# Do disadvantaged students benefit from attending classes with more skilled colleagues? Evidence from a top university in Brazil*

Rodrigo Oliveira† Henrique Motte‡ Alei Santos §

December 2, 2024

**Abstract**

Using two rich administrative data sets and a rule of admission at one top university in Brazil that splits students into two classes, we apply a regression discontinuity design to study the effect of class allocation on academic performance and labor market outcomes. The last student of the first class will have higher-ability peers but a lower ordinal rank than the first student of the second class. These effects usually play in different directions. The main results suggest that the academic outcomes of affirmative action students in technology and health sciences majors are negatively impacted by being the last students in the first class. However, this negative effect does not translate into the labor market outcomes.

**Keywords:** affirmative action, peer effect, ranking effect, Brazil, education.
**JEL Codes:** I24. I25, I28, J15

*We thank Daniel da Mata, Edson Severnini, Fernanda Estevan, Robson Tigre, Diana Gonzaga, Vinícius Mendes, Yuri Barreto, the participants of the Brazilian Econometric Society Conference and UNU-Wider Summer School participants for all the comments. Usual disclaimer applies. Authors order was randomized using the AEA randomization tool (Confirmation code: vdf9O4zrTiCC).

†Contact information: Oliveira, Rodrigo: United Nations University, World Institute for Development Economics Research (UNU-WIDER). E-mail: oliveira@wider.unu.edu. Corresponding author.

‡Contact information: Motte, Henrique: Department of Economics, Federal University of Bahia (UFBA). E-mail: henriquezmotte@gmail.com

§Contact information: Santos, Alei: Getúlio Vargas Foundation, SP. E-mail: santos.alei@gmail.br

# 1 Introduction

Lower performance and mismatch – a situation where students would be better off if not enrolled in an elite university or prestigious major – is one of the main concerns surrounding affirmative action (AA) policies in higher education. Several studies estimate the causal effect of AA policies on students' performance and labor market outcomes in the US (Bleemer, 2022; Arcidiacono, 2005), India (Bagde et al., 2016), and Brazil (Francis-Tan and Tannuri-Pianto, 2018; Oliveira et al., 2023). However, these studies have not yielded conclusive results, though some suggest the presence of a certain degree of mismatch. The existing literature and experts suggest that mismatch primarily stems from misconceptions about the chosen major and inadequate preparation. In addition, other factors may contribute to the lower performance of affirmative action students, such as peer quality, peer pressure, a lack of sense of belonging, motivation, and future expectations. Nonetheless, the causal effects of peers and ranking on affirmative action students' academic and labor market outcomes in developing country contexts are not well documented. Therefore, this paper investigates whether a lower ranking outweighs the potential benefits of having better peers.

We leverage a unique rule of class assignment in a prestigious Brazilian university to examine the impact of class composition on AA students' academic and labor market outcomes. Under this assignment rule, first-year students are divided into classes that commence either at the beginning or in the middle of the subsequent academic year.[1] This division is based on their ranking in the entry exam and their applicant status – whether an AA applicant or not. The highest-ranked students are assigned to the first semester, which starts between February and March, while the remaining students begin classes in the second semester, between July and August. Within the same major and admission year the last student in the first class holds a lower relative rank compared to the first student in the second class, despite both having similar entrance exam scores. Furthermore, based on entrance exam scores, the peers in the first class are of higher quality. Therefore, the class allocation rule allows us to study the impact of being the last AA student among the best, compared to being the best AA student among the worst, on academic and labor market outcomes.

The identification strategy employed in this study utilizes a regression discontinuity design, which is based on the class assignment rule and the entrance exam score as the running variable. Before taking the exam, students are required to select the major they wish to apply for. They are not allowed to apply for more than one major or change their application within the same year. Once approved and enrolled, switching majors becomes

[1] The academic year in Brazil differs from the academic year in the US and Europe. Freshman students start the major in February or March of each year.

highly challenging. The most common way for students to switch majors is to drop out and retake the entry exam the following year.[2] The vestibular, a mandatory entrance exam, is conducted annually in November. Based on the exam results, a ranking is established and the top-performing students are selected, subject to slots being available. The number of slots is announced before the exam takes place. Irrespective of their applicant status or entrance exam scores, students are not permitted to choose their starting semester. AA and regular candidates do not compete directly against each other, and each type of student possesses its own rank based on the vestibular. Forty-five percent of the major slots in each major are reserved for AA students. Notably, both types of students will be present in the March and August first-year classes, and each class must have a 45-55 percent proportion of AA-regular students.

We use a rich administrative data set that comprises 4,843 AA students and 7,640 regular students who enrolled at the Federal University of Bahia (UFBA) between 2006 and 2012 in majors that employed a two-class division system. The data allow us to evaluate the class composition's impact on students' GPAs at the beginning and end of the major, dropouts, and failures. In addition, we merge the UFBA academic data with employee–employer administrative records, allowing us to track students' trajectories in the labor market between one to eight years after enrollment. This enables us to estimate the class composition effect on formal employment and earnings.

UFBA makes an interesting case study as one of the pioneering Brazilian federal universities to implement AA policies. Since 2005 the policy has established a reservation of 45% available slots in each major for former high-school students from the public education system. The policy also required at least one year of elementary schooling in a public institution. Additionally, UFBA is situated in Bahia state, which boasts the highest proportion of Black or mixed-race individuals among all Brazilian states, accounting for 83% of the population according to PNAD data.[3] Given the significant representation of the Black and mixed-race population in Bahia state and the eligibility criteria primarily based on the former high-school type, it becomes challenging for professors to discriminate based on observable characteristics such as race. Importantly, professors do not have access to information regarding whether a student is an AA applicant or not, nor do they possess knowledge of students' scores in the vestibular entrance exam.

The main findings suggest that being among the last AA students of the first class is worse than being among the first AA students of the second class for students in technology and health majors. Our identification strategy allows us to compare the first and last students within their group. For instance, for each last AA student of the first class, the comparison students will be the first AA student of the second class. The last AA

[2] For an extensive discussion about major switching at UFBA, see Oliveira et al. (2023).

[3] Brazilian Annual Household Survey, conducted by the Brazilian Institute of Statistics (IBGE).

students of the first class enrolled in technology majors have an average first-year GPA of 0.434 standard deviations (s.d.) lower than the first AA student in the second class. A similar result is observed when considering the average GPA in the final of the major. Moreover, these students fail in 4.33 more courses compared to the top-performing AA students in the second class. For AA students enrolled in health sciences majors, the negative effect on GPA emerges similarly in the first year and in the final GPA, while they fail in 4.14 more courses. Conversely, the results for the last AA students in the first class within social sciences majors indicate lower probability of dropout by –0.076 percentage points. We also estimate the effect of class allocation on the probability of employment in the formal market and wages. Leveraging a comprehensive nationwide employee–employer administrative data set, we are able to track all enrolled students for a period ranging from one to eight years after their enrollment, depending on the year they began the major. The main results suggest no effect for all outcomes, which differs with recent evidence for non-AA students (Ribas et al., 2020; Roux and Riehl., 2022).

