assume that the outcome of interest is earnings.

The solid lines show the average potential earnings for students attending the selective institution, conditional on their score and affirmative action type, $E[Y_{i1}|\, t_i, s_i]$. The dashed lines depict the average potential earnings from attending the outside option, $E[Y_{i0}|\, t_i, s_i]$. In the absence of affirmative action (i.e. $\omega = 0$), there is a single admission cutoff c', for all students. By focusing on

[26] Take the case of targeted individuals. The upper limit of integration in Equation (2) is given by the fact that degree j is always feasible students with score $s_{ij} \geq c_{tj}(\mu)$, thus not affecting their degree decision. The lower limit of integration comes from the fact that degree j is never feasible for students with scores $s_{ij} < c_{tj}(\mu)$. The logic for non-targeted students is the reverse.

students scoring around c', we can easily observe that implementing affirmative action is efficient for the marginal student.[27]

Now assume the planner implements an affirmative action policy such that targeted and non-targeted students face different cutoffs, represented by $c_{t=1}$ and $c_{t=0}$ respectively. Targeted students with scores $s_i \in [c_{t=1}, c']$ are now offered admission, while non-targeted students with scores $s_i \in [c', c_{t=0}]$ are displaced from the selective institution. Area A depicts the mean gain for benefited individuals, while the mean cost incurred by displaced individuals is represented by area B. In the figure, we observe that the consequences of affirmative action depend crucially on heterogenous returns for benefited and displaced individuals, and also on how this difference varies with the admission cutoff.

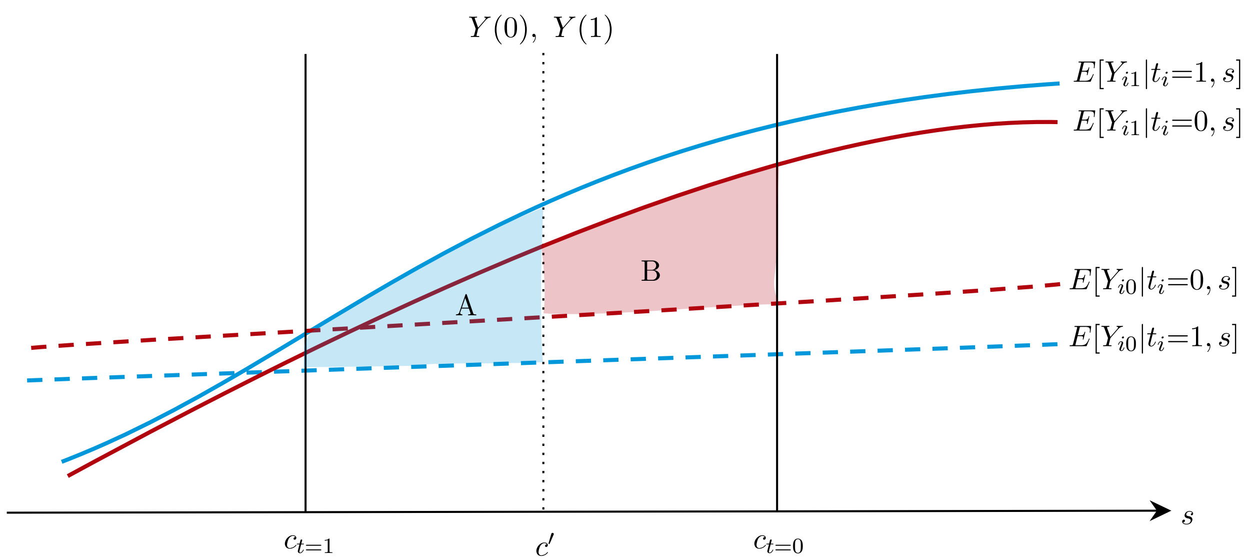


Figure 5: Conceptual Framework

Notes: This figure presents the distributional impacts of an affirmative action policy in a centralized mechanisms. Each of the lines denote the mean potential outcome for targeted (blue line) and non-targeted (yellow line) individuals. While the solid line present the expected outcome of attending the selective degree, the dashed line presents the expected outcome of attending the outside option.

6 Econometric model

In Section **??**, we showed that, in our model, affirmative action works via increasing the number of degrees available to targeted individuals at the cost of reducing the number of options available to non-targeted individuals. As a result of these changes in choice sets, individuals may change their preferred degrees, which may in turn impact students' realized outcomes. In this section, we show how to bring these objects to the data.

[27]First, the expected earnings of attending the selective institution are higher for targeted than for non-targeted students. Second, non-targeted students have a better outside option than targeted individuals. As a result, the returns of attending the selective institution are larger for targeted than for non-targeted students.

6.1 Object of interest

Our conceptual framework uses a potential outcomes framework to characterize the conditional average treatment effect of increasing the share of spots reserved for targeted students. In particular, our goal is to estimate the effect of increasing the affirmative action schedule from ω to ω' for individuals of type θ. Using the stability condition, we can rewrite Equation (1) as:

$$\begin{aligned}\tau(\theta) &= \mathbb{E}[Y_i(\mu') - Y_i(\mu) \,|\, \theta_i = \theta] \\ &= \mathbb{E}\left[\sum_j \mathbb{1}\{\mu'(\theta_i) = j\} \cdot Y_{ij} - \sum_j \mathbb{1}\{\mu(\theta_i) = j\} \cdot Y_{ij} \;\middle|\; \theta_i = \theta\right] \end{aligned} \tag{3}$$

where μ and μ' represent the matching function for each affirmative action schedule, and Y_{ij} denotes the potential value of some outcome of interest for individual i if they attend degree j.

Accordingly, estimating $\tau(\theta)$ requires estimating the matching functions for each of the affirmative action schedules, together with the potential outcomes. The estimation of each of these objects presents its specific challenges, which we address in turn. First, to estimate the matching function, we leverage the rules of the mechanism φ, together with its inputs (Θ, q, ω). The main difficulty emerges from the fact that, in practice, researchers never observe the full set of student types Θ, as individuals rarely exhaustively rank the full list of preferences $\succ_i$. As such, in Section 6.2, we introduce an empirical school choice model to estimate student preferences for degrees.

Second, we ought to estimate the potential outcomes for every student and degree combination Y_{ij}. One possible direction is to impose a linear projection of Y_{ij} on student and degree characteristics and estimate the conditional expectation of Y_{ij} using OLS. However, identifying these parameters is difficult because of the standard problem of selection into degrees. That is, the decision to attend a given degree is likely correlated with unobservables of the potential outcome, thus biasing the parameter estimates. For example, students who choose to attend the most selective degrees are commonly also the best prepared to perform well in these–be it due to relevant prior training, a comfort with the environment, useful socio-emotional traits, or other unobservables.

To overcome this problem, in Section 6.3, we follow Abdulkadiroglu et al. (2020) and construct selection-corrected estimates of the parameters using a control function approach. This approach jointly models the choice of degrees (selection equation) together with the potential outcomes (outcome equation). The identification strategy leverages the fact that choice sets are a discontinuous function of test-scores, while potential outcomes are affected by test-scores in a continuous manner.

6.2 School choice model

To recover student preferences, we summarize the observed choices in the centralized platform by fitting a random utility model. Specifically, student i's indirect utility for attending degree j is:

$$\begin{aligned} u_{ij} &= V_{ij} + \eta_{ij} \\ &= \delta_j^t + \gamma_j^t \cdot s_{ij} + \kappa^t \cdot d_{ij} + \eta_{ij} \end{aligned} \tag{4}$$

where V_{ij} captures the part of the utility that varies according to the observed characteristics of students and degrees, and η_{ij} captures unobserved tastes for degrees. We parametrize V_{ij} as a function of degree fixed effects δ_j^t, the student's test scores s_{ij}, and a dummy d_{ij} indicating whether the student lives in the same municipality where the degree is offered. We allow flexible heterogeneity in tastes by estimating preference models separately for each affirmative action type t. We model the unobserved taste η_{ij} as following an extreme value type I distribution, conditional on vectors s_i and d_i. The outside option aggregates all degrees offered by municipal, state, and private universities, or not enrolling at all. This option is available to everyone and has a deterministic utility V_{i0} normalized to zero.

Implicit in Equation (4) is an assumption that V_{ij} is independent of the matching $\tilde{\mu}$. As such, preference parameters $(\delta_j^t, \gamma_j^t, \kappa^t)$ can rationalize true utilities for degrees under any realization of the matching function. This assumption implies that preferences for degrees are independent of the observed allocation of students, ruling out preferences for peers. Although this assumption might sound restrictive, most existing empirical approaches abstract away from equilibrium sorting based on preferences for peers (Agarwal and Somaini, 2020).[28]

Let $\tilde{\mu}$ be a realized matching that we observe in the data. The mechanism's stability property implies a discrete choice model with observable and personalized choice sets $\Omega_i = \Omega_i(\tilde{\mu})$. Also, it implies that the degree to which the student is assigned is also their preferred feasible option ex-post, that is $D_i = D_i(\tilde{\mu}) = \tilde{\mu}(\theta_i)$. One of the main advantages of final stable allocations is that researchers can exploit revealed preference relations based only on assignment data. As such, this approach can be implemented in any centralized system using strict priorities with stable assignments. The stability property, as an ex-post optimality condition is not necessarily guaranteed when students do not have complete information about the admission cutoffs. In Appendix E, we show that in the Brazilian setting students have little uncertainty over the final admission cutoffs due to the iterative property of the mechanism.

Following Fack et al. (2019), we restrict the choice set Ω_i to all degrees available to student i based on their score and on the admission cutoff specific to their degree and affirmative action

[28] Allende (2021) is a notable exception.

type. The logit model implies that the probability that individual i selects degree j is given by:

$$Pr(D_i = j) = \frac{\exp V_{ij}}{\sum_{k \in \Omega_i} \exp V_{ik}} \tag{5}$$

We estimate $(\delta_j^t, \gamma_j^t, \kappa^t)$ for $t \in \{0, 1\}$ using a maximum likelihood estimator.

6.3 Potential outcome equation

Next, we switch our focus to constructing selection-corrected estimates of a potential outcome model using a control function approach. We follow a similar approach to that of Abdulkadiroglu et al. (2020), who link school choices to potential outcomes to estimate schools' value-added in New York City. We use our estimated parameters to predict the potential outcomes for attending a given counterfactual preferred degree.

We project the potential outcome of student i at counterfactual degree j on degree fixed effects and student and degree characteristics:

$$Y_{ij} = \alpha_j^t + X'_{ij}\beta_j^t + \varepsilon_{ij} \tag{6}$$

where α_j^t and β_j^t are population parameters for student group t, implying $\mathbb{E}[\varepsilon_{ij}] = \mathbb{E}[X_{ij}\varepsilon_{ij}] = 0$. The variable X_{ij} is a vector of observed covariates, including student test scores s_{ij}, and a dummy d_{ij} indicating whether the student lives in the municipality where the degree is offered. Our goal is to recover the parameters of the potential outcome equations defined above.

The mean outcome Y_i observed in the data for a given matching $\tilde{\mu}$ is given by:

$$\mathbb{E}[Y_{ij}|X_{ij}, D_i = j] \quad = \quad \alpha_j^t + X'_{ij}\beta_j^t + \mathbb{E}[\varepsilon_{ij}|X_{ij}, D_i = j] \tag{7}$$

The OLS estimation of this equation would likely yield biased parameters due to selection into degrees based on unobservable preferences. To recover unbiased estimates we would need to assume that $\mathbb{E}[\varepsilon_{ij}|X_{ij}, D_i = j] = 0$, thus implying that degree choices and potential outcomes are not correlated after accounting for student and degree observed characteristics.

To account for selection on unobservables, we link the outcome equation to the school choice model by conditioning Equation (6) on the vector of unobserved tastes $\eta_{i1}, ..., \eta_{iJ}$:

$$\mathbb{E}[Y_{ij}|X_{ij}, \eta_{i1}, ..., \eta_{iJ}] \quad = \quad \alpha_j^t + X'_{ij}\beta_j^t + \mathbb{E}[\varepsilon_{ij}|X_{ij}, \eta_{i1}, ..., \eta_{iJ}] \tag{8}$$

This model allows expected potential outcomes to vary across students with different preferences for degrees in a way that is not captured by students' observables. To estimate Equation (8) we use the multinomial logit selection model of Dubin and McFadden (1984), which imposes a linear relationship between potential outcomes and the unobserved logit errors. Imposing such a parametric approximation yields:

$$\mathbb{E}[Y_{ij}|X_{ij},\eta_{i1},...,\eta_{iJ}]=\alpha_j^t+X_{ij}'\beta_j^t+\sum_{k=0}^{J}\psi_k^t\cdot(\eta_{ik}-\mu_\eta)+\rho^t\cdot(\eta_{ij}-\mu_\eta) \tag{9}$$

where $\mu_\eta \equiv E[\eta_{ij}]$ is Euler's constant. As pointed out by Abdulkadiroglu et al. (2020), this parametric relationship allows for a wide range of selection on unobservables in the context of school choice. The parameter ψ_k captures the effect of the preference for degree k that is common across all potential outcomes. For example, students with high preferences for a given type of institution may have higher outcomes in all other degrees in a way that is not fully captured by student observables (e.g., social status). We refer to this term as selection on levels. The coefficient ρ represents the match effect of preferring degree j. This unobserved match component allows, for instance, for students to sort into degrees based on potential outcome gains. We refer to this type of selection as selection on gains (Roy, 1951).

By iterated expectations, the mean outcome observed in the data for a given assignment $\tilde{\mu}$ is:

$$\mathbb{E}[Y_i|X_{ij},D_i=j]=\alpha_j^t+X_{ij}'\beta_j^t+\sum_{k=0}^{J}\psi_k^t\lambda_{ik}+\rho^t\lambda_{ij} \tag{10}$$

where $\lambda_{ik} \equiv \lambda_k(X_{ij},D_i) \equiv \mathbb{E}[\eta_{ik}-\mu_\eta|X_{ij},D_i]$ is the expectation of the unobserved preference conditional on the student's characteristics, and degree of choice. These functions have a closed-formed solution and can be computed using the logit functional form. These objects serve as control functions to correct for selection on unobservables.

The main concern is that the equation above is not identified. We can further condition on Ω_i

$$\mathbb{E}[Y_i|X_{ij},D_i=j,\Omega_i] \quad = \quad \alpha_j^t+X_{ij}'\beta_j^t+\sum_{k=0}^{J}\psi_k^t\cdot\lambda_{ik}(\Omega_i)+\rho^t\cdot\lambda_{ij}(\Omega_i) \tag{11}$$

where $\lambda_k(X_{ij},\Omega_i,D_i) \equiv \mathbb{E}[\eta_{ik}-\mu_\eta|X_{ij},\Omega_i,D_i]$ is the expectation of the unobserved preference conditional on the student's characteristics, choice set, and degree of choice. These functions have a closed-formed solution and can be computed using the logit functional form. These objects serve as control functions to correct for selection on unobservables.