An important contribution of our study lies in its examination of the dynamics within different fields of study. Each field attracts students with distinct skills and profiles. For instance, the health field has a higher proportion of female students, while the technology field has more male students and requires stronger mathematical abilities. The campuses for each field are located in different neighborhoods of the same city, reducing the interaction of students from different fields. Additionally, the majors in each field are exposed to different labor market conditions. These factors can significantly influence students' academic performance, future labor market outcomes, and their choices of majors.

The observed negative results for AA students in technology and health majors can be attributed to lower math and science aptitude among students from disadvantaged backgrounds.[4] We explore this issue by looking at specific courses students enroll in for the first semester. As expected, the class allocation has a higher impact on math- and science-related courses. Being among the last AA students of the first class significantly decreases calculus grades by -1.25 s.d. and microbiology and anatomy by -0.834 s.d. and -0.514 s.d., respectively.

We further investigate how peer quality and relative ranking can explain our results. We use a similar strategy proposed by Ribas et al. (2020) and try to disentangle the effects of peers and ranking. When analyzing AA students, our results differ from Ribas et al. (2020) for students in technological majors. When the peer quality difference is higher, the class allocation effect for AA students is stronger. This suggests that having better peers can

[4]In general, Brazilian public schools have lower quality than private schools at the elementary, middle, and high school levels. The performance of public school students relative to private school students is unusually low in Brazil, even when compared to similar countries. See, for example, Figures 3.13 to 3.15 in this 2021 OECD report comparing the quality of public and private school education in the OECD countries versus developing countries: https://www.oecd.org/publications/education-in-brazil-60a667f7-en.htm.

act as negative peer pressure when students are in the lower part of the ability distribution (Booij et al., 2016). However, there is no clear pattern for health major students. The lower the peer quality, the higher the negative effect on grades, while the higher the peer quality, the stronger the positive effect on failures. For social science students, the positive results are stronger when the peer quality difference is small.

This study makes at least three contributions to the literature. First, to the best of our knowledge this is the first paper to study the effect of class composition on AA students at college and university levels. Additionally, we add by looking for a representative and respected university in a low- and middle-income country. Prior studies have addressed these matters for regular students in two other universities in developing countries (Roux and Riehl., 2022; Ribas et al., 2020).

The novelty of our study lies in its focus on the outcomes of AA enrolled students. Nonetheless, after presenting the main results for AA students we also show the results for regular students to facilitate interpretation and comparison with the previous articles that solely concentrate on regular students. It is important to note, however, that AA and regular students are not directly comparable.[5] Therefore, caution must be exercised when comparing our results with theirs. Our findings diverge from theirs. Being the last regular student in the first class proves more advantageous than being the first regular student in the second class. These last students exhibit a 0.22 standard deviation higher first-semester GPA, 0.23 s.d. higher first-year GPA, and 0.22 s.d. higher final-year GPA.

These results are primarily driven by students enrolled in social sciences majors. It is important to highlight that although our study bears similarities to the aforementioned research, our specific context differs. The last regular student in the first class is not the overall last student in the class. Consequently, the mechanisms through which peers and ranking affect regular students diverge from those discussed in the published articles. In addition, we show that before the AA policy – using data for enrolled students in 2003 – the difference disappeared, which is more consistent with Ribas et al.'s (2020) and Roux and Riehl.'s (2022) findings.

The second contribution shows that the poor academic performance may appear when students in the lower part of the ability distribution curve are placed among high-skilled peers. There is extensive literature on disadvantaged students entering top colleges, suggesting no clear evidence of whether they benefit from enrolling at better colleges, with some studies pointing out that there is no mismatch (Bagde et al., 2016; Dale and Krueger, 2014; Bleemer, 2022) and others finding that there is (Arcidiacono, 2005; Arcidiacono and

[5] Appendix Table A.1 shows that the best regular students in the second class are, on average, at the top of their class, while the best AA students are not. Also, the worst AA student in the first class is, on average, the worst in their class, whereas this is not the case for the lowest-performing regular student. Additionally, because the lowest-performing regular students in the first class are not the lowest in the entire class, they are exposed to different peers and mechanisms compared to the previous studies by Ribas et al. (2020) and Roux and Riehl. (2022).

Lovenheim, 2016; Oliveira et al., 2023) at least in a small level. The mechanisms that explain these results are still unclear, and our findings suggest that peer and ranking effects may explain part of them. Our third contribution relates to the impact of class composition on labor market outcomes. Although there is solid evidence about the impact of class composition on academic outcomes, little is known about the labor market effects (Ribas et al., 2020; Roux and Riehl., 2022). We are the first study to link class composition, labor markets, and affirmative action.

As discussed, class composition may affect individuals' outcomes through peer and ranking effects. Peers play an important role in cognitive and non-cognitive skills formation and in labor market outcomes. The extensive literature on the topic shows evidence of a positive relationship to having better-performing peers.[6] However, there is also evidence that peers can be harmful (Bursztyn and Jensen, 2015; Bursztyn et al., 2019), and the negative effects are stronger when students are not at the top of the ability distribution (Booij et al., 2016). In other words, conditional on students' rank, better-performing peers can reduce academic success.

Other studies have shown that ordinal rank can explain academic success (Zeidner and Schleyer, 1999; Elsner and Isphording, 2016; Murphy and Weinhardt, 2020; Dasgupta et al., 2020; Elsner et al., 2021), or labor market outcomes (Ribas et al., 2020; Roux and Riehl., 2022) . The main explanation is that students with higher abilities in a low-ability group may have a misconception about their absolute ability and thus invest more in their education. This phenomenon is known as the "big fish in a small pond" effect (Marsh and Parker, 1984; Zeidner and Schleyer, 1999). The ordinal rank can affect students' achievement because better-ranked students' social networks can be more supportive. They can also be more motivated and self-confident (Elsner and Isphording, 2016).

# 2 Institutional background

Established in 1808, UFBA holds the distinction of being Brazil's first university. Today, it stands as one of the largest and most prestigious higher education institutions in the country and the second largest in the Northeast Region, both in terms of its physical infrastructure and student enrollment.[7] UFBA is entirely tuition-free, making it a crucial gateway to higher education, and particularly to a flagship college, especially for students from socioeconomically disadvantaged backgrounds.