To provide intuition, suppose there are two available degree choices, A and B, and two individuals, 1 and 2. Individuals are identical in observables and have to select a degree from randomly assigned choice sets. Suppose that individual 1 can only choose option A, while individual 2 can pick between options A and B. If both individuals choose degree A, we can use a revealed preferences argument to learn that individual 2 has an unobserved taste for option A that is higher in expectation than that revealed by individual 1. The selection parameters capture whether this expected difference in unobserved preferences is relevant for the potential outcome. The estimation of equation (11) proceeds in two steps. First, we compute $\hat{\lambda}_k(\cdot)$ using the estimated parameters from the choice Equation (4). Next, we plug $\hat{\lambda}_k(\cdot)$ into Equation (11), and estimate parameters $(\alpha_j, \beta_j, \psi_j, \rho)$ using OLS.

The identification of Equation (11) relies on the assumption that choice sets $\Omega_i(\tilde{\mu})$ are exogenous to unobserved tastes $\eta_{i1}, ..., \eta_{iJ}$ after conditioning on X_{ij}. This identification assumption is implied by the model, as the only student inputs of the personalized choice set function $\Omega_i(\tilde{\mu}) = \{j \in \mathcal{J} \mid s_{ij} \geq c_{t_i j}(\tilde{\mu})\}$ are student test-scores s_{ij}, and affirmative action types t_i. Implicit in this assumption is the fact that students take admission cutoffs c_{tj} as given and cannot manipulate them with their own specific application behavior.

In Appendix G we provide a formal identification proof. The intuition behind the identification argument is similar to that of a standard regression discontinuity design. Take two individuals with almost identical observable characteristics and unobservable tastes for degrees. While individuals who barely crossed the admission cutoff of degree j will be eligible to attend the degree, individuals just below the cutoff will not be admitted to the degree. Thus, differences in potential outcomes can only be explained by discontinuous degree availability (and attainment). The assumption on linearity allows us to extrapolate the treatment effects away from the admission discontinuities using the linear functional form of the potential outcomes described in Equation (6).

6.4 Test-scores, Ability and Exogenous Shifters [SKIP THIS SUBSECTION]

Our exogenous variation exploits random differences in scores in the written essay component of the exam stemming from students being assigned to graders of varying strictness. Because the allocation of graders is completely exogenous to the student, we can construct a leave-one-out measure of grader leniency for every student. We argue that this measure is uncorrelated to the student but directly affects their score.

The ENEM exam consists of four multiple choice tests -on Math, Language, History, and Science- and a written essay.[29] Each essay is marked by two randomly assigned graders. Each year, over 5,000 individuals are involved in the grading process, each of whom is assigned to more than 500 exams. The only restriction imposed in the assignment process is that graders must be from different states to their assigned students. The essay is marked based on five different competencies, each of them with a score ranging from 0 to 200. The final grade is the simple average of the total points given by each of the graders. If the score difference across graders is large for any exam, that exam is graded by a third individual, and the final grade is then the average of the two closest scores.[30]

[29] The essay usually has a 20% weight for most degrees. This holds true across different fields of study.

[30] A third grader intervenes in two situations: (i) If the difference between scores across graders is larger than 100 points, (ii) if the difference in one or more of the competencies is larger than 80 points. The grading rubric is available from the following link:
http://download.inep.gov.br/educacao_basica/enem/guia_participante/2018/manual_de_redacao_do_enem_2018.pdf

Our exogenous variation exploits random differences in scores in the written essay component of the exam stemming from students being assigned to graders of varying strictness. Because the allocation of graders is completely exogenous to the student, we can construct a leave-one-out measure of grader leniency for every student. We argue that this measure is uncorrelated to the student but directly affects their score. Specifically, we define our instrument z_i as:

$$z_i \;=\; \frac{1}{2}\sum_{j=1}^{2}\left(\frac{1}{(N_{g(i,j)}-1)}\sum_{k\in(N_{g(i,j)}/i)} e_{kg(i,j)}\right)$$

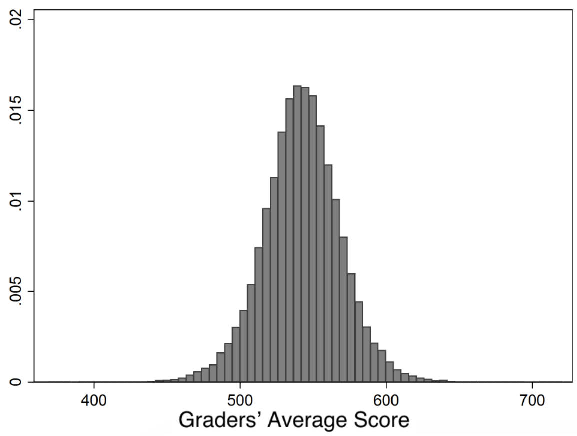
where $g(i,j)$ denotes jth grader assigned to individual i, e_{kg} represents the score given by grader g to student k, and N_g denotes the set of exams received by grader g. In words, for a given student i, the leave-one-out (jackknife) instrument is the average score their two graders gave to all other students they graded. To construct the leniency measure, we only use the first two assigned graders.

Existence of First Stage: In order for the instrument to be valid, we need a first-stage relationship between leniency and student scores. Panel (a) in Figure 6 reports a remarkably wide dispersion in graders' average scores, with a range of 150 points (or about 2 test score standard deviations). Given the large number of essays marked by each grader, in the absence of any leniency differences, the leave-one-out mean instrument for each grader should be concentrated around 550 points, the mean score for the exam. Panel (b) reports the first-stage relationship between grader leniency and student scores. We observe a linear and strong relationship between student scores and the leave-one-out instrument.

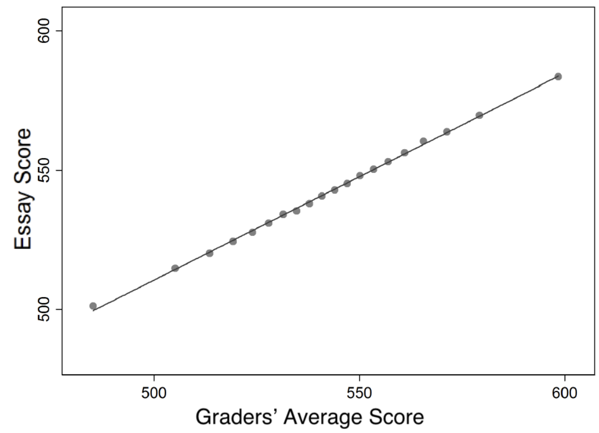
Reduced-Form on Educational Outcomes: To show that our instrument shifts students scores in a meaningful way, in Figure 7 we present the reduced-form relationship between our instrument and educational outcomes. Panel (a) suggests that a student that got a grader with a leniency score in the top 5% is 2.7 percentage points more likely to attend college than one who got a grader in the bottom 5%. In Panel (b) we show that grader leniency is not only impacting the extensive margin decision of attending a higher education institution but also the quality of the degree one attends, as measured by the average peer score on the ENEM.

Exclusion Restriction: Given that the allocation of graders is random, our leniency measure should be uncorrelated with student ability and any other important confounders. The exclusion restriction will likely hold as we expect grader leniency only to impact students by shifting their essay scores. To test the randomness of the assignment we provide one balance test and one placebo test. Panel (a) in Figure 8 shows no correlation between grader leniency and student performance on the multiple choice part of the exam. Likewise, Panel (b) reports a null relationship between our instrument and previous college attendance. Consistent with a satisfactory randomized assignment, we find no correlation between the instrument and any relevant student characteristics.

Figure 6: First Stage



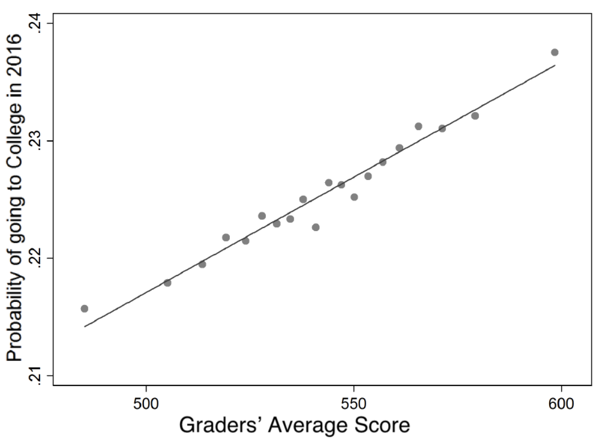
(a) Histogram



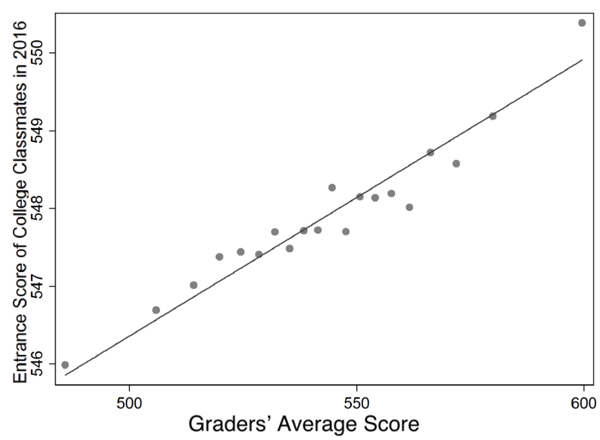
(b) First Stage on Essay Score

Notes: The x-axis measures leniency, and is the average of the leave-out mean instrument for each grader. Panel (a) reports the histograms of graders leniency measure, and Panel (b) presents the first-stage with respect to students score in the essay section.

Figure 7: Reduced Form



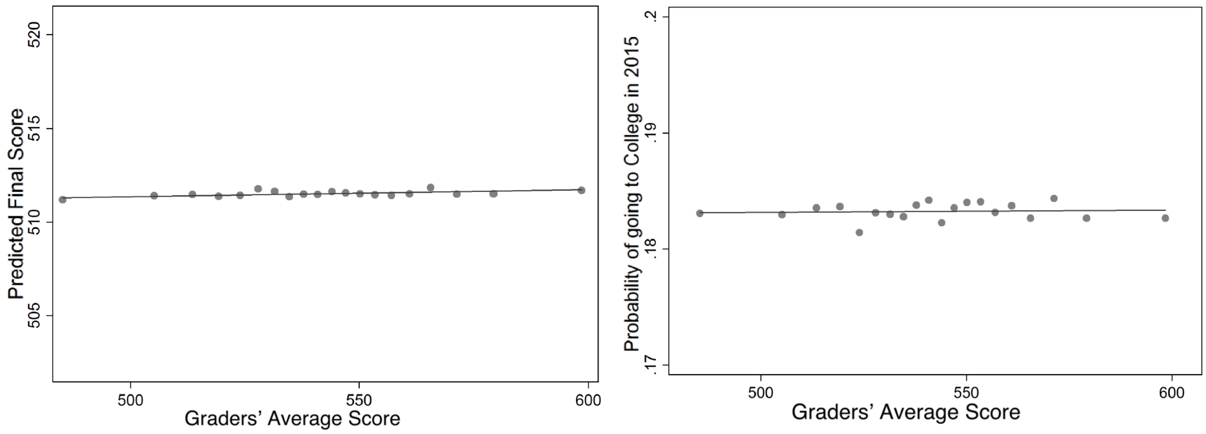
(a) On College Attendance in $t+1$



(b) On Peer Quality

Notes: The x-axis measures leniency, and is the average of the leave-out mean instrument for each grader. Panel (a) shows the reduced-form between grader leniency and college attendance during the year after taking the exam. Panel (b) reports the relationship between graders leniency and peer quality of the chosen.

Figure 8: Placebo Test



(a) On previous college attainment

(b) On predicted score

Notes: The x-axis measures leniency, and is the average of the leave-out mean instrument for each grader. Panel (a) shows no correlation between grader leniency and students performance on the multiple choice part of the exam. Panel (b) reports relationship between graders leniency and previous college attendance.

7 Parameter estimates

In this section, we present the parameter estimates. We first discuss the preference parameters that arise from the estimation of the school choice model. We then discuss the estimates from the potential outcome model.

7.1 Estimating sample

To estimate the model, we focus on the round of applications to higher education institutions, which involved taking the ENEM in 2015 and using the SISU platform in 2016. We restrict the sample to every applicant submitting an application to any institution in the state of Minas Gerais. We do this exclusively for computational power constraints. In total, we have 12 institutions offering 320 degrees (including the outside option). Our sample size consists of 200,900 students, of which 56% are targeted and 44% are non-targeted. We expect to expand to others states in the next version of the draft.

7.2 School choice model

In Panel A of Table 2 we present the parameters estimates from the school choice model. We find positive coefficients for the location dummies, indicating that students have strong preferences for studying close to where they reside. The location dummy is 3.32 for targeted individuals, which implies that [add interpretation]. This coefficient is slightly smaller for non-targeted students, with a magnitude of 3.23.

Table 2: Preference Parameters

	Mean	
Targeted status	AA	NA
Panel A: School Choice Model		
Degree fixed effect (δ_j)	10.46	31.20
	-	-
Score parameter (γ_j)	-0.025	-0.053
	-	-
Location dummy (κ)	3.321	3.232
	$(\cdot)$	$(\cdot)$
Panel B: Potential Outcomes Model		
Degree fixed effect (α_j)	-2,103	-1,164
	-	-
Score parameter (β_j)	7.35	4.94
	-	-
Selection in levels (ψ_j)	-70.25	-319.31
	-	-
Selection in gains (ρ)	69.45	342.29
	(11.83)	(25.41)
Share of students (%)	56	44
Number of students	200,900	
Number of degrees	320	
Number of institutions	12	

Notes: This table summarizes the parameters estimates. AA denote the targeted students, while NA refers to non-targeted students. Panel A presents coefficients from the school choice model. Panel B display parameters from the potential outcomes model.

Parameters related to degree fixed effects (δ_j) or those associated with student test scores (γ_j) are very high-dimensional, as we have one different estimated coefficient per degree and affirmative action status tuple. To visualize these parameters, we estimate the average valuation for a given degree $\bar{V}_j^t$. To make estimates comparable across targeted and non-targeted groups, we fix the test score variable to equal the admission cutoff score for open spots for that degree, c_{0j}, and also fix the location dummy equal to 0 (i.e., the student lives outside the municipality where the degree is offered). Specifically,

$$\bar{V}_j^t \quad = \quad \hat{\delta}_j^t + \hat{\gamma}_j^t c_{0j}$$

The value of $\bar{V}_j^t$ indicates the valuation of a given degree relative to the outside option. The larger $\bar{V}_j^t$, the more likely it is that the student would choose degree j over the outside option if only these two degrees were offered to them.

In Figure 10(a), we show how these valuations vary by degree selectivity and affirmative action status. First, we note that, relative to the outside option, degrees have a low valuation. This is not surprising as the outside option pools a collection of different alternatives (e.g., working, attending a private institution, waiting another year to go to college) and its value is the maximum value of all these alternatives. Second, we find that targeted students put a higher value on the outside option relative to attending a federal degree when compared to their non-targeted counterpart. If targeted students are in greater need of working, for example, they will be more likely to choose the outside option and to enter the labor market, rather than enrolling in any given degree. Third, there is a very strong correlation between degree selectivity and average degree valuation. Fourth, the valuation of degrees is non-linear and substantially higher for very selective degrees. These degrees mostly consist of medicine and engineering-related programs.

We use these parameters to assess the in-sample model fit of the model by comparing the admission thresholds predicted by the model to those observed in the data. The correlation coefficient of the slope is virtually 1 for both open and reserved seats. We discuss the construction of the predicted admission thresholds and present the model fit figures in Appendix D.