Prospective students seeking admission to UFBA are required to participate in a com-

[6]For example, in elementary education (Rao, 2019; Hoxby, 2000), high school (Sund, 2009; Anelli and Peri, 2017; Eisenkopf, 2010), college (Ribas et al., 2020), and labor market Lepine and Estevan (2021)

[7]In 2017, UFBA had 105 undergraduate courses, 136 postgraduate courses (82 masters and 54 PhDs), and 42 postgraduate specialization courses. That same year, the UFBA budget was US$413,446,423.98 and it offered 8,875 vacancies to new students that did the vestibular in 2016.

prehensive entry exam known as the "vestibular." Held once per year, this exam evaluates students' proficiency in various subjects, including Portuguese grammar and reading, math, physics, chemistry, geography, biology, foreign language (English or Spanish), history, and philosophy. Students are then ranked based on their performance in the exam and admitted according to the number of available slots for each major. Importantly, students must choose their major before the entry exam. They cannot change their choice afterwards to adapt to their preparation effort – changing majors between the entrance exam and the admission process is not allowed. Students can change majors in very few cases after finishing the first semester following enrollment. As extensively discussed by Oliveira et al. (2023), switching majors accounts for only 3% of UFBA-enrolled students. Note also that because students choose the majors before the entrance exam, they do not know the minimum score to be selected, which depends on the performance of their competitors.[8]

UFBA was the second federal university to adopt affirmative action policies for admissions, but the first to offer 45% of the slots and to not focus only on racial criteria.[9] Enacted in 2004, this policy aimed to offer the opportunity to enter the state flagship university to students who only had access to lower-quality primary education, predominantly those from economically disadvantaged backgrounds. The eligibility criteria is being a former student from a public institution during high-school education. Furthermore, 85% of the enrolled AA students must be Black or mixed race.

Until 2012, 23 courses selected students into two classes. Students who achieved higher scores in the entry exam were assigned to start their university studies in March, while those with lower scores would start in August. Students could not choose which semester to start; this allocation was done using the entrance exam ranking for the major they applied for. Should a student be selected for the first semester but decide not to enroll at UFBA during that period, they would be required to retake the vestibular the following year.[10] The selection process was carried out independently for each group of students, obeying the descending order of the overall score calculated from the student's performance in the entry exam. For instance, if the course has 100 available slots for the first semester and 100 for the second semester, 45 of these slots in each semester will be filled by AA applicants.

The practice of dividing students into two classes at UFBA does not serve as a tracking policy, as one might initially assume based on Duflo et al. (2011) and Card and Giuliano (2016). Tracking policies aims to group students with similar abilities with the argument

[8] After 2013 the vestibular was replaced by a national selection process called SISU (*Sistema de Seleção Unificada*). Since SISU was adopted, UFBA stopped collecting information about students' grades and socioeconomic characteristics; therefore, we work only with students who did the vestibular up to 2013.

[9] The University of Brasilia (UNB) was the first university to create an AA policy in Brazil. It reserved 20% of slots for Black and mixed-race students.

[10] However, this is not an expected behavior. The vestibular selection process is one of the main causes of mental health problems among young people in Brazil. See, for example, https://www.hospitaloswaldocruz.org.br/imprensa/releases/pressao-pre-vestibular-pode-afetar-saude-e-bem-estar-dos-estudantes/

that teachers can tailor their teaching to different groups, increasing students' learning (Betts, 2011). However, interviews with professors and university employees suggested that this is not the case. Some UFBA majors decided to have two classes starting at different times since they lacked the resources to allocate the total number of students selected through the entrance exam to the semester beginning in March.

The implementation of AA at UFBA holds particular significance due to the university's location in the state of Bahia. According to PNAD data, 83% of the population is Black or mixed race, the highest percentage among all Brazilian states. Because of the high share of the Black and mixed-race population in Bahia state, it becomes challenging for professors to discriminate according to students' observable characteristics. Additionally, because the type of high school attended creates the AA status, and is information not available to professors, they do not know which students are AA enrollees.[11] Lastly, it is important to highlight that all students within the same major at UFBA have equal access to college infrastructure, curriculum, and professors.

# 3 Data

In this study we use two rich administrative data sets that have been matched using a unique identifier known as the CPF (*Cadastro de Pessoas Físicas*), a nine-digit individual taxpayer identification number.

**UFBA: academic performance**. The administrative records comprise two data sets. The first includes the basic socioeconomic questionnaire administered during the entrance exam day and contains the entrance exam grades of all enrolled students. However, it does not indicate whether the students were approved for the first or second semester. The second data set contains the academic history of the students who enrolled at UFBA, including their starting semester. The administrative records of UFBA provide detailed information on students' grades in each course and failures from 2006 to 2017, and whether the student graduated up to 2021. We use this information to calculate the GPA at the beginning (first-semester and first-year GPA) and end of the major (GPA in the tenth semester), dropouts, and failures measured by the number of courses the students failed between the first and tenth semesters.[12]

The sample has 12,494 students – 4,854 AA and 7,640 regular students – who completed the registration process to enroll at the university. The data also allows us to identify

[11] After some interviews with professors and former students we concluded that it is easy to guess which students are not AA students. However, they face difficulties in distinguishing whether a Black or mixed-race student is an AA student.

[12] To measure the performance at the end of the major, we used the GPA and failures in the tenth semester instead of when they completed the major because this information would be missing for students that drop out.

whether they enrolled in the first or second semester, which is based on the student's entrance exam rank. These two data sets are merged using the unique CPF information available only for UFBA-enrolled students. Importantly, UFBA did not have a minimum score to admit students. The only requirement was that they score above zero in the dissertation and the second phase exams. Table A.2 shows the average share of A.A student per entry class and major.

It is also important to mention that Brazilian federal universities had no centralized data at the time.[13] Therefore, each university's information must be collected directly from each institution using a formal request that respects Brazil's individual data protection law. Each university's legal department has the right to approve making the data available. Only UFBA introduced such a large AA policy in 2005.[14] These facts explain the use of data from a single university in this study.

Regarding GPA, we calculate three indicators: (1) the weighted-average grade in the first semester, where the grades are weighted by the total hours in each subject; (2) the weighted-average grade in the first year at the university; and (3) the weighted-average grade in the tenth semester as a proxy for the GPA at the end of the course. Then, for each major–year we standardize these three variables so that they have zero means and their standard deviation is equal to 1. With these three new variables we intend to measure student performance at different times over the major. These measures allow us to compare the students as freshman and bachelor candidates. Failures are the number of courses the student was not approved. Dropout is a dummy equal to 1 if the student did not graduate.