7.3 Potential outcome equation

The potential outcome equation establishes a structural relationship between predicted income, degree attendance, and student characteristics. We define the value added as the gains of attending a given degree relative to the outside option. To calculate the value added, we use the population parameters α_j and β_j, and fix the test score variable to equal the admission cutoff score for open spots c_{0j}, and also fix the location dummy equal to 0.

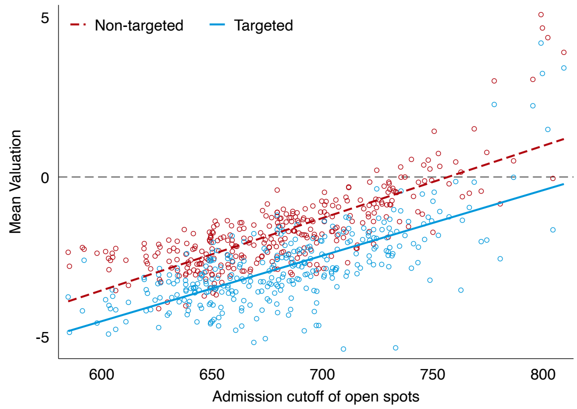
$$\mathrm{VA}_j = (\hat{\alpha}_j^t - \hat{\alpha}_0^t) + (\hat{\beta}_j^t - \hat{\beta}_0^t) \cdot c_{0j}$$

In Figure 10(c), we plot the value-added of each degree against its selectivity level. We find that more selective degrees offer higher value added. In contrast to the degree valuations from Figure 10(a), we find than the average degree offers higher value than the outside option (i.e., positive value added). Moreover, targeted students obtain higher value added than their non-targeted counterpart. This is likely driven by non-targeted students having outside options with high returns, such as attending similar programs in private universities.

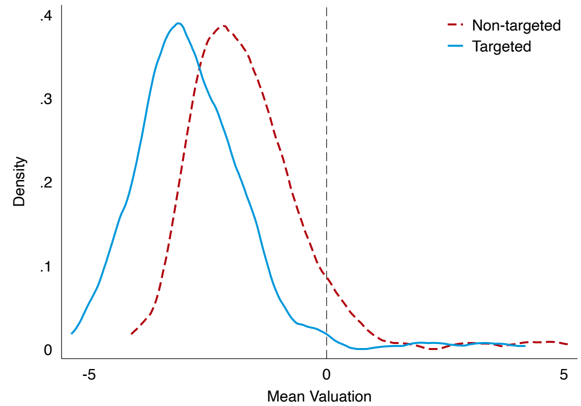
8 Counterfactual estimation

In this section, we use the parameter estimates from our model to compute student allocations, as well as their respective potential outcomes under different affirmative action schedules. We then compare the overall predicted income gains and losses for targeted and non-targeted individuals, and for the system as a whole.

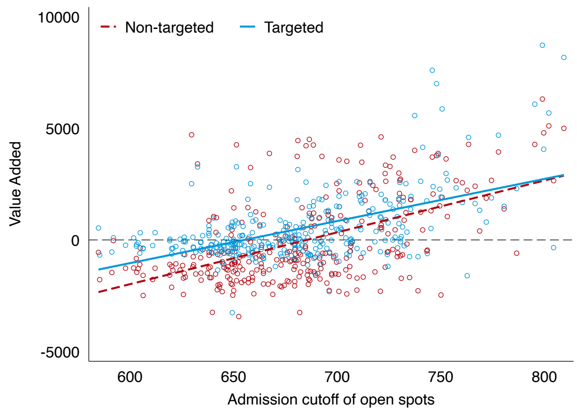
Figure 9: Mean valuations and value added



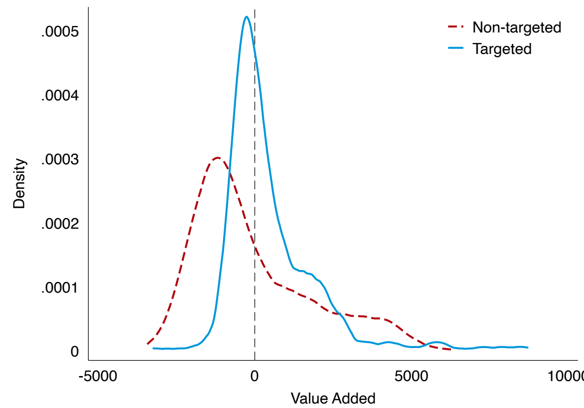
(a) Mean valuation and degree selectivity



(b) Mean valuation density



(c) Value added and degree selectivity



(d) Mean valuation density

Notes: This figure present the mean valuation and value added of degrees based on the parameter estimates. Panel (a) shows the relationship between degree selectivity and students' mean valuation for them. Panel (b) shows the relationship between degree selectivity and degree's value added. In Panels (a) and (c), the x-axis denotes degree selectivity as measured by the admission cutoff of open seats c_{0j}. Panel (b) and (d) show the distribution of degree valuation and value added for each affirmative action type.

8.1 Estimating counterfactual allocations

The goal is to simulate student allocations for any given affirmative action schedule ω by leveraging the rules of the mechanism. We start by recovering the inputs of the matching function $\varphi(\Theta, q, \omega) = \mu$, namely student types, Θ, and degree capacities, q.

The first step is to compute the set of student types $\Theta = \bigcup_{i \in \mathcal{I}} \theta_i$, where $\theta_i = (\succ_i, s_i, t_i)$. We recover preferences, $\hat{\succ}_i$, by evaluating indirect utilities from the school choice model introduced in Section 6.2 using preference parameter estimates from Section **??**. To recover the remaining inputs to Θ, we assume no behavioral responses to the regulation. This allows us to recover inputs directly from the data. Specifically, we assume that the composition of applicants, $\mathcal{I}$, priority scores, s_{ij}, and affirmative action status, t_i, are fixed and invariant to changes in the affirmative action schedule. The main concern related to ruling out behavioral responses is whether the policy effects are well predicted by ignoring them. In Appendix F, we use the empirical design introduced in Section **??** to test whether these assumptions are supported by the data. We discuss each of them in turn.

First, in terms of compositional responses, we find no evidence of the affirmative action regulation inducing changes in the number of ENEM takers or the share of targeted students who take the test.

Second, we examine behavioral responses in terms of student test scores. Based on theory, affirmative action can have heterogeneous effects on exam preparation effort depending on a student's position in the ability distribution and their affirmative action status (Bodoh-Creed and Hickman, 2018).[31] In Appendix F.1, we show that we find no evidence of changes in preparation effort for either group at any point in the test score distribution. This is in line with the results of Estevan et al. (2018), who examine an affirmative action policy in a large state university in Brazil and find no evidence of behavioral reactions concerning exam preparation effort.[32]

Third, we assess whether the regulation induces strategic responses from individuals gaming their affirmative action eligibility. As explained in Section **??**, the AA regulation targeted students who had attended public high school from grades 10 to 12. In Appendix F.2, we evaluate the impact of the regulation on high school choice. Similar to Mello (2020), we find supportive evidence of students switching from private to public schools between grades 9 and 10 to obtain AA eligibility. Nonetheless, we neglect this margin of response from the analysis for two reasons. The first one is conceptual. We are interested in learning the implications of affirmative action regulation which targets fixed demographics, such as race or socioeconomic background. As such, we do not deem it important to include eligibility manipulation as a potential margin of response in the model. The second reason is empirical; ignoring this margin of response is very unlikely to bias our estimates. We focus our analysis on the 2016 SISU sample, where roughly 70% of all applicants had graduated from high school before 2015. Consequently, given the timing of the regulation, most applicants did not have any strategic incentives to switch schools.[33] Our reduced form estimates, imply that only 0.05% of all 2016 SISU applicants had switched schools in response to the AA regulation.

[31] Following Cotton et al. (2020)'s comparative static predictions, targeted individuals in the upper part of the ability distribution may reduce their effort profiting from lower admission test scores in their preferred degrees. On the other hand, the AA regulation may place within reach degrees which would otherwise be unattainable for targeted individuals, thus incentivizing them to increase their preparation effort. The predictions are reversed for the case of non-targeted students. For empirical evidence in other contexts, see Akhtari et al. (2020) and Cotton et al. (2021).

[32] These null effects on test scores are consistent with effort cost functions being sufficiently convex. The highly competitive nature of the exam might have already induced students to be at a sufficiently convex point of the effort cost function where the elasticity of effort to changes in admission thresholds is low.

[33] The regulation was announced in 2012 and implemented in 2013. Thus, students attending grades 10th or higher in private schools in 2012 (i.e. graduating before 2015) could have not switched to a public school and become eligible for the reserved spots.

Moreover, in Appendix F.2 we show that most strategic school-switchers come from low-performing private schools, suggesting that these movers are very unlikely to affect the final allocations of the mechanism.

The second step to recovering the inputs to the matching function, is to recover degree capacities, q. This variable, as opposed to the inputs of Θ, is a policy choice that is decided together with the affirmative action schedule. Since we are interested in learning about the affirmative action consequences we keep this variable fixed and recover it from the observed data. After recovering, $\hat{\Theta}$ and q, we simulate the matching function as $\varphi(\hat{\Theta}, q, \omega) = \hat{\mu}$.

8.2 Estimating counterfactual outcomes

The next step after estimating the matching function, is to compute the realized outcomes associated with such assignment. We define our object of interest for individual of type θ_i as $\mathbb{E}[Y_i(\hat{\mu}(\theta_i)) \mid \theta_i]$, where $\hat{\mu}$ denotes the matching function estimated above. Using Equations (3), (8), and (9) we parametrize this object as:

$$\begin{aligned}
\mathbb{E}[Y_i(\hat{\mu}(\theta_i)) \mid \theta_i] &= \sum_j \mathbb{1}\{\hat{\mu}(\theta_i) = j\} \cdot \mathbb{E}[Y_{ij} \mid \theta_i] \\
&= \sum_j \mathbb{1}\{\hat{\mu}(\theta_i) = j\} \cdot \big(\alpha_j + X_{ij}'\beta_j + \mathbb{E}[\varepsilon_{ij} \mid \theta_i]\big) \\
&= \sum_j \mathbb{1}\{\hat{\mu}(\theta_i) = j\} \cdot \left(\alpha_j + X_{ij}'\beta_j + \sum_{k=0}^{J} \psi_k \cdot (\eta_{ik} - \mu_\eta) + \rho \cdot (\eta_{ij} - \mu_\eta)\right) \quad (12)
\end{aligned}$$

This parametrization assumes that potential outcomes are invariant to the affirmative action regulation. There are two implications that we deem important to discuss. The first one is that this assumption rules out, for instance, peer effects in the production function of degrees, as well as changes in the value of Y_{ij} coming from stigmatization or reduction of the signaling value of degrees as a result of the AA regulation. Second, it implies that the value of the outside option is fixed. This assumption would be violated if non-targeted students displaced from public institutions were to crowd out other students from private institutions. This effect could potentially create a crowding-out cascade extending throughout the whole system. However, in Appendix H we show that most private institutions, including those comparable to federal universities, are far from being capacity constrained. Thus the crowding-out concerns are of second order in our specific setting.

We use Equation 12 together with the selection, and selection-corrected potential outcome parameter estimates, to simulate student allocations for a given schedule ω. In Appendix I we describe the simulation procedure in detail.

8.3 Counterfactual admission thresholds

Next, we perform counterfactual exercises where we vary the share of reserved seats ω from 0 to 100%. For each affirmative action schedule ω, we compute students' allocations and estimate their expected potential outcome in terms of predicted income as described in Section 3.4. We proceed by computing the admission thresholds associated with counterfactual allocations.

For expositional purposes, in Figure 10 we show the distribution of admission cutoff scores across degrees for reserved and open spots under two different counterfactual scenarios. The first one is a counterfactual similar to the current policy that reserves 50% of the seats to targeted students ($\omega = 0.5$). The second one corresponds to a laissez-faire situation in which there is no affirmative action at all ($\omega = 0$). Under this scenario, all students compete over the same spots and thus face the same admission thresholds.

We find that under the affirmative action counterfactual, where $\omega = 0.5$, cutoffs for open spots students are substantially higher than those reserved for targeted ones. This implies that targeted students can get admitted into selective degrees with much lower scores. In the absence of affirmative action, when $\omega = 0$, the distribution of cutoff scores (dashed black line) becomes uniform across affirmative action types and closer to the distribution of cutoff scores for open seats in the presence of affirmative action. In Appendix J, we use quantile-quantile plots to show how admission thresholds change for open and reserved spots under different affirmative action schedules ranging from 0 to 100%. The overall results suggest that, by removing the affirmative action program, admission into selective degrees becomes much harder for targeted students but not substantially easier for non-targeted ones (see Figures J.2 and J.3).

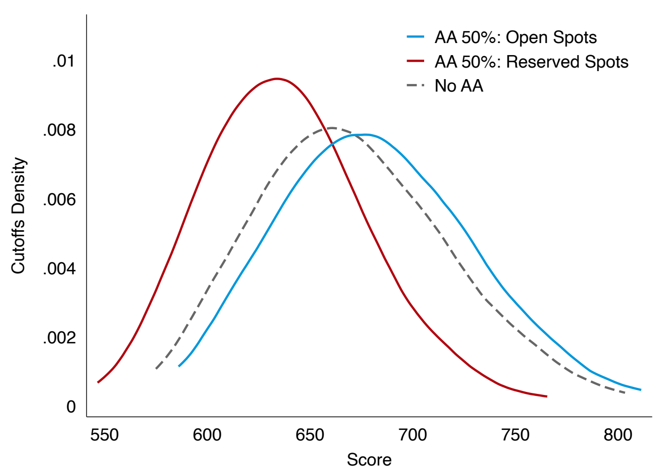


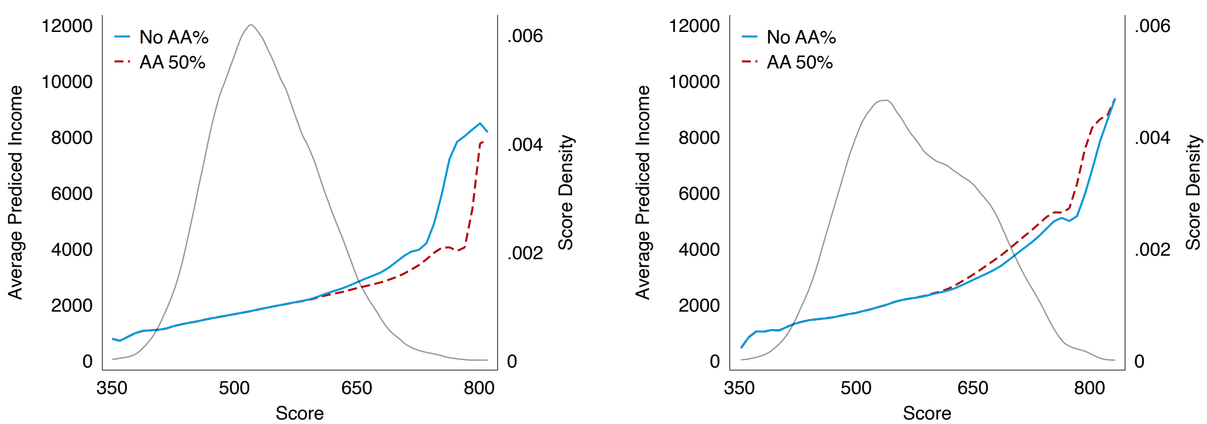
Figure 10: Distribution of cutoff scores

Notes: This figure shows the distribution of equilibrium cutoffs for open and reserved spots under different affirmative action schedules. The blue line shows the distribution of admission cutoffs for open spots when $\omega = 0.5$. The red line shows the distribution of admission cutoffs for reserved spots when $\omega = 0.5$. The dashed black line shows the distribution of admission cutoffs faced by all students when $\omega = 0$.

8.4 Counterfactual outcomes

In Figure 11 we show the expected predicted income for targeted and non-targeted students under each of the counterfactuals. The red dashed line denote a counterfactual scenario without AA, while the blue solid line indicates a counterfactual with 50% reserved seats. The grey line in the background (and measured by the right-hand side axis) shows the distribution of students over ENEM scores. In Figure 11(a) we show that the affirmative action program induces large gains on targeted individuals, especially on those students with high scores that now can access more selective degrees with higher value-added. As expected, targeted individuals at the very top of the score distribution are not affected by the AA regulation as all degrees are within reach even in the absence of the regulation.

These gains for targeted students, however, come at the cost of displacing non-targeted students from those very selective degrees. In Figure 11(b) we show that, even though non-targeted students are worse off under the affirmative action policy, their losses are small relative to the gains of targeted individuals. This is mostly explained by the fact that non-targeted individuals have lower chances of ending up in the outside option (see Figure 10(a)) and that their outside option has a relatively higher value-added (see Figure 10(c)).



(a) Targeted students

(b) Non-targeted students

Figure 11: Students' income with and without affirmative action

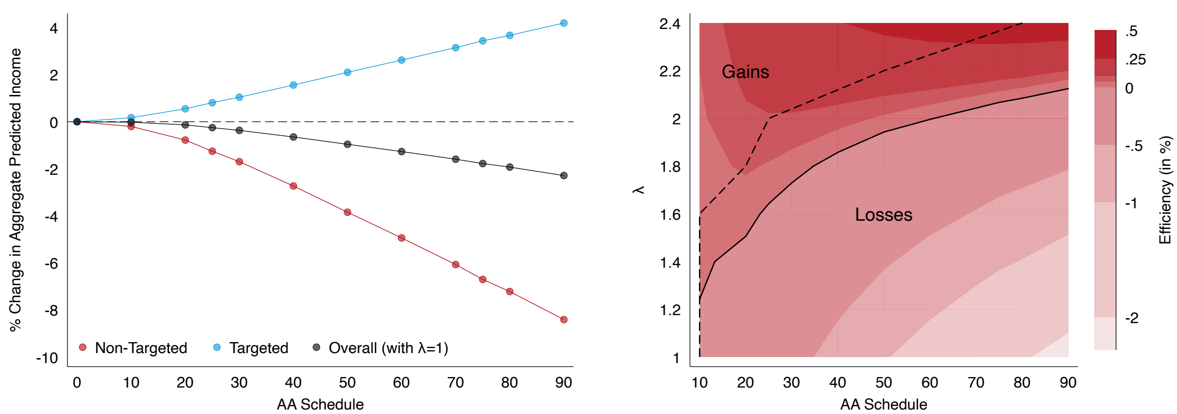
Notes: This figure shows the expected outcome for targeted and non-targeted students with and without affirmative action across the score distribution. The blue line represents the expected outcome when $\omega = 0$, and the red line represents the outcome when $\omega = 0.5$.

Next, we estimate the aggregate effects of AA on predicted income. Let $\Delta_t(\omega) = \Delta_t(\omega, 0)$ denote the aggregate gains for affirmative action group t of moving from no AA to an ω AA schedule, as indicated by Equation (2). The overall aggregate gains over targeted and non-targeted individuals are defined as:

$$\Delta(\omega, \lambda) \quad = \quad \Delta_0(\omega) + \lambda \Delta_1(\omega) \tag{13}$$

where λ denote the welfare weights capturing society's concerns for fairness with respect to the targeted group.

In Figure 12(a) we present the group-specific, as well as the overall aggregate gains in terms of predicted income. We normalize these gains in terms of the aggregated predicted income for affirmative action group t when $\omega = 0$. The blue and red lines denote the gains and losses for targeted and non-targeted individuals, respectively. The grey line denote the normalized overall aggregate change using equal welfare weights across groups, that is $\lambda = 1$. We find that, at the current affirmative action schedule of $\omega = 0.5$, the average targeted individual sees an increase of 2.1% in their predicted income, while the average non-targeted students faces a drop of 3.8%. The predicted income of the average student across affirmative action groups reduces by 0.96%. These results are consistent with those presented in Figure 11. Although the gains for targeted students are greater than the losses experienced by non-targeted individuals, these gains are added over fewer targeted individuals and thus do not compensate for the overall losses of non-targeted individuals.



(a) Predicted income

(b) By welfare weights

Figure 12: Predicted income under different affirmative action schedules

Notes: This figure shows the overall gains and losses of affirmative action in terms of predicted income. In Panel (a), we normalize the outcome with respect to the aggregate predicted income in the absence of AA. We perform a separate normalization for each of the groups: targeted, non-targeted and overall. The overall label denotes the outcome when individuals across both groups are given equal weights. Panel (b) shows a heat map of the overall efficiency gains when we vary the affirmative action schedule and the welfare parameter λ. We normalize the outcome for a given λ, relative to a counterfactual scenario without AA. The solid black line denotes the values of λ and ω, such that there are zero efficiency gains. The dashed black line indicates the values of λ and ω, such that the efficiency gains are maximized.

To take a stance on the efficiency of the policy, in Figure 12(b) we present the overall aggregate gains, $\Delta(\omega, \lambda)$, under different affirmative action schedules for different values of λ an. For a given λ, we normalize the efficiency gains relative to the overall aggregate predicted income in the absence of AA. The darker areas denote higher predicted income. The solid black line denotes the values of λ and ω such that $\Delta(\omega, \lambda) = 0$. The area to the left of the black solid line are net positive efficiency gains, while the area to the right reflects net losses. The black dashed line indicates the AA schedule ω that is associated to highest $\Delta(\omega, \lambda)$ for a given λ. We find that a welfare weight of $\lambda = 2.2$ rationalizes the current policy of $\omega = 0.5$ as the optimal AA schedule.

9 Discussion and Conclusion

9.1 Discussion

i) what parameters are driving these results,

ii) how these results connect to the initial motivation of the paper

9.2 Conclusion

In this paper, we study the distributional consequences of affirmative action policies in centralized admission systems. We develop and estimate a model where that link preferences over degrees together with the potential outcomes from attending each of them. We find that the affirmative action policy increases enrollment of targeted students into more selective degrees. In terms of outcomes, we focus on the impact of affirmative action policies on academic progress and the expected income of attending selective degrees. Our results suggest that these policies create large benefits for targeted students while imposing a smaller cost to non-targeted individuals. However, the effects of the policy cascade through the system affecting a larger number of non-targeted individuals than the number of targeted students that it benefits. Taken altogether we find that the efficiency impacts on the system largely depends on the welfare weight given to targeted students.

This paper focuses on the first order welfare trade-off between targeted and non-targeted individuals. However, these type of policies can also be affect other margins of the educational market not considered in this paper. By increasing diversity, universities can affect student outcomes through peer effects and influencing inter-group attitudes. A more diverse composition of the student body may change the production function of degrees, by affecting student's academic outcomes, social behavior and preferences. Understanding how these different margins interact with the direct distributional effects of affirmative action policy is key to understand its overall role in shaping a higher education sector that fosters social mobility and promotes healthy democracies.

References

Abdulkadiroglu, A. (2005). College admissions with affirmative action. *International Journal of Game Theory 109*(4), 535–549.

Abdulkadiroglu, A., J. Angrist, Y. Narita, and P. Pathak (2017). Research design meets market design: Using centralized assignment for impact evaluation. *Econometrica 85*(5), 1373–1432.

Abdulkadiroglu, A., P. Pathak, J. Schellenberg, and C. R. Walters (2020). Do parents value school effectiveness? Forthcoming, American Economic Review.

Agarwal, N. and P. Somaini (2018). Demand analysis using strategic reports: An application to a school choice mechanism. *Econometrica 86*(2), 391–444.

Agarwal, N. and P. Somaini (2020). Revealed preference analysis of school choice models. *Annual Review of Economics 12*(1), 471–501.

Akhtari, M., N. Bau, and L. Jean-William (2020). Affirmative action and pre-college human capital. Technical report, Working Paper.

Allende, C. (2021). Competition under social interactions and the design of education policies. Working paper, Princeton University.

Alon, S. (2015). *Race, Class, and Affirmative Action*. Russell Sage Foundation.

Alon, S. and O. Malamud (2014). The impact of israel's class-based affirmative action policy on admission and academic outcomes. *Economics of Education Review 40*, 123 – 139.

Arcidiacono, P. (2005). Affirmative action in higher education: How do admission and financial aid rules affect future earnings? *Econometrica 73*(5), 1477–1524.

Arcidiacono, P. and M. Lovenheim (2016). Affirmative action and the quality-fit trade-off. *Journal of Economic Literature 54*(1), 3–51.

Aygun, O. and I. Bo (2021). College admission with multidimensional privileges: The brazilian affirmative action case. Forthcoming, American Economic Journal: Microeconomics.

Azevedo, E. M. and J. D. Leshno (2016). A supply and demand framework for two-sided matching markets. *Journal of Political Economy 124*(5), 1235–1268.

Bagde, S., D. Epple, and L. Taylor (2016). Does affirmative action work? caste, gender, college quality, and academic success in india. *American Economic Review 106*(6), 1495–1521.

Bo, I. and R. Hakimov (2020). Iterative versus standard deferred acceptance: Experimental evidence. Forthcoming, Economic Journal.

Bodoh-Creed, A. and B. Hickman (2019). Pre-college human capital investment and affirmative action: A structural policy analysis of u.s. college admissions. Working paper, Haas School of Business, University of California, Berkeley.

Bodoh-Creed, A. and B. R. Hickman (2018). College assignment as a large contest. *Journal of Economic Theory 175*, 88–126.

Card, D. and A. B. Krueger (2005). Would the elimination of affirmative action affect highly qualified minority applicants? evidence from california and texas. *ILR Review 58*(3), 416–434.

Chetty, R., J. N. Friedman, E. Saez, N. Turner, and D. Yagan (2020, 02). Income Segregation and Intergenerational Mobility Across Colleges in the United States*. *The Quarterly Journal of Economics 135*(3), 1567–1633.

Cotton, C., B. R. Hickman, and J. P. Price (2020). Affirmative action,shifting competition, and human capital accumulation: A comparative static analysis of investment contests. Technical report, Working Paper.

Cotton, C., B. R. Hickman, and J. P. Price (2021). Affirmative action and human capital investment: Evidence from a randomized field experiment. Forthcoming, Journal of Labor Economics.

Dobbie, W., J. Golding, and C. Yang (2018). The effects of pre-trial detention on conviction, future crime, and employment: Evidence from randomly assigned. *American Economic Review 108*(2), 2 01–240.

Doyle, J. (2008). Child protection and adult crime. *Journal of Political Economy 116*(4), 746–770.

Dubin, J. A. and D. L. McFadden (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica 52*(2), 345–362.

Duflo, E. and A. V. Banerjee (2019). Good economics for hard times. In *Schooling, Experience, and Earnings*. Public Affairs.

Dur, U., S. D. Kominers, P. A. Pathak, and T. Sonmez (2018). Reserve design: Unintended consequences and the demise of boston's walk zones. *Journal of Political Economy 126*(6), 2457–2479.

Dur, U., P. A. Pathak, and T. Sonmez (2020). Explicit vs. statistical targeting in affirmative action: Theory and evidence from chicago's exam schools. Forthcoming, Journal of Economic Theory.

Epple, D., R. Romano, and H. Sieg (2008). Diversity and affirmative action in higher education. *Journal of Public Economic Theory 10*(4), 475–501.

Estevan, F., T. Gall, and L.-P. Morin (2018, 05). Redistribution Without Distortion: Evidence from an Affirmative Action Programme at a Large Brazilian University. *The Economic Journal 129*(619), 1182–1220.

Fack, G., J. Grenet, and Y. He (2019, April). Beyond truth-telling: Preference estimation with centralized school choice and college admissions. *American Economic Review 109*(4), 1486–1529.

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

Gale, D. E. and L. S. Shapley (1962). College admissions and the stability of marriage. *American Mathematical Monthly 69*(1), 9–15.

Howell, J. S. (2010). Assessing the impact of eliminating affirmative action in higher education. *Journal of Labor Economics 28*(1), 113–166.

Long, M. C. (2004). Race and college admissions: An alternative to affirmative action? *Review of Economics and Statistics 86*(4), 1020–1033.

Mello, U. (2020). Does affirmative action in undergraduate education impact high school choice? Technical report, Working Paper.

Pathak, P. A., A. Rees-Jones, and T. Sonmez (2020). Reversing reserves. Working Paper 26767, NBER.

Roy, A. D. (1951, 06). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers 3*(2), 135–146.

Sander, R. H. (2004). A systemic analysis of affirmative action in american law schools. *Stanford Law Review 57*(2), 367–483.

Singer, P. (2011). *Practical Ethics.* Cambridge University Press; 3rd edition.

Sonmez, T. and M. B. Yenmez (2021). Affirmative action in india via vertical, horizontal , and overlapping reservations. Working paper, Boston College.

Appendices

A Additional Tables

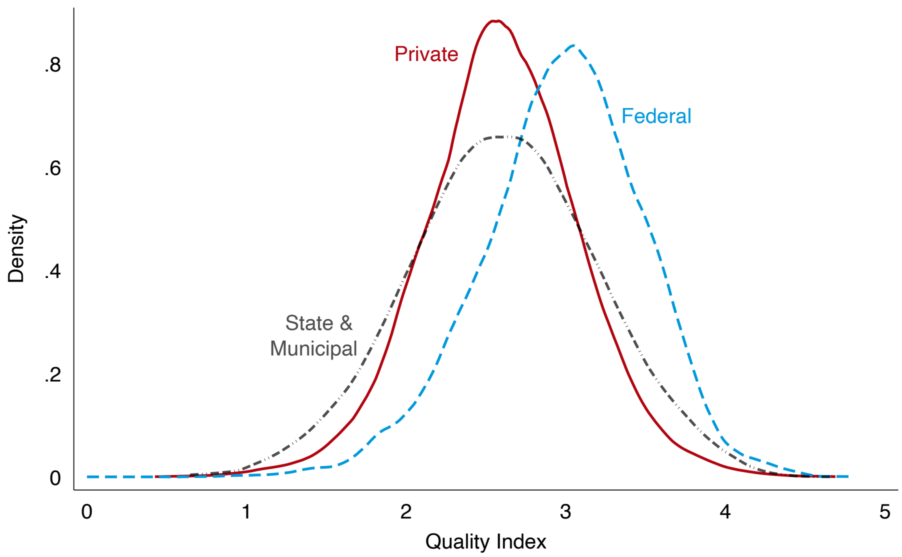
Table A.1: Administrative sector and type categories (as of 2016)

	Enrollment (# in 1,000)	Enrollment (share %)	Institutions	Degrees
Panel A: by sector, all students				
Federal	1,249	15.7	107	6,361
State	623	7.8	123	3,541
Municipal	47	0.6	45	264
Private	6,058	75.9	2,110	22,827
Panel B: by institution type, only federal enrollment				
University	1,083	86.7	63	5,005
Vocational	165	13.3	44	1,356

Notes: Calculations based on the Brazilian Higher Education micro data of 2016. Panel A in this table shows the breakdown of the number and share of students enrolled by sector. Panel B shows the enrollment number and share by institutions type among students enrolled in federal institutions. The total number of students is in 1,000s. Sector refers to whether the higher education institution is public (Federal, State, or Municipal administered) or private. Type refers to whether the institution is a university, or a vocational institution.