**RAIS: employment and income**. The labor market outcomes stem from RAIS (*Relação Anual de Informações Sociais*), a matched employee–employer data set maintained by Brazil's Ministry of Economy. The RAIS data set contains information on each formal worker at each plant in Brazil, as all formal establishments in Brazil are legally obligated to submit information to RAIS. We use yearly information for the period 2006–2020, which allows us to follow all students from one to eight years after enrollment. We construct a set of dummies of formal employment, which equals one if the individual is formally employed in December of each year and 0 otherwise. We also collect information on earnings in December of each year. For a few individuals who have two or more jobs, we considered only the job with higher earnings. Note that RAIS has information only for workers in the formal labor sector.

[13] After SISU implementation in 2012, the Ministry of Education developed a centralized data center that can be accessed in a safe room conditional on formal requests.

[14] For a detailed description of the AA policies implemented in Brazilian higher education institutions, see Vieira (2019).

## 3.1 Descriptive statistics

Table 1 shows the descriptive statistics for the variables used in the study. Columns 1–4 present the statistics for all AA students, while columns 5–8 present only the statistics for those around the cutoff in the regression discontinuity design (RDD). The bandwidth used in this table for all characteristics to define students around the cutoff is the one used for the tenth-semester GPA.[15] This table shows that for AA students the explanatory covariates are well balanced. However, for the vestibular score and the outcome variables it is possible to see that they are balanced only around the cutoff.

The entry exam score average is higher in the first class, showing that students in these classes have peers with better skills. The number also suggests that regular students need a higher score in the vestibular exam to enter the university. The average score of regular students in the second class is higher than the average score of the AA students in the first class. This is simple evidence that AA plays an important role in providing access to UFBA for disadvantaged students. Figure A.2 complements the evidence by showing the distribution of vestibular scores for AA and regular students.

**Differences in field of study entrance exam score.** An important part of the study is the results splitting the sample into different fields. As explained in the introduction, there are many reasons to expect that peer and ranking effects may differ between the three fields of study. Figure 1 shows the entrance exam score for AA students. The figure shows that students in technology fields have higher scores, followed by health sciences[16] and social sciences. Appendix Table A.3 details the majors in each field. The difference between the entrance exam score by fields is even more pronounced when looking at the regular students' distributions in appendix figure A.1.

**Students' position in the rank distribution.** As highlighted before, we focus on AA students and provide separate estimates for regular students. The two groups are not subject to the same mechanism, and the results are not comparable. First, each group has its own rank, which implies a different relative rank than the other group. For example, Table A.1 shows that the best regular students in the second class are, on average, at the top of that class, while the best AA students are around the median students in terms of the entrance exam score. This table also shows that the worst-ranked AA student is, on average, the last of the first class, while the worst regular student is not. We also show in Appendix Figure A.2 that the distribution of entrance exam scores among regular students between the first and second semesters is more similar than the distribution of entrance

[15] Using the optional bandwidth in the RDD analysis for other outcome measures does not change the interpretation of the results for this table.

[16] The second peak in health sciences is explained by students enrolled in medicine.

exams of AA students. This evidence is also supported by a lower value of the D-statistic of the Kolmogorov–Smirnov test for comparing distributions.

# 4 Empirical strategy

## 4.1 Method

We use a sharp discontinuity regression to estimate the effect of class assignment on educational and labor market outcomes. We assume that, for each course and AA status, the last student joining the first class is similar to the first student of the second class. These students have the same classes, teachers, and college infrastructure. The differences between them are their ordinal ranks and their classmates. The last student in the first class had the lower ordinal rank and better classmates in terms of entrance exam scores. In all estimations, the sample is split between AA and non-AA students.

The main underlying assumption is that students cannot manipulate their entrance exam scores around the cutoff. Because each student needs to choose the major before the exam, they do not know the minimum score to be approved in the university or in the first class of each major. The university does not fix the minimum score to be admitted to each major, and there is no minimum threshold – it is defined by the number of slots in each major and the entrance exam ranking in the specific year. Students thus cannot predict the cutoff because the minimum score also depends on other students' efforts. In addition, Appendix B shows the coefficients and graphs of the manipulation test proposed by Cattaneo et al. (2020). The results suggest that there is no evidence of manipulation around the cutoff.

An important fact is that the university does not override the class assignment rule after students enroll and classes start. We can simplify the explanation with a hypothetical example. Suppose the economics major has 100 slots yearly and divides students into two classes. For simplicity, assume there is no AA policy. If 150 candidates apply to the economics major, the candidates ranked 1–50 will start the course in March and candidates 51–100 will start in September. Suppose the university releases the results in January, students enroll in February, and classes start in March. Before March, students 10 and 11 decided not to enroll. UFBA automatically moves students 51 and 52 to the class starting in March and invites students 101 and 102 to enroll in September. Therefore, the last student in the first class is student 52, and we will use her entrance exam score as the cutoff point. After the classes start in March, if students 30 and 31 drop out no one else is invited to fill their slots.

Therefore, in our setting, the cutoff is the score of the last student $i$ of the group $g$ that entered during the first class (beginning in March) in each cohort $c$ – with $g = 1$

if AA beneficiary and 0 otherwise – for major $m$. As in Francis-Tan and Tannuri-Pianto (2018) and Zimmerman (2019), the running variable is the normalized score calculated as $NS_{igmc} = \frac{(S_{igmc} - T_{gmc})}{SD_{gmc}}$, where $S_{igmc}$ is the score of student $i$ in the admission process; $T_{gmt}$ is the threshold or cutoff – the score of the last student of group $g$ in the first class of major $m$ and cohort $c$ – and $SD_{gmc}$ is the standard deviation of the score for group $g$, major $m$, and cohort $c$. This standardizing and pooling approach yields consistent estimates for the local average treatment effect (LATE) (Cattaneo et al., 2016; Duryea et al., 2023). In particular, we estimate the non-parametric local linear model of the form:

$$Y_{igmc} = \beta_0 + \beta_1 NS_{igmc} + \beta_2 A_{igmc} + \beta_3 NS_{igmc} \times A_{igmc} + \beta_4 female_{igmc} + \gamma_m + \rho_c + \varepsilon_{igmc} \quad (1)$$

where $Y_{igmc}$ is an outcome for student $i$ in group $g$, major $m$, and cohort $c$. $A_{igmc} = 1\{NS_{igmc} \geq 0\}$, with positive values meaning those students with a normalized score equal to or above the minimum to be in the first class. We add the major fixed effect, $\gamma_m$, and the cohort in which they started the major, $\rho_c$, to account for possible cohort-related unobserved factors.