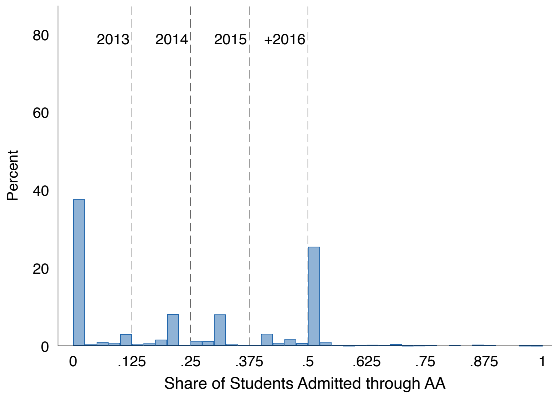
B Additional Figures

Figure B.1: Quality distribution

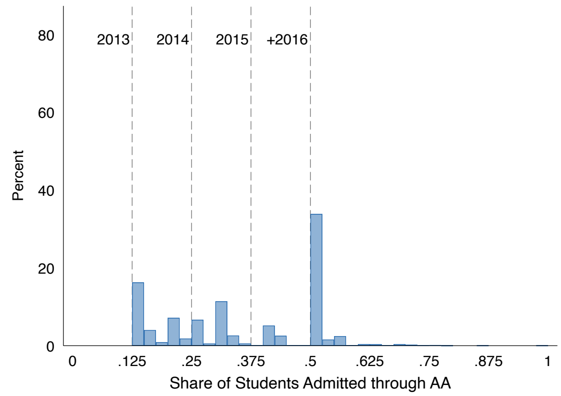


Notes: This figure shows the distribution of quality as measured by an index ranging from 1 to 5 prepared by the Ministry of Education of Brazil to evaluate degrees between years 2014 and 2016. An observation is a degree, and each observation is weighted by the number of students enrolled in the degree. Degrees from state and municipal institutions are pooled together.

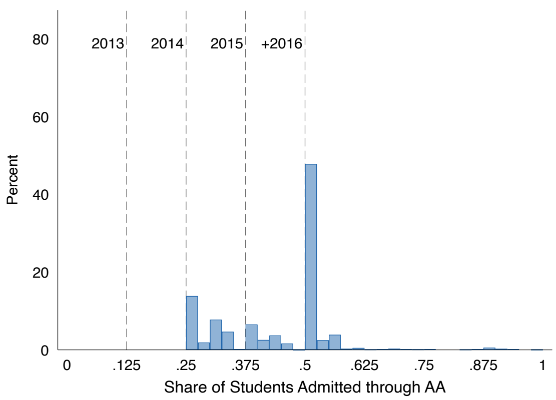
Figure B.2: Staggered Implementation of the policy



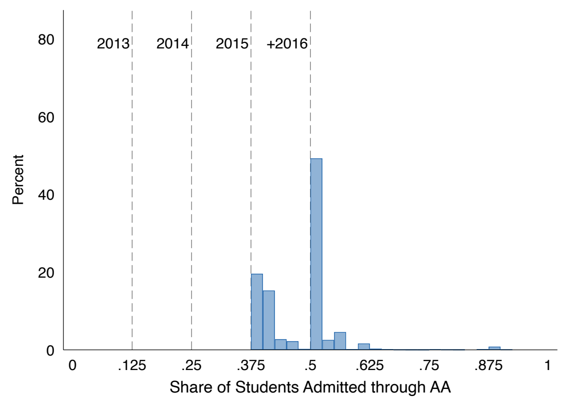
(a) 2012



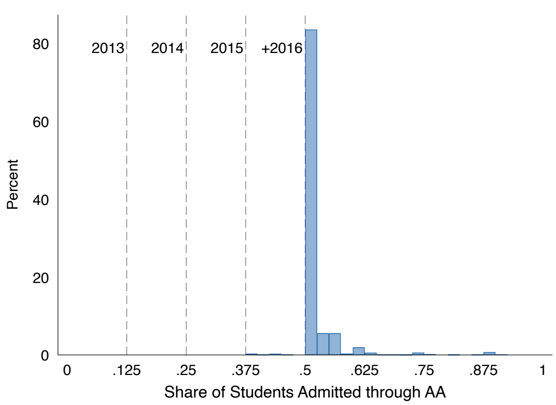
(b) 2013



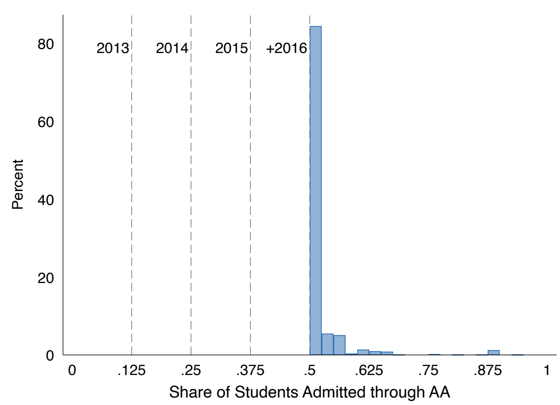
(c) 2014



(d) 2015



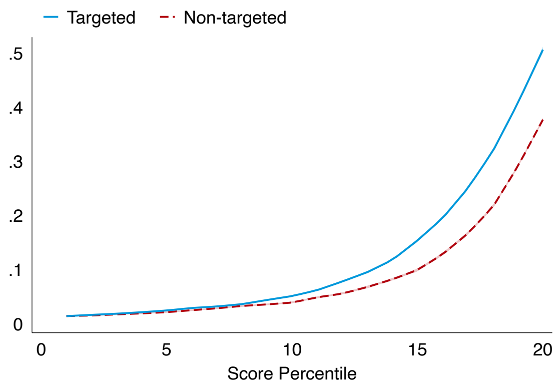
(e) 2016



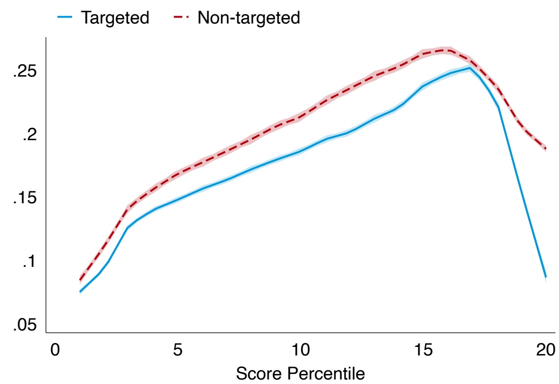
(f) 2017

Notes: These figures describe the implementation of the affirmative action regulation. An observation is a degree program, and the y-axis measures the share of student that got admitted through the affirmative action admission track in each of the years. The dashed lines represents 12.5%, 25%, 30% and 50%. This figure only uses data from institutions participating from the SISU process.

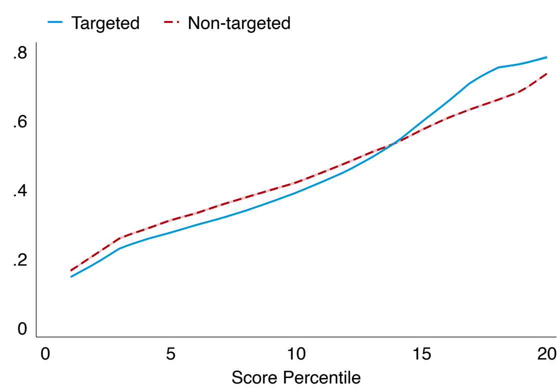
Figure B.3: College Attendance in 2016



(a) Enroll Federal



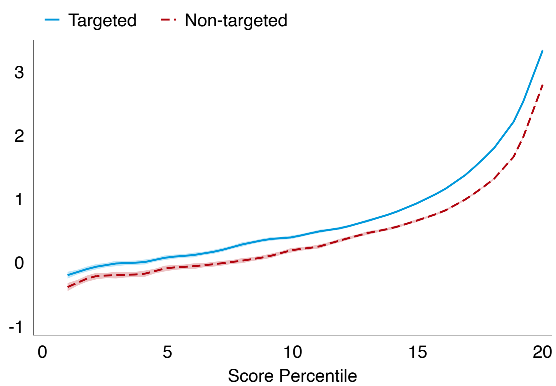
(b) Enroll Private



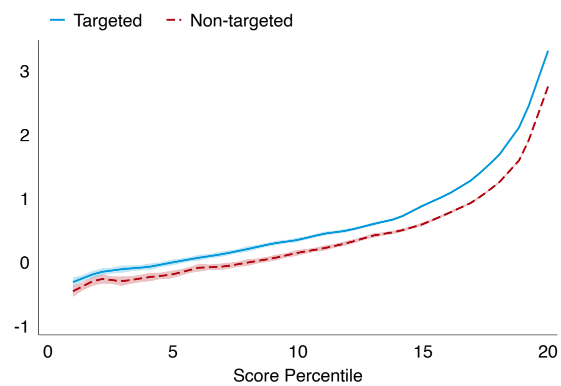
(c) Attend Any

Notes: These figures describe...

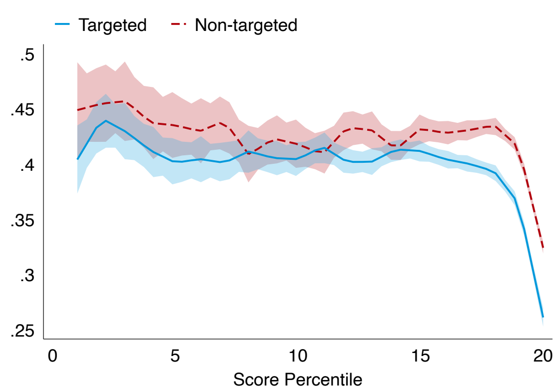
Figure B.4: College Outcomes — Enrolling Federal in 2016



(a) Quality 2016



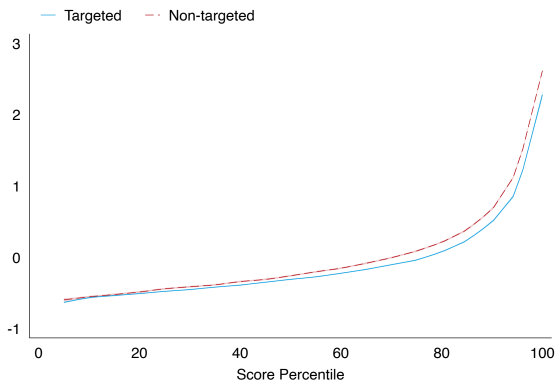
(b) Quality 2019



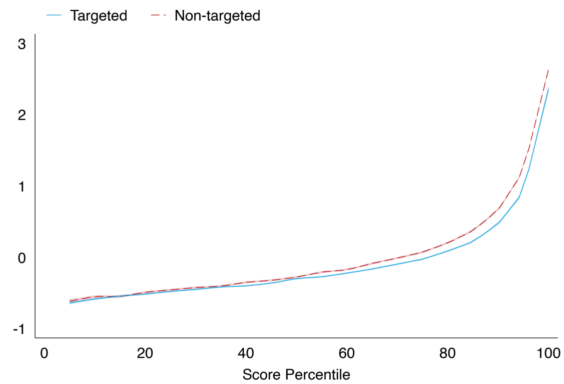
(c) 4-year Dropout

Notes: These figures describe...

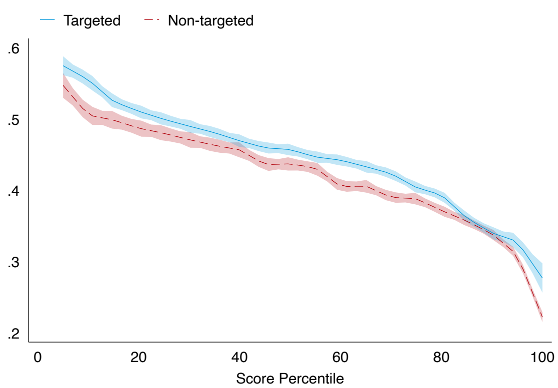
Figure B.5: College Outcomes — Enrolling Private in 2016



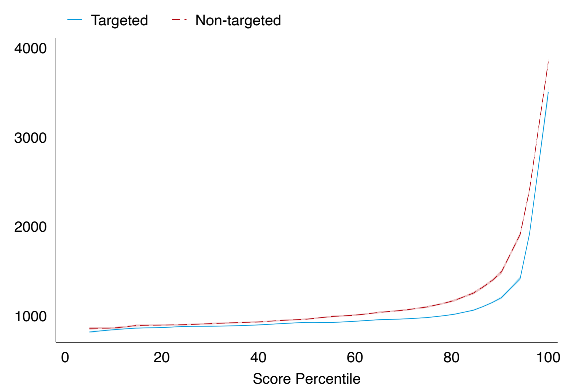
(a) Quality 2016



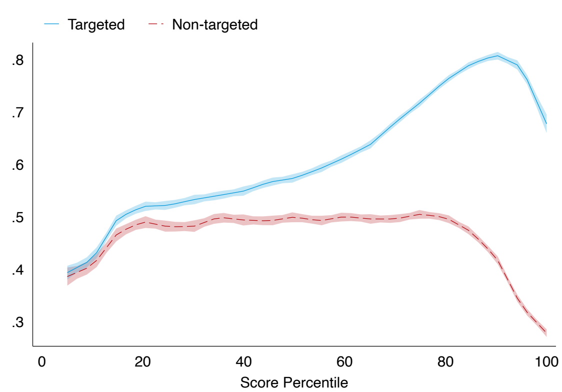
(b) Quality 2019



(c) 4-year Dropout



(d) Degree Price



(e) Financial Aid

Notes: These figures describe...