We included a female dummy to control for differences in academic achievement by gender (Ribas et al., 2020). Calonico et al. (2019) and Frölich and Huber (2019) showed that including pre-treatment covariates increases the precision of the estimates. We use a triangular kernel function for weighting the observations, and we apply the estimator and the optimal bandwidth selection proposed by Calonico et al. (2014). $\beta_2$ is the parameter of interest, which identifies the class composition effect on students' outcomes. A $\beta_2 < 0$, would mean that the ranking effect is more important or that the peer effects are not very important in this setting.[17] Unfortunately, the RDD design does not allow us to disentangle the effect of peers and ranking. To do that, we conduct an exercise similar to Ribas et al. (2020). First, we calculated the absolute peer quality by summing up the entrance exam score of each individual in the class. Then, we calculated the difference between classes in the same major and year. Finally, we split the sample below and above the median of absolute peer quality difference. The results are presented in section 5.3.

A potential issue with our approach is the endogeneity of the allocation cutoff to the first class. The compliance to accept enrolling in the first class can be higher than the compliance to enroll in the second class. The solution presented by Chaisemartin and Behaghel (2020) to re-balance is to drop the last first-class enrolled student for each major and year.[18] Therefore, appendix tables A.4 and A.5 present the estimates for the main

[17] An alternative interpretation for the peer effect is that there could be negative peer pressure of having better peers(Booij et al., 2016; Bursztyn et al., 2019).

[18] See Duryea et al. (2023) for an application of Chaisemartin and Behaghel (2020) in a RDD context. Note, however, that in our case, all students accepted to enroll. In the case of the mentioned article, they apply this procedure because they observe students who are offered a slot for each class and whether they decide to enroll. In addition, because they have a fuzzy setting, their empirical strategy is more similar

results after doing this procedure. The main results do not change.

# 5 Results

We divide the results into four parts. First, we provide the results on academic performance (GPA, dropout, and failure) and the heterogeneity analyses by social sciences, technology, and health sciences. Second, we perform the analyses for employment and wages. Third, we investigate whether peer quality or ranking explains the results. Fourth, we present the results for regular students.

Figure 2 presents the regression discontinuity plots for AA students for the main outcomes. This figure suggests that class allocation negatively affects the last students of the first class in most of the outcomes. Appendix Figures A.3–A.5 display the same graphs for each of the main UFBA fields. These figures suggest a negative impact of being the last AA student of the first class for the technology and health majors.

## 5.1 Academic performance

Table 2 presents the main baseline results of the impact of being among the last students of the first class. Column 2 shows that for AA students enrolled in technology majors, being among the last of the first class reduces their average grade in the first year by –0.434 standard deviations. Column 3 shows that the effect persists at the end of the major, with those students having an accumulated GPA in the 10th semester reduced by -0.367 standard deviations. Column 5 shows that the last of the first class fail 3.97 more courses than students enrolled in the second class.

For the last AA students enrolled in the first class of health majors, column 2 shows a lower GPA in the first year, and column 3 shows a lower GPA in the tenth semester. Column 5 shows they failed in 4.14 more courses. A different scenario is observed for the last first-class AA students enrolled in social sciences majors. The only statistically significant coefficient is a negative effect on dropouts.

Table 2 shows that being among the last AA students in the first class reduces students' academic achievement for technology and health majors. Our central hypothesis for a negative effect on technology and health fields' academic outputs is that AA students have worse math and science-related abilities before college, which translates into learning problems during the major – mainly at the beginning of it.[19] We investigated this hypothesis using a sample of enrolled students in different courses in the first semester.

First, we selected the courses in the highest 1% percentile of class size. However, the 12 selected courses are unbalanced between the main fields. Therefore, we selected the

to using a randomized waiting list as an instrument.

[19]This hypothesis corroborates Oliveira et al.'s (2023) findings using data for the same university.

four courses with the most enrolled students per field.[20] The results presented in Table 3 support the assumption that AA students have a lower background in math and science skills. The effects for the calculus courses are much higher than for other courses and much stronger than those observed for the three GPA measures presented in Table 2. More specifically, being among the last AA students in the first class reduces their grades by −1.25 s.d., almost three times the effect for the aggregated GPA. The table also shows an increase in the probability of failing the course by 0.65 percentage points. The last first-class AA students enrolled in health sciences also observed a reduction in their grades in microbiology and anatomy courses, with the microbiology effect size almost two times higher than the observed effect on health majors' GPAs.

The analysis at the course level also enables us to address a potential concern of the study. If professors changing the level of the classes to adapt to students' performance is the main cause of academic achievement, we would expect the results from Table 3 to be similar to the first-semester GPA results in column 1 of Table 2, which does not occur. In addition, the estimations in Table 3 also include professor fixed effects. Appendix Table A.7 shows that removing the professor fixed effects does not significantly change the results.

## 5.2 Employment and income

In the last subsection we showed that the last AA students placed among high-ability peers do worse in technology and health majors. The opposite happens to AA students enrolled in social sciences majors. A natural question is whether these results would translate into labor market effects. We re-estimate equation 1 forlabor market outcomes. Because the last cohort enrolled in 2012, we are able to follow students for 1–8 years after enrollment.

We follow Roux and Riehl.'s (2022) approach to defining our dependent variables. The first is an indicator of having formal employment between 1–8 years after enrollment. The second is the number of appearances in the labor market between 1–8 years after enrollment. The third is the total inflation-corrected earnings 1–8 years after enrollment, including zeroes for individuals without income in any period. The fourth is the log of the mean of the inflation-corrected earnings 1–8 years after enrollment, which does not include zeroes for those without formal employment. Table 4 shows the results. Unlike the previous analyses, in this section, there are no potential unobserved effects because of professors' behavior or students' decision to enroll at UFBA.

Table 4 shows the results for AA students. The results indicate no impact on the probability of having a job for the last student in the first class. We also show no impacts on the number of jobs and earnings. This result aligns with the findings of Ribas et al. (2020) and Roux and Riehl. (2022).

[20] The total number of courses in the sample is 1123.