C Predicted income

Examples of trajectories. Assume 0 is the outside option, and 1 and 2 are degree programs in public and private institutions respectively

	Year 1	Year 2	Year 3	Year 4
Student A	1	1	1	1
Student B	1	0	0	0
Student C	1	2	2	2
Student D	0	0	0	2
Student E	2	2	2	2

Our working sample are all ENEM takers from years 2009, 2010, 2011 and 2012 We observe 2017 income from matched employer-employee records (RAIS) Non-parametric matching is very demanding for degrees with small samples For now we summarize the trajectory using degree attainment in year 1 and year 4

$$y_{i,T} = \alpha_t + X_i\beta + \delta_{J(i,t+1)} + \delta_{J(i,t+4)} + \varepsilon_i$$

X is a vector of covariates including, test scores and traits targeted by the AA regulation α_t is the ENEM year, and T denotes the year in which we observe income δ are degree fixed effects, $J(i,t)$ is a function indicating the degree that student i attends in period t We use these coefficients to predict earnings of 2015 ENEM takers (SISU 2016)

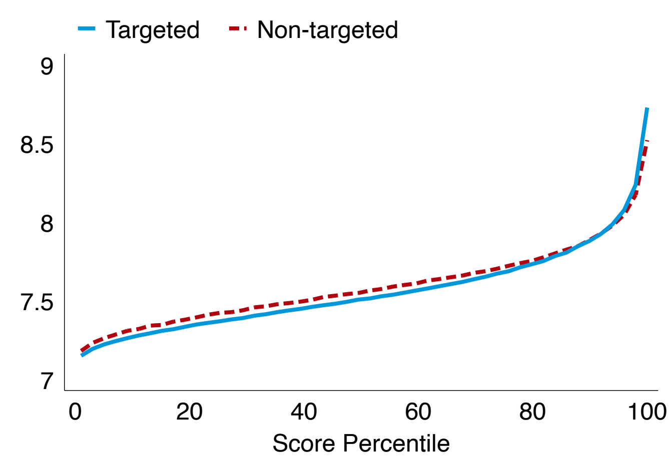


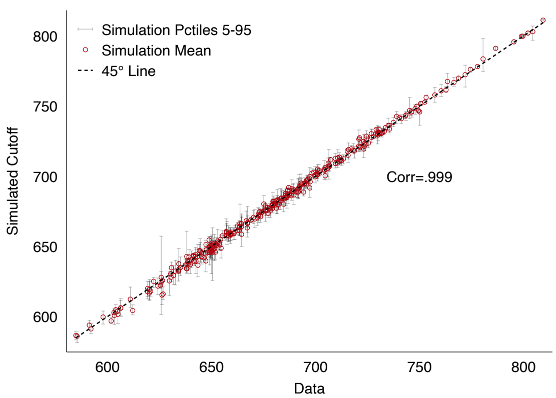
Figure C.1: Log (Predicted Income)

D Model Fit

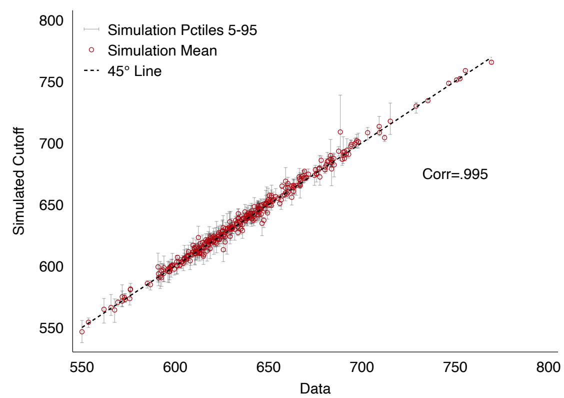
To verify that our parameters can recover the observed allocation, we simulate the admission cutoffs predicted by the model and contrast them against those observed in the data. To recover the admission cutoffs we follow steps 1 to 5 described in Appendix I, and simulate 20 different admission cutoffs for each of the different degree programs.

Panel (a) and (b) in Figure D show the model fit for the admission cutoffs of open and reserved seats, respectively. In the horizontal axis, we plot the observed admission cutoff scores, while in the vertical axis, we plot the average simulated admission cutoff score across all simulations, and the 95% coverage interval. It is important to note that in our data we observe four different admission cutoffs for the reserved spots: one for each affirmative action type. To construct the observed cutoff we take a weighted average of the four admission cutoffs where the weights correspond to the share of students of each affirmative action type applying on SISU. We observe that admission cutoffs predicted by our model are very similar to those observed in the data. The correlation coefficient of the slope is virtually 1 for both open and reserved seats.

Figure D.1: Model fit



(a) Open



(b) Reserved

Notes: This figure shows the in-sample model fit for the admission cutoffs of open and reserved seats. In the horizontal axis, we plot the observed cutoff score, while in the vertical axis, we plot the average cutoff score across all simulations and the 95% coverage interval.

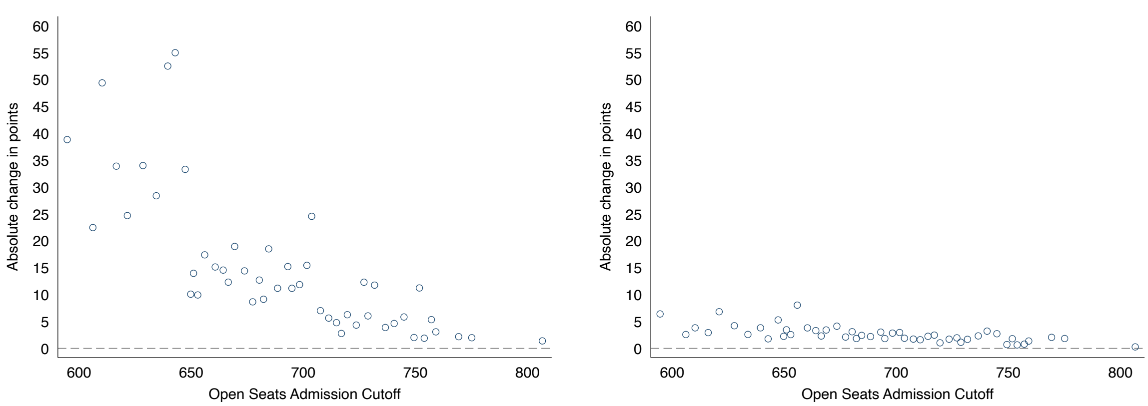
E Iterative Admission Cutoffs

The centralized admission system uses an iterative deferred acceptance mechanism. Under this system, students are sequentially asked to submit rank ordered lists over the course of several "trial" day. At the end of each day, the system produces a cutoff grade representing the lowest grade necessary to be accepted at a specific program. Each degree reports five admission cutoffs, one for each affirmative action admission track. In this appendix we explore the change in admission cutoffs over the course of the application period.

Unfortunately, there are no administrative records of the admission cutoffs reported by the system throughout the application period. The system only saves the final admission cutoffs. To circumvent this issue, we scrapped online data of the degrees offered by the Federal University of Minas Gerais (UFMG) during the 2021 admission period. In total we observe 450 admission cutoffs (90 degrees with 5 admission tracks each) over the 9 days of the application process. In Figure E.1 we show how the admission cutoffs evolve over time. Panel (a) displays the absolute difference between the final admission cutoffs (day T in the figure) and the one reported in the first day of the admission period. We observe large differences in admission cutoffs, especially for less selective degrees. Panel (b) displays the absolute difference between the final admission cutoffs (day T in the figure) and the one reported in last day of the admission period. The pattern in the data shows that at the last day of the application period, most degrees have converged or are very close to converging to the final admission cutoff.

Overall, these data suggest that the ex-ante and ex-post eligibility into degrees are very similar.

Figure E.1: Absolute Change in admission cutoffs



(a) Day 1 - Day T

(b) Day T-1 - Day T

Notes: This figure shows the absolute change in admission cutoffs over degree selectivity, as defined by the degree's final admission cutoffs of the open seats. Panel (a) display the difference between the absolute difference between the final admission cutoffs and those reported in the first day of the system. Panel (b) displays the difference between the absolute difference between the final admission cutoffs in those reported in the penultimate day of the system.

F Behavioral Responses

In this Appendix we explore the behavioral responses to the affirmative action regulation. We explore behavioral responses on three different margins. First, students might have responded by changing their pre-college human capital accumulation. Second, some students may have switched from private to public high-schools to gain eligibility as affirmative action students. Third, the affirmative action regulation might have changed the composition of applicants. We study each of these responses in turn.

F.1 High school movers

F.2 High school movers

To be eligible for the affirmative action admission tracks, students need to have completed grades 10 to 12 in a public high-school. In this appendix we study whether the affirmative action regulation impacted the switching rates from private to public high schools at grade 10. We use micro data from the school census in Brazil. This dataset is at the student leveI and allows us to follow the universe of students over time and across schools between 2009 and 2017.

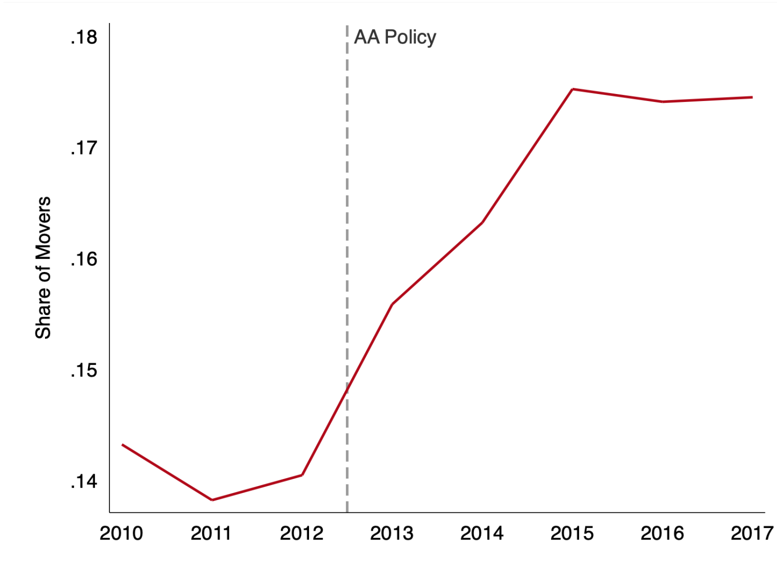
We start by analyzing the switching behavior of 9th graders who were enrolled in regular private schools.[34] The private sector represents about 20% of the total enrollment in a given year. In 2012, around 250,000 students enrolled in 10th grade after completing 9th grade in a private school the year before. From this sample, 14% enrolled in a public institution and the remaining stayed in the private sector. It is important to note that between grades 9 and 10, students transition from middle school to high school. Since students are forced to change schools, switching rates from the private to the public sector are higher than in any other grade.

In Figure F.1 we show how the share of movers from private to public institutions in 10th grade changes over time. The x-axis represents the year in which students attend 10th grade. We observe a big increase in the share of movers after the law was introduced in 2013. The rate of switchers revolves around 14% before the policy and stabilizes around 17.5% in 2015. This change in the trend of the switching rate is consistent with changes due the affirmative action policy but it could also mask other contemporaneous shocks affecting the decision to switch from the private to the public sector (e.g. improvements in the public sector infrastructure).

In order to have a comparison group that helps isolate the effect of the the policy from other confounding factors we use students switching to the public sector between grades 10 and 11. In contrast to individuals who switch between grades 9 and 10, these students cannot gain affirmative action status as a result of switching to a public institution. As such, it is unlikely that these movers are motivated by strategic reasons related to the affirmative action policy.

[34] We restrict our sample to students between 13 and 17 years of age.

Figure F.1: Switching rates to public schools between grades 9 and 10



Notes: This figure shows the switching rates from private to public schools between grades 9 and 10. The x-axis represents the year after the move. The y-axis represents the share of switchers over the universe of students enrolled in private schools in 9th grade.

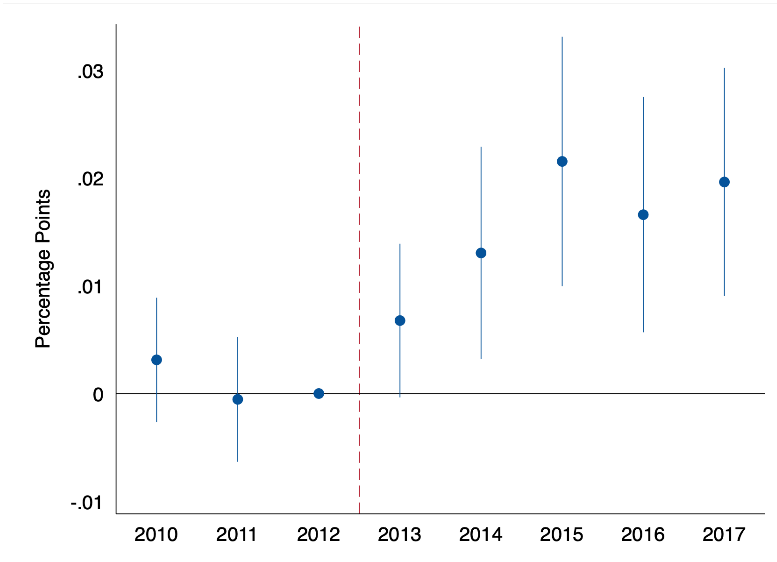
We estimate the following differences in differences model:

$$switch_{it} = \sum_j \beta_j \cdot Treat_i \cdot 1\{t = j\} + \alpha \cdot Treat_i + \delta_{d(i)t} + \varepsilon_{it} \tag{14}$$

where $d(i)$ and t stand for school district and year, respectively, and $Treat_i$ is an indicator variable taking the value 1 if student i is in 9th grade and 0 otherwise. We make the normalization $\beta_{2012} = 0$, so that all β_j coefficients represent differences in outcomes relative to the year before the policy was implemented. The estimates are presented in Figure F.2. We observe that before the affirmative action regulation is passed there are no differences in the pre-trend between switching rates of 9th graders relative to 10th graders. The coefficient grows between 2013 and 2015, and then stabilizes around 2 percentage points.

In 2016, roughly 70% of all applicants to SISU had graduated from high school before 2015. Thus, this subset of applicants did not have any strategic incentives to switch schools. The remaining 30% graduated from high school in 2015, out of which 27% came from a private high schools. Taking the estimates from Equation 14 at face value, implies that the total share of students from private high schools would have been 2 percentage points higher 0.3×0.27

Figure F.2: Dynamic differences-in-differences estimates



Notes: This figure shows the differences in differences β_j estimates from Equation 14. The coefficient are year-specific coefficient for individuals in 9th grade relative to individuals in 10th grade.

G Identification Proof

G.1 Setting

Assume there are $J = 4$ degrees $\mathcal{J} = \{A, B, C, D\}$. For simplicity we order them terms of selectivity such that the admission thresholds of A, B and C are given by $c_A > c_B > c_C$, and D is the outside option with unlimited capacity.

Students are characterized by their preferences $\succ_i$ and score s_i. All degrees give the same weight to all tests such that score does not depend on j. Note that everything can be generalize to having X_i, where s_i is just a component of X_i. In addition, every student-degree combination has a potential outcome given by:

$$\begin{aligned} Y_{ij} &= \mathbb{E}[Y_{ij}|s_i] + \varepsilon_{ij} \\ &= Y_j(s_i) + \varepsilon_{ij} \end{aligned}$$

In the data we observe choice sets Ω_i, choices probabilities $\pi_j(\Omega_i, s_i) = Pr(D_i = j \,|\, \Omega_i, s_i)$ and outcomes. The non-parametric form of the mean observed outcome in the data is:

$$E[Y_i|D_i = j, \Omega_i, s_i] \quad = \quad Y_j(s_i) + E[\varepsilon_{ij}|D_i = j, \Omega_i, s_i]$$

Our goal is to recover $Y_j(s_i)$. We explore two cases:

1. **Random allocation of choice sets**: By assigning choice sets at random, we have the following combinations of choice sets:

$$\Omega_i \in \{(A, B, C, D), (A, B, D), (A, C, D), (A, D), (B, C, D), (B, D), (C, D), (D)\}$$

2. **Deferred acceptance algorithm**: Under this mechanism, we would observe the following choice sets in the data:

$$\Omega_i \in \{(A, B, C, D), (B, C, D), (C, D), (D)\}$$

G.2 Random Choice Sets

When choice sets are random, we can flexibly condition on any given score $s_i = s$. Let Δ_j^{m-n} denote the difference in mean outcomes of individuals attending j between individuals with choice set Ω^m and Ω^n. Note that Δ_j^{m-n} is observed. We prove identification of $Y_j(s_i)$ for $j = A$. A similar argument applies for the rest of degrees.

From the definition of Δ_j^{m-n} we know that:

$$\begin{aligned}
\Delta_A^{(A,B,C,D)-(A,B,D)} &= E[Y_i|D_i = A, \Omega_i = (A,B,C,D), s_i = s] - E[Y_i|D_i = A, \Omega_i = (A,B,D), s_i = s] \\
\Delta_A^{(A,B,C,D)-(A,D)} &= E[Y_i|D_i = A, \Omega_i = (A,B,C,D), s_i = s] - E[Y_i|D_i = A, \Omega_i = (A,D), s_i = s] \\
\Delta_A^{(A,C,D)-(A,D)} &= E[Y_i|D_i = A, \Omega_i = (A,C,D), s_i = s] - E[Y_i|D_i = A, \Omega_i = (A,D), s_i = s] \\
\Delta_A^{(A,B,D)-(A,D)} &= E[Y_i|D_i = A, \Omega_i = (A,B,D), s_i = s] - E[Y_i|D_i = A, \Omega_i = (A,D), s_i = s]
\end{aligned}$$

Using Equation (15) we can rewrite the previous system of equations as:

$$\begin{aligned}
\Delta_A^{(A,B,C,D)-(A,B,D)} &= E[\varepsilon_{iA}|D_i = A, \Omega_i = (A,B,C,D), s_i = s] - E[\varepsilon_{iA}|D_i = A, \Omega_i = (A,B,D), s_i = s] \\
\Delta_A^{(A,B,C,D)-(A,D)} &= E[\varepsilon_{iA}|D_i = A, \Omega_i = (A,B,C,D), s_i = s] - E[\varepsilon_{iA}|D_i = A, \Omega_i = (A,D), s_i = s] \\
\Delta_A^{(A,C,D)-(A,D)} &= E[\varepsilon_{iA}|D_i = A, \Omega_i = (A,C,D), s_i = s] - E[\varepsilon_{iA}|D_i = A, \Omega_i = (A,D), s_i = s] \\
\Delta_A^{(A,B,D)-(A,D)} &= E[\varepsilon_{iA}|D_i = A, \Omega_i = (A,B,D), s_i = s] - E[\varepsilon_{iA}|D_i = A, \Omega_i = (A,D), s_i = s]
\end{aligned}$$

which leads to a system of 4 equations and 4 unknowns. This means that the unobserved component of equation (15) is identified. Since we know $E[\varepsilon_{ij}|D_i, \Omega_i, s_i]$, and we observe $E[Y_i|D_i, \Omega_i, s_i]$, we use Equation (15) to recover $Y_j(s_i)$.

G.3 Deferred acceptance mechanism

When individuals are assigned using a deferred acceptance mechanism, we can only estimate parameters for individuals in the neighborhood of the admission threshold scores $s_i = c_A$, $s_i = c_B$, and $s_i = c_C$.

Let Δ_j^s denote the difference in mean outcomes of individuals attending j between individuals with score $s^+ = s + \epsilon$ and score $s^- = s - \epsilon$ (with $\epsilon \to 0$). Then:

$$\begin{aligned}
\Delta_B^{c_A} &= E[Y_i|D_i = B, \Omega_i = (A,B,C,D), s_i = c_A^+] - E[Y_i|D_i = B, \Omega_i = (B,C,D), s_i = c_A^-] \\
\Delta_C^{c_A} &= E[Y_i|D_i = C, \Omega_i = (A,B,C,D), s_i = c_A^+] - E[Y_i|D_i = C, \Omega_i = (B,C,D), s_i = c_A^-] \\
\Delta_D^{c_A} &= E[Y_i|D_i = D, \Omega_i = (A,B,C,D), s_i = c_A^+] - E[Y_i|D_i = D, \Omega_i = (B,C,D), s_i = c_A^-] \\
\Delta_C^{c_B} &= E[Y_i|D_i = C, \Omega_i = (B,C,D), s_i = c_B^+] - E[Y_i|D_i = C, \Omega_i = (C,D), s_i = c_B^-] \\
\Delta_D^{c_B} &= E[Y_i|D_i = D, \Omega_i = (B,C,D), s_i = c_B^+] - E[Y_i|D_i = D, \Omega_i = (C,D), s_i = c_C^-] \\
\Delta_D^{c_C} &= E[Y_i|D_i = D, \Omega_i = (C,D), s_i = c_C^+] - E[Y_i|D_i = D, \Omega_i = (D), s_i = c_C^-]
\end{aligned}$$

Using Equation (15) we can rewrite the previous system of equations as:

$$\begin{aligned}
\Delta_B^{c_A} &= E[\varepsilon_{iB}|D_i = B, \Omega_i = (A,B,C,D), s_i = c_A^+] - E[\varepsilon_{iB}|D_i = B, \Omega_i = (B,C,D), s_i = c_A^-] \\
\Delta_C^{c_A} &= E[\varepsilon_{iC}|D_i = C, \Omega_i = (A,B,C,D), s_i = c_A^+] - E[\varepsilon_{iC}|D_i = C, \Omega_i = (B,C,D), s_i = c_A^-] \\
\Delta_D^{c_A} &= E[\varepsilon_{iD}|D_i = D, \Omega_i = (A,B,C,D), s_i = c_A^+] - E[\varepsilon_{iD}|D_i = D, \Omega_i = (B,C,D), s_i = c_A^-] \\
\Delta_C^{c_B} &= E[\varepsilon_{iC}|D_i = C, \Omega_i = (B,C,D), s_i = c_B^+] - E[\varepsilon_{iC}|D_i = C, \Omega_i = (C,D), s_i = c_B^-] \\
\Delta_D^{c_B} &= E[\varepsilon_{iD}|D_i = D, \Omega_i = (B,C,D), s_i = c_B^+] - E[\varepsilon_{iD}|D_i = D, \Omega_i = (C,D), s_i = c_B^-] \\
\Delta_D^{c_C} &= E[\varepsilon_{iD}|D_i = D, \Omega_i = (C,D), s_i = c_C^+] - E[\varepsilon_{iD}|D_i = D, \Omega_i = (D), s_i = c_C^-]
\end{aligned}$$

This leads to a system of 6 equations and 12 unknowns, and therefore is not identified. We need to make additional assumptions.

We define $\lambda_k(\cdot)$ as a function that maps from students unobservables to choice probabilities. We can choose any functional form of λ as long as [We need to check if this is true]

$$\lambda_k(D_i, \Omega_i, s_i) > \lambda_h(D_i, \Omega_i, s_i) \Leftrightarrow Pr(D_i = k|\Omega_i, s_i, \succ_i) > Pr(D_i = h|\Omega_i, s_i, \succ_i)$$

Given the functional form that we choose, $\lambda_k(\cdot)$ is observed.

A potential assumption is to restrict the way in which ε_i relates to choices. Following Dubin and McFadden (1984), we impose a linear relationship between potential outcomes and the unobserved error component

$$\begin{aligned}
E[\varepsilon_{ij}|D_i, \Omega_i, s_i] &= \psi_A(s_i)\lambda_A(D_i, \Omega_i, s_i) + \psi_B(s_i)\lambda_B(D_i, \Omega_i, s_i) + \psi_C(s_i)\lambda_C(D_i, \Omega_i, s_i) \\
&\quad + \psi_D(s_i)\lambda_D(D_i, \Omega_i, s_i) + \rho(s_i)\lambda_j(D_i, \Omega_i, s_i)
\end{aligned}$$

With this restriction in place, we can rewrite our identifying equations:

$$
\begin{aligned}
\Delta_B^{c_A} \quad = \quad & \psi_A(c_A)\left[\lambda_A(D_i = B, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_A(D_i = B, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& + \psi_B(c_A)\left[\lambda_B(D_i = B, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_B(D_i = B, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& + \psi_C(c_A)\left[\lambda_C(D_i = B, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_C(D_i = B, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& + \psi_D(c_A)\left[\lambda_D(D_i = B, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_D(D_i = B, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& + \rho(c_A)\left[\lambda_B(D_i = B, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_B(D_i = B, \Omega_i = (B,C,D), s_i = c_A)\right] \\
\Delta_C^{c_A} \quad = \quad & \psi_A(c_A)\left[\lambda_A(D_i = C, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_A(D_i = C, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& + \psi_B(c_A)\left[\lambda_B(D_i = C, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_B(D_i = C, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& + \psi_C(c_A)\left[\lambda_C(D_i = C, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_C(D_i = C, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& + \psi_D(c_A)\left[\lambda_D(D_i = C, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_D(D_i = C, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& + \rho(c_A)\left[\lambda_C(D_i = C, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_C(D_i = C, \Omega_i = (B,C,D), s_i = c_A)\right] \\
\Delta_D^{c_A} \quad = \quad & \psi_A(c_A)\left[\lambda_A(D_i = D, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_A(D_i = D, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& + \psi_B(c_A)\left[\lambda_B(D_i = D, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_B(D_i = D, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& + \psi_C(c_A)\left[\lambda_C(D_i = D, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_C(D_i = D, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& + \psi_D(c_A)\left[\lambda_D(D_i = D, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_D(D_i = D, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& + \rho(c_A)\left[\lambda_D(D_i = D, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_D(D_i = D, \Omega_i = (B,C,D), s_i = c_A)\right]
\end{aligned}
$$

and so on for the other thresholds...

Note that here we have only 3 equations and 5 parameters ($\psi_A(c_A)$,$\psi_B(c_A)\psi_C(c_A)$,$\psi_D(c_A)$,$\rho(c_A)$) and therefore the model is not identified. We need an additional restriction.

By imposing that $\psi_j(s) = \psi_k$ and that $\rho(s) = \rho$, I will show now that the model is identified.

We can rewrite our previous conditions:

$$
\begin{aligned}
\Delta_B^{c_A} \;=\;& \psi_A\left[\lambda_A(D_i = B, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_A(D_i = B, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& +\psi_B\left[\lambda_B(D_i = B, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_B(D_i = B, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& +\psi_C\left[\lambda_C(D_i = B, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_C(D_i = B, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& +\psi_D\left[\lambda_D(D_i = B, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_D(D_i = B, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& +\rho\left[\lambda_B(D_i = B, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_B(D_i = B, \Omega_i = (B,C,D), s_i = c_A)\right] \\
\Delta_C^{c_A} \;=\;& \psi_A\left[\lambda_A(D_i = C, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_A(D_i = C, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& +\psi_B\left[\lambda_B(D_i = C, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_B(D_i = C, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& +\psi_C\left[\lambda_C(D_i = C, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_C(D_i = C, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& +\psi_D\left[\lambda_D(D_i = C, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_D(D_i = C, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& +\rho\left[\lambda_C(D_i = C, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_C(D_i = C, \Omega_i = (B,C,D), s_i = c_A)\right] \\
\Delta_D^{c_A} \;=\;& \psi_A\left[\lambda_A(D_i = D, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_A(D_i = D, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& +\psi_B\left[\lambda_B(D_i = D, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_B(D_i = D, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& +\psi_C\left[\lambda_C(D_i = D, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_C(D_i = D, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& +\psi_D\left[\lambda_D(D_i = D, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_D(D_i = D, \Omega_i = (B,C,D), s_i = c_A)\right] \\
& +\rho\left[\lambda_D(D_i = D, \Omega_i = (A,B,C,D), s_i = c_A) - \lambda_D(D_i = D, \Omega_i = (B,C,D), s_i = c_A)\right] \\
\Delta_C^{c_B} \;=\;& \psi_A\left[\lambda_A(D_i = C, \Omega_i = (B,C,D), s_i = c_B) - \lambda_A(D_i = C, \Omega_i = (C,D), s_i = c_B)\right] \\
& +\psi_B\left[\lambda_B(D_i = C, \Omega_i = (B,C,D), s_i = c_B) - \lambda_B(D_i = C, \Omega_i = (C,D), s_i = c_B)\right] \\
& +\psi_C\left[\lambda_C(D_i = C, \Omega_i = (B,C,D), s_i = c_B) - \lambda_C(D_i = C, \Omega_i = (C,D), s_i = c_B)\right] \\
& +\psi_D\left[\lambda_D(D_i = C, \Omega_i = (B,C,D), s_i = c_B) - \lambda_D(D_i = C, \Omega_i = (C,D), s_i = c_B)\right] \\
& +\rho\left[\lambda_C(D_i = C, \Omega_i = (B,C,D), s_i = c_B) - \lambda_C(D_i = C, \Omega_i = (C,D), s_i = c_B)\right] \\
\Delta_D^{c_B} \;=\;& \psi_A\left[\lambda_A(D_i = D, \Omega_i = (B,C,D), s_i = c_B) - \lambda_A(D_i = D, \Omega_i = (C,D), s_i = c_B)\right] \\
& +\psi_B\left[\lambda_B(D_i = D, \Omega_i = (B,C,D), s_i = c_B) - \lambda_B(D_i = D, \Omega_i = (C,D), s_i = c_B)\right] \\
& +\psi_C\left[\lambda_C(D_i = D, \Omega_i = (B,C,D), s_i = c_B) - \lambda_C(D_i = D, \Omega_i = (C,D), s_i = c_B)\right] \\
& +\psi_D\left[\lambda_D(D_i = D, \Omega_i = (B,C,D), s_i = c_B) - \lambda_D(D_i = D, \Omega_i = (C,D), s_i = c_B)\right] \\
& +\rho\left[\lambda_D(D_i = D, \Omega_i = (B,C,D), s_i = c_B) - \lambda_C(D_i = D, \Omega_i = (C,D), s_i = c_B)\right]
\end{aligned}
$$

$$
\begin{aligned}
\Delta_D^{c_C} \quad = \quad & \psi_A \left[\lambda_A(D_i = D, \Omega_i = (C,D), s_i = c_C) - \lambda_A(D_i = D, \Omega_i = (D), s_i = c_C)\right] \\
& +\psi_B \left[\lambda_B(D_i = D, \Omega_i = (C,D), s_i = c_C) - \lambda_B(D_i = D, \Omega_i = (D), s_i = c_C)\right] \\
& +\psi_C \left[\lambda_C(D_i = D, \Omega_i = (C,D), s_i = c_C) - \lambda_C(D_i = D, \Omega_i = (D), s_i = c_C)\right] \\
& +\psi_D \left[\lambda_D(D_i = D, \Omega_i = (C,D), s_i = c_C) - \lambda_D(D_i = D, \Omega_i = (D), s_i = c_C)\right] \\
& +\rho \left[\lambda_D(D_i = D, \Omega_i = (C,D), s_i = c_C) - \lambda_C(D_i = D, \Omega_i = (D), s_i = c_C)\right]
\end{aligned}
$$

Note that now we have 6 equations and 5 unknown, and therefore the parameters $(\psi_A, \psi_B, \psi_C, \psi_D, \rho)$ are identified. Once those parameters are identified, we can plug them in in equation (15) and recover $Y_j(s_i)$.