### 5.3 Ranking effect versus peer quality

So far we have shown that being among the last AA students in the first class reduces students' academic achievement for technology and health majors, but has no effect on labor market outcomes. Therefore, our results suggest that class allocation can potentially affect students' academic outcomes, but the direction of the effect depends on the student's major field. To further understand whether ranking effect or peer quality explains our results, we conduct an exercise similar to Ribas et al. (2020). We create a measure of the difference in absolute peer quality difference between classes and split the sample above and below the median of this measure. The absolute peer quality is the sum of the entrance exam scores of all enrolled students in the same major and year. As Ribas et al. (2020) explains, the smaller the difference in peer quality, the weaker the contribution of peer quality to the class composition effect or the more important the ranking effect. The results for AA students, displayed in Table 5, are quite heterogeneous and somewhat different from their work. In general, for all students, the rank effect seems to explain the negative effect on the first semester GPA and the positive effect on failures. Next, we interpret the results separately for each field to ease understanding.

**Technology majors.** When the peer quality difference is small, the results are positive and statistically significant for the number of failures and surprisingly positive for graduation on time. When the peer quality difference increases, we observe a stronger negative effect for technology students in the final GPA, failures, and dropouts. The positive impact on failures is higher. This suggests that peer composition is important for AA students' learning. More importantly, it indicates that better peers can be harmful (Bursztyn and Jensen, 2015; Bursztyn et al., 2019; Booij et al., 2016). It also corroborates the evidence that peer effects can have heterogeneous impacts depending on the student rank, being unfavorable when students are at the bottom of the distribution (Booij et al., 2016). We do not find any effects on labour market outcomes.

**Health majors.** For the last student enrolled in the health majors' first class, the negative GPA effects seem to be driven by students enrolled in classes below the median of the peer quality difference. However, surprisingly, a lower GPA does not translate into higher dropouts and failures. When the peer quality is higher, there is no effect on GPA, but there is a positive and stronger effect on failures. There is a positive effect on the probability of finding a job when the difference in peer quality is low.

**Social sciences Majors.** For students enrolled in social sciences majors, the results suggest the lower the peer quality difference, the lower the dropout and higher the graduation on time. Surprisingly, the effect of failures turns positive.

## 5.4 Results for regular students

In this section we will replicate the previous analysis for the regular students. It is important to highlight that AA and regular students placed at the top and at the bottom of their relative rank have different positions in the absolute class rank. Appendix Table A.1 makes this point very clear. The best regular student in the second class is the best student in the class, while the worst regular student in the first class is not the worst in the class. Therefore, our setting is different from that of Ribas et al. (2020) and Roux and Riehl. (2022).

Table 6 shows that our findings also diverge from those aforementioned studies, which find negative effects on students' academic performance in a context without a diversity of student types. Being the last regular student in the first class, but not the last of the class, proves more advantageous than being the first student in the second class. These last students exhibit a 0.22 s.d. higher first-semester GPA, 0.23 s.d. higher first-year GPA, and 0.22 s.d. final-year GPA. This result is primarily driven by students enrolled in social sciences majors. They also have a lower probability of dropout and a higher probability of graduating on time. Reinforcing that our specific context differs, the last student in the first class is not the overall last student in the class. Consequently, the mechanisms through which peers and ranking affect regular students diverge from those discussed in the published articles. Appendix Table A.8 shows no impacts of class allocation on labor market outcomes.

Even though the results between groups are not exactly comparable, we can learn something from the regular students finding. First, the positive results for both groups weaken the argument that an adverse effect is expected because second-class students have six months to prepare before joining the university. Second, it also weakens the argument that professors are less rigorous with the second-class exams. We will discuss these two situations in more detail in the following section. Third, although regular and AA students have different relative rankings, they have the same peers. Therefore, it suggests that peers may have heterogeneous effects on learning, depending on the student type.

Finally, to make a direct comparison with the previous studies, we used the data for 2003, the only available data for a period before the AA policy.[21] Using the year 2003 provides a direct comparison to Ribas et al. (2020) and Roux and Riehl. (2022) because all students are regular students. Appendix Table A.9 shows no impact of the class allocation on academic outcomes, such as in the aforementioned studies.

[21] The data for 2004 is also available, but because the four-month strike in the second half of 2004 most of the students admitted to start in September 2004 only enrolled in February 2005.

# 6 Potential threats to the identification

The empirical strategy faces four main concerns. Unfortunately, we cannot rule them out with our data set and identification strategy, but we provide some evidence that helps to understand whether there is some potential bias in our results. Note that similar threats are also discussed by Ribas et al. (2020) and Roux and Riehl. (2022), and they cannot provide a full explanation.

The first threat is that the classes in the first and second semesters can have different sizes, which adds a potential omitted bias to the estimation. Table A.3 show that the average difference in class size between the first and the second semester is only 1.1. This is not a reasonable value to support the assumption that class size can explain the results. From the 24 majors in the sample, only management has an average difference between the first and second semesters greater than five students.

The second concern is that professors with more experience at UFBA could know that students in the first class have better vestibular scores. They could reduce the level of the exams for the second class, which could create a downward bias in the estimated coefficients. Unfortunately, it is not possible to observe and measure professors' behavior. However, professors can identify which class they are teaching more clearly in the first year. Therefore, if there are teacher effects, it would probably be more pronounced for the GPA in the first year. Besides, we assume that if professors', and not students', characteristics drive the results, we should expect a negative result for all fields and for AA and regular students. However, this is not what happens. While AA students in technology and health have negative effects, students in social sciences have a positive effect. Regular students also observed a positive effect of being allocated to the first class.

One may also think that grading standards may affect the results. However, there is no grading policy at UFBA. Each professor has the freedom to decide how to evaluate their students. Unfortunately, this is another mechanism that we cannot rule out. Notwithstanding, to provide some evidence on the extent to which this may happen, we estimate equation (2) for first-semester courses using the same sample of the main analysis. First semester courses are the moment when instructors know exactly which students are enrolled in the first or second semester. The estimating equation is:

$$\begin{aligned} Y_{imcjp} = \alpha + \beta_0 AAStudent_{imcjp} + \beta_1 SecondClass_{imcjp} + \\ \beta_2 Score_{imcjp} + \beta_3 AAStudent_{imcjp} \times SecondClass_{imcjp} + \\ \nu_c + \psi_m + \nu_{jp} + \epsilon_{imcjp} \end{aligned} \tag{2}$$

where $Y_{imcjp}$ is the grade of student $i$ from major $m$ and cohort $c$ in course $j$ taught by instructor $p$, $SecondClass_{imcjp}$ is an indicator of whether the student enrolled in the second class, $Score$ is the student score in the entrance examination, $\nu_c$ is a set of cohort fixed effects, $\psi_m$ is a set of major fixed effects, and $\nu_{jp}$ is a set of instructor-by-course fixed effects. Hence, $\beta_0$ implicitly reveals how instructors who taught the same first-semester courses behaved with first and second classes.

Appendix Table A.10 shows that even after controlling for entry exam score, the second-class coefficient is negative, which means that, if anything, professors who teach the same first-semester course to first and second classes discriminate against the latter, contrary to what we should expect from grading-on-a-curve behavior.

A third potential concern to our strategy is that students who were approved for the second class could decide not to enroll. High-skill regular students come from families with higher incomes, and some of them can choose not to study at UFBA and to go to a high-level private university if they go to the second class, increasing the dropout rates. This behavior should not expected from AA students because they come from poor families that cannot afford a flagship private university. However, both groups of students approved for the second semester may decide to go to the labor market between March and August.[22] If this happens they may choose not to enroll at UFBA. This assumption may be stronger for AA students at the bottom of the second-class distribution because they probably come from the poorest families.

Another assumption could be that students approved for the second semester will use this time for preparation. After meetings with university staff, students, and former students, we strongly believe that this hypothesis can be discarded.[23]

The fourth potential concern is that students approved for the second semester have a lower probability of enrollment. Unfortunately, the lack of data on admissions does not allow us to test this assumption. However, it is important to highlight that if this is true, and the student who does not enroll is at the bottom of the distribution of the second class, it will not affect our RDD estimates because they are outside the bandwidth range. If it is the students at the top, this would suggest that the negative coefficients estimated for AA students in Table 2 can be interpreted as lower-bound effects.

[22]Unfortunately, our data set only allows us to verify if students have formal employment in December. The other information available is whether they have another job before December of each year, but there is no information about the month if they had another job.

[23]UFBA is located in one of the most beautiful cities in Brazil, also known for its strong cultural and nightlife. Most of the information we collect indicates that students use this free time between March and August for leisure. The vestibular year is known as a very stressful time for young Brazilians. See, for example, https://www.hospitaloswaldocruz.org.br/imprensa/releases/pressao-pre-vestibular-pode-afetar-saude-e-bem-estar-dos-estudantes/. Furthermore, to prepare for the major during these six months prospective students would need to predict the future professors, the level of the course, and the course syllabus to decide what to study. We do not think these are reasonable assumptions.

## 7 Conclusion

This paper exploits the class allocation rule to study whether the class environment improves or harms academic performance and labor market outcomes. The main results suggest that being the last among the better students is harmful to AA students enrolled in technology and health fields. However, the negative effects happen only for the academic outcomes.

Our findings show that the background in math and science skills and the class composition matter in explaining disadvantaged students' lower performance in college. However, the potential peers and ranking effects do not affect the labor market outcomes. Therefore, policies that help disadvantaged students close the educational and opportunity gap may improve welfare and boost the effects of AA policies.

## References


Anelli, M., and G. Peri (2017). 'The Effects of High School Peers' Gender on College Major, College Performance and Income'. *The Economic Journal*, 129(618): 553–602. https://doi.org/10.1111/ecoj.12556

Arcidiacono, P. (2005). 'Affirmative Action in Higher Education: How Do Admission and Financial Aid Rules Affect Future Earnings?' *Econometrica*, 73(5): 1477–524. https://doi.org/10.1111/j.1468-0262.2005.00627.x

Arcidiacono, P., and M. Lovenheim (2016). 'Affirmative Action and the Quality–Fit Trade-Off'. *Journal of Economic Literature*, 54(1): 3–51. https://doi.org/10.1257/jel.54.1.3

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): 1495–521. https://doi.org/10.1257/aer.20140783

J. R. Betts (2011). The economics of tracking in education. *Handbook of the economics of education, vol 3.*

Booij, A.S., E. Leuven, and H. Oosterbeek (2016). 'Ability Peer Effects in University: Evidence from a Randomized Experiment'. *Review of Economic Studies*, 84(2): 547–78. https://doi.org/10.1093/restud/rdw045

Bleemer, Z.(2022). Affirmative Action, Mismatch, and Economic Mobility after California's Proposition 209. *Quarterly Journal of Economics*, 137(1): 115–160. https://doi.org/10.1093/qje/qjab027

Bursztyn, L., G. Egorov, and R. Jensen (2019). 'Cool to be Smart or Smart to be Cool? Understanding Peer Pressure in Education'. *Review of Economic Studies*, 86(4): 1487–526. https://doi.org/10.1093/restud/rdy026

Bursztyn, L., and R. Jensen (2015). 'How Does Peer Pressure Affect Educational Investments?' *Quarterly Journal of Economics*, 130(3): 1329–67. https://doi.org/10.1093/qje/qjv021

Calonico, S., M.D. Cattaneo, and R. Titiunik (2014). 'Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs'. *Econometrica*, 82(6): 2295–326. https://doi.org/10.3982/ECTA11757

Calonico, S., M.D. Cattaneo, M.H. Farrell, and R. Titiunik (2019). 'Regression Discontinuity Designs Using Covariates'. *Review of Economics and Statistics*, 101(3): 442–51. https://doi.org/10.1162/rest_a_00760

D. Card, P. Giuliano. 2016. 'Can Tracking Raise the Test Scores of High-Ability Minority Students¿ *American Economic Review, 106* 10.1257/aer.20150484

Cattaneo, M. D. and Keele, L. and Titiunik, R. and Vazquez-Bare, G. (2016). 'Interpreting Regression Discontinuity Designs with Multiple Cutoffs'. *The Journal of Politics*, 78(4): 1229-1248. 10.1086/686802

Cattaneo, M.D., M. Jansson, and X. Ma (2020). 'Simple Local Polynomial Density Estimators'. *Journal of the American Statistical Association*, 115(531): 1449–55. https://doi.org/10.1080/01621459.2019.1635480

de Chaisemartin, Clément and Behaghel, Luc (2020). 'Waiting lists, non-takers, non compliance, instrumental variable, local average treatment effect, randomized controlled trials'. *Econometrica*, 88(4): 1453-1477. https://doi.org/10.3982/ECTA16032

Dale, S.B., and A.B. Krueger (2014). 'Estimating the Effects of College Characteristics Over the Career Using Administrative Earnings Data'. *Journal of Human Resources*, 49(2): 323–58. https://doi.org/10.3368/jhr.49.2.323

Dasgupta, U., S. Mani, S. Sharma, and S. Singhal (2020). 'Effects of Peers and Rank on Cognition, Preferences, and Personality'. *Review of Economics and Statistics*. https://doi.org/10.1162/rest_a_00966

Duflo, Esther and Dupas, Pascaline and Kremer, Michael (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in *American Economic Review*

Eisenkopf, G. (2010). 'Peer Effects, Motivation, and Learning'. *Economics of Education Review*, 29(3): 364–74. https://doi.org/10.1016/j.econedurev.2009.08.005