G.4 Adding grader instrument

Before discussing identification using graders, it is important to change the model so that potential outcomes depend on ability instead of score:

$$E[Y_i|D_i, \Omega(s_i), a_i] = Y(D_j, a_i) + E[\varepsilon_{ij}|D_i, \Omega(s_i), a_i]$$

Note here that the choice sets still depend on the score but everything else on ability.

Adding the graders allows us to relax some of our parametric assumptions we had before. Note that in the extreme case that graders move students across the whole score distribution, we can identify a relatively flexible model in which $(\psi_A(a)$,$\psi_B(a)\psi_C(a)$,$\psi_D(a)$,$\rho(a))$ vary with ability. If the graders are not enough to cross all thresholds, we may need to add additional assumptions (e.g. $\psi_A(a) = \psi_A^0 + \psi_A^1 \cdot a_i$). I'm not going to write all equations because they are too many, but you hopefully get the point.

H Capacity constraints in the private sector

I Simulation procedure

In this Appendix we describe the procedure used to simulate the counterfactuals. Let M denote the number of Monte Carlo simulations. The m^{th} simulations works as follows:

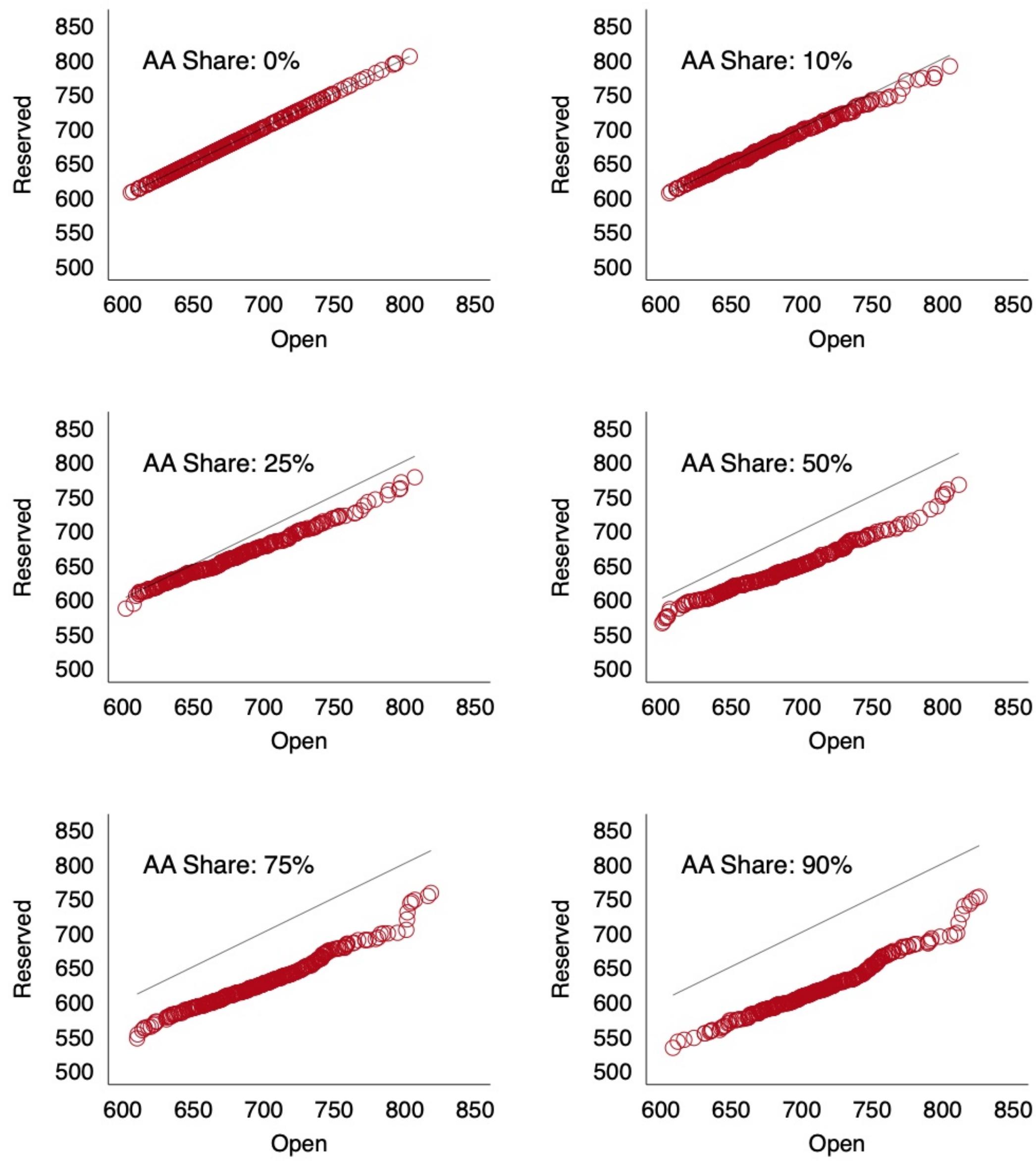
1. Simulate a vector of unobserved tastes $\eta_{ij}^m \sim EVT1$
2. Compute preferences $\hat{\succ}_i^m$ reflecting indirect utilities $\hat{u}_{ij}^m$ using preferences estimates $(\hat{\delta}_j^t, \hat{\gamma}_j^t, \hat{\kappa}^t)$ together with η_{ij}^m
3. Construct the set of student types as $\hat{\Theta}^m = \bigcup_i \hat{\theta}_i^m$, where $\hat{\theta}_i^m = (\hat{\succ}_i^m, s_i, t_i)$
4. Compute the matching function $\varphi(\hat{\Theta}^m, q, \omega) = \hat{\mu}^m$ based on mechanism φ's allocation rules.
5. Calculate the cutoff scores $c_{jt}^m(\hat{\mu}^m)$ that are consistent with the equilibrium
6. Use Equation (9) to compute the predicted potential outcome as $\hat{Y}_{ij} = \hat{\alpha}_j + X_{ij}'\hat{\beta}_j + \sum_{k=1}^J \hat{\psi}_k \cdot (\hat{\eta}_{ik}^m - \mu_\eta) + \hat{\rho} \cdot (\hat{\eta}_{ij}^m - \mu_\eta)$
7. Compute the expect potential outcome for the corresponding matching function $\hat{Y}_i(\hat{\mu}^m) = \sum_j \mathbb{1}\{\hat{\mu}^m(\theta_i) = j\} \cdot \hat{Y}_{ij}$

Thus, the expected potential outcome for individual i under affirmative action schedule ω is:

$$\bar{\hat{Y}}_i(\omega) = \frac{1}{M}\sum_m \hat{Y}_i(\hat{\mu}^m) \tag{15}$$

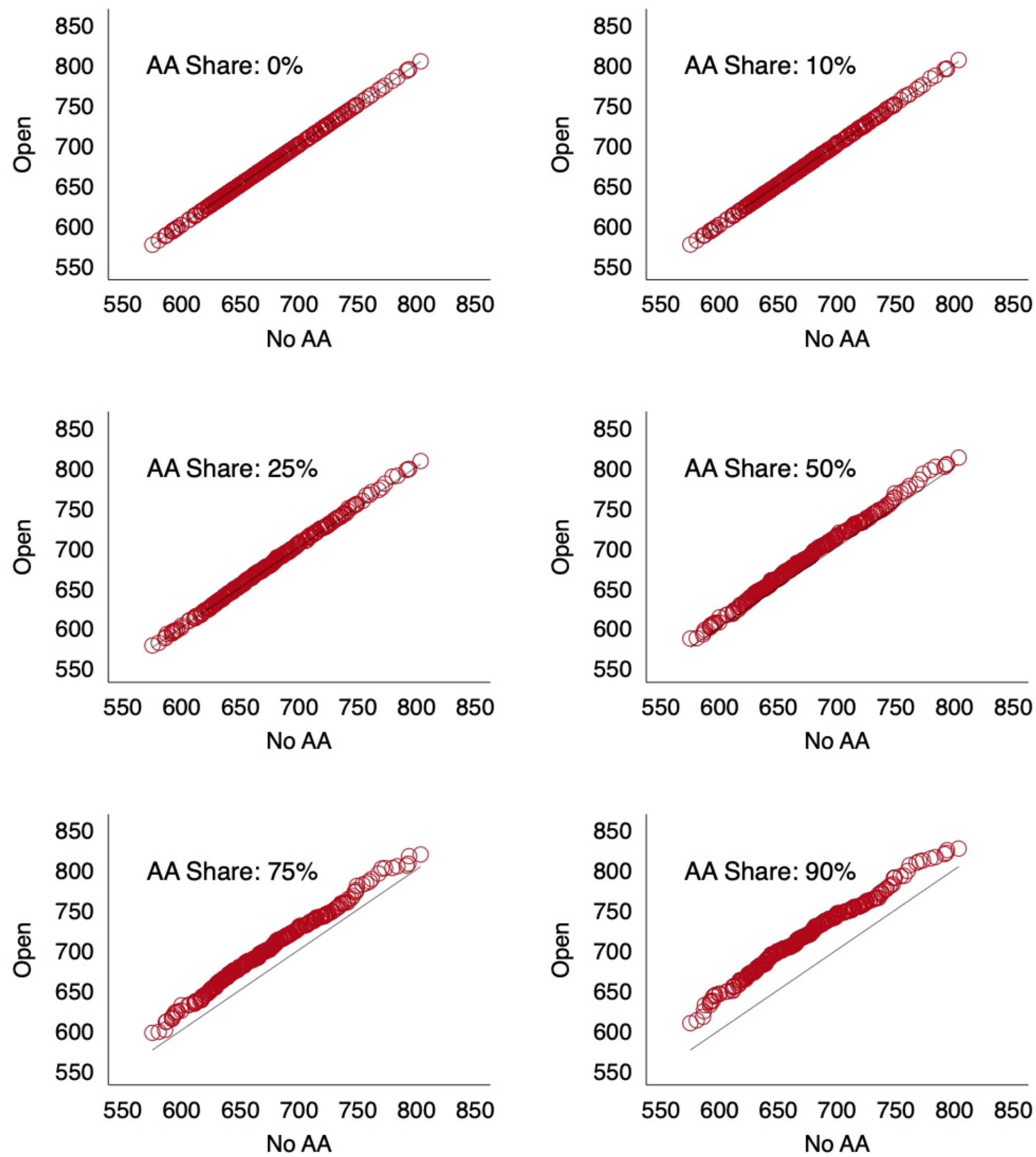
J Quantile-Quantile Plots

Figure J.1: Quantile-quantile admission cutoffs, Open vs Reserved



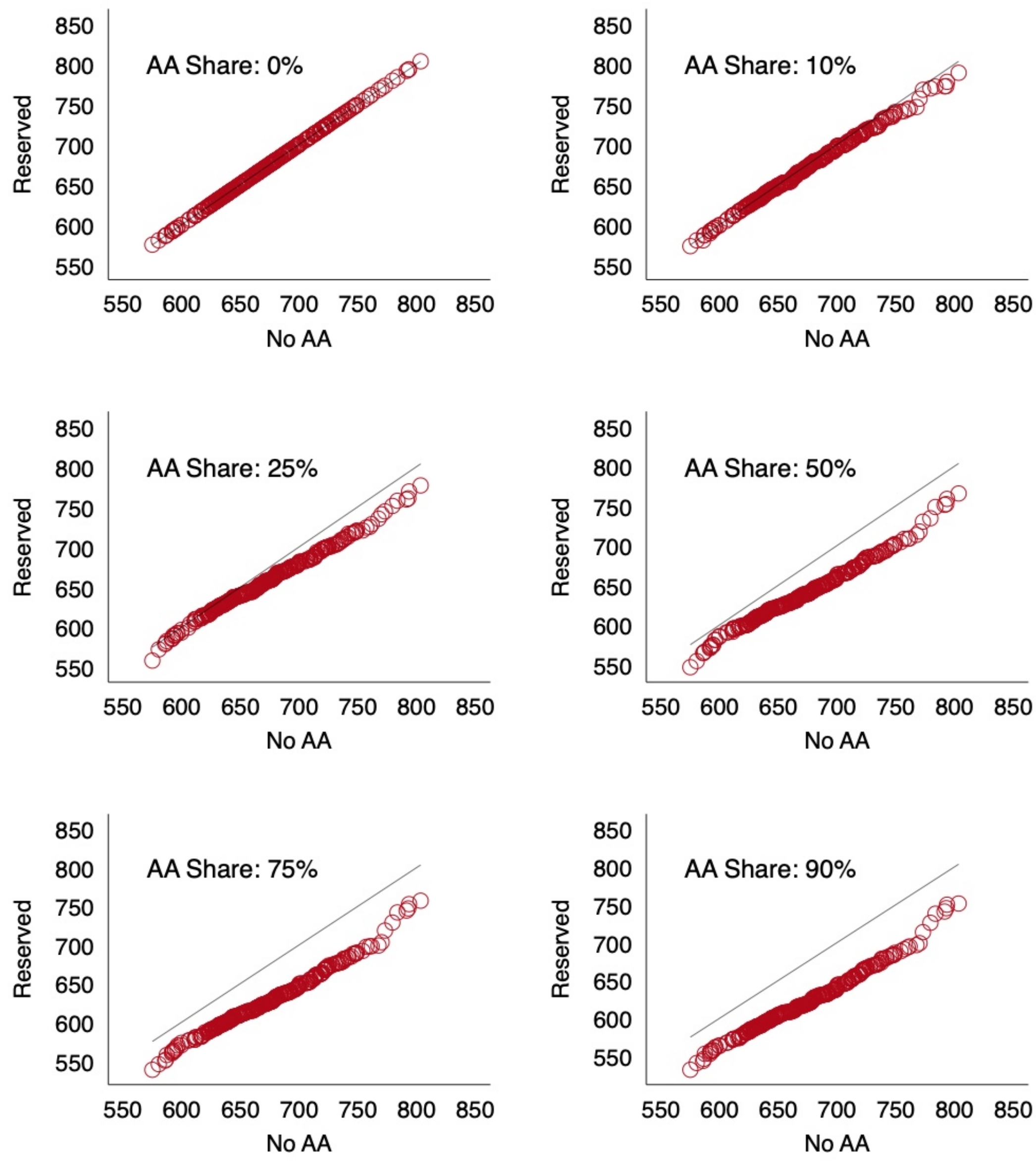
Notes: This figure shows the quantile-quantile plot of admission cutoffs of open and reserved spots, under different affirmative action schedules. The solid line is a 45 degree line. The x-axis and the y-axis represent the admission cutoff of open and reserved spots, respectively. The red circles denote the quantile associated to each of the admission cutoff value.

Figure J.2: Quantile-quantile plots, No AA vs Open



Notes: This figure shows the quantile-quantile plot of admission cutoffs of open spots. The solid line is a 45 degree line. The x-axis represent the admission cutoff of spots in the absence of the AA policy (i.e. $\omega = 0$). The y-axis represent the admission cutoff of open spots under different affirmative action schedules. The red circles denote the quantile associated to each of the admission cutoff values.

Figure J.3: Quantile-quantile plots, No AA vs Reserved



Notes: This figure shows the quantile-quantile plot of admission cutoffs of reserved spots. The solid line is a 45 degree line. The x-axis represent the admission cutoff of spots in the absence of the AA policy (i.e. $\omega = 0$). The y-axis represent the admission cutoff of reserved spots under different affirmative action schedules. The red circles denote the quantile associated to each of the admission cutoff values.