Elsner, B., and I. Isphording (2016). 'A Big Fish in a Small Pond: Ability Rank and Human Capital Investment'. *Journal of Labor Economics*, 35. https://doi.org/10.1086/690714

Elsner, B., I. Isphording, and U. Zölitz (2021). 'Achievement Rank Affects Performance and Major Choices in College'. *Economic Journal*. https://doi.org/10.1093/ej/ueab034

Francis-Tan, A. and M. Tannuri-Pianto (2018). 'Black Movement: Using Discontinuities in Admissions to Study the Effects of College Quality and Affirmative Action'. *Journal of Development Economics*, 135: 97–116. https://doi.org/10.1016/j.jdeveco.2018.06.017

Frölich, M., and M. Huber (2019). 'Including Covariates in the Regression Discontinuity Design'. *Journal of Business & Economic Statistics*, 37(4): 736–48. https://doi.org/10.1080/07350015.2017.1421544

Hoxby, C. (2000). 'Peer Effects in the Classroom: Learning from Gender and Race Variation'. Working Paper 7867. Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w7867

Lepine, A., Estevan, F. (2021). 'Do ability peer effects matter for academic and labor market outcomes?'. *Labour Economics. 71.* https://doi.org/10.1016/j.labeco.2021.102022

Marsh, H. W., Parker, W. (1984). 'Determinants of student self-concept: is it better to be a relatively large fish in a small pond even if you don't learn to swim as well?'. *J. Pers. Soc. Psychol. 47, 213–231.* https://doi.10.1037/0022-3514.47.1.213

Murphy, R., and F. Weinhardt (2020). 'Top of the Class: The Importance of Ordinal Rank'. *Review of Economic Studies*, 87(6): 2777–826. https://doi.org/10.1093/restud/rdaa020

Oliveira, R., Santos, A., Severnini, E (2023). Bridging the Gap: Mismatch Effects and Catch-Up Dynamics in a Brazilian College Affirmative Action *National Bureau of Economic Research (NBER).* Working paper number 31403. https://doi.org/10.3386/w31403

Rao, G. (2019). 'Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools'. *American Economic Review*, 109(3): 774–809. https://doi.org/10.1257/aer.20180044

Ribas, R.P., B. Sampaio, and G. Trevisan (2020). 'Can Better Peers Signal Less Success? The Effect of Class Assignment on Career Investment. *labor Economics, 64*. https://doi.org/10.1016/j.labeco.2020.101835

N. Roux., E. Riehl (2022). 'Do college students benefit from placement into higher-achieving classes? *Journal of Public Economics, Forthcoming.* .

Sund, K. (2009). 'Estimating Peer Effects in Swedish High School Using School, Teacher, and Student Fixed Effects'. *Economics of Education Review*, 28(3): 329–36. https://doi.org/10.1016/j.econedurev.2008.04.003

Vieira, R., Arends-Kuenning, M. (2019). 'Affirmative action in Brazilian universities: Effects on the enrollment of targeted groups'. *Economics of Education Review*, 73. https://doi.org/10.1016/j.econedurev.2019.101931

Zeidner, M., and E.J. Schleyer (1999). 'The Big-Fish–Little-Pond Effect for Academic Self-Concept, Test Anxiety, and School Grades in Gifted Children'. *Contemporary Educational Psychology*, 24(4): 305–29. https://doi.org/10.1006/ceps.1998.0985

Zimmerman, S. (2019). 'Elite Colleges and Upward Mobility to Top Jobs and Top Incomes'. *American Economic Review*, 109(1): 1–47. 10.1257/aer.20171019

Duryea, S. and Ribas, R. and Suzanne, D. and Ribas, R. and Sampaio, B. and Sampaio, G. and Trevisan, G. 'Who benefits from tuition-free, top-quality universities? Evidence from Brazil'. *Economics of Education Review*, 95. https://doi.org/10.1016/j.econedurev.2023.102423

# Tables and Figures

## Table 1: Descriptive Statistics For Affirmative Action Students

| | *All students* | | | | *Students around the cutoff* | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) First Class | (2) Second Class | (3) Mean Difference | (4) Obs | (5) First Class | (6) Second Class | (7) Mean Difference | (8) Obs |
| Age | 21.4 | 21.56 | -0.15 | 4853 | 21.52 | 21.65 | -0.13 | 2547 |
| | (5.64) | (5.82) | [0.35] | | (5.69) | (5.95) | [0.58] | |
| Male | 0.48 | 0.42 | 0.05 | 4854 | 0.46 | 0.44 | 0.02 | 2548 |
| | (0.5) | (0.49) | [0.00] | | (0.5) | (0.5) | [0.45] | |
| Black or Mixed Race | 0.97 | 0.96 | 0.01 | 4014 | 0.97 | 0.96 | 0.01 | 2102 |
| | (0.17) | (0.2) | [0.01] | | (0.16) | (0.2) | [0.09] | |
| Share of Students in STEM | 0.21 | 0.22 | -0.01 | 4854 | 0.22 | 0.2 | 0.02 | 2548 |
| | (0.41) | (0.41) | [0.43] | | (0.41) | (0.4) | [0.33] | |
| Entry Exam Score | 13181.44 | 12066.14 | 1115.3 | 4854 | 12552.44 | 12239.5 | 312.95 | 2548 |
| | (1450) | (1174.29) | [0.00] | | (1132.18) | (1169.04) | [0.00] | |
| Standardized Entry Exam Score | -0.37 | -0.91 | 0.54 | 4854 | -0.68 | -0.83 | 0.15 | 2548 |
| | (0.70) | (0.56) | [0.00] | | (0.55) | (0.56) | [0.00] | |
| Entry Exam Score Transformation | 0.32 | 0.2 | 0.11 | 4854 | 0.25 | 0.22 | 0.03 | 2548 |
| | (0.15) | (0.12) | [0.00] | | (0.12) | (0.12) | [0.00] | |
| 1st Sem. GPA | 6.93 | 6.41 | 0.52 | 4529 | 6.68 | 6.53 | 0.16 | 2411 |
| | (1.58) | (1.56) | [0.00] | | (1.59) | (1.55) | [0.01] | |
| 1st Year GPA | 6.84 | 6.38 | 0.46 | 4660 | 6.62 | 6.48 | 0.14 | 2466 |
| | (1.47) | (1.43) | [0.00] | | (1.4) | (1.43) | [0.01] | |
| 10th Sem. GPA | 6.84 | 6.44 | 0.4 | 4700 | 6.66 | 6.53 | 0.14 | 2490 |
| | (1.47) | (1.42) | [0.00] | | (1.43) | (1.41) | [0.02] | |
| Performance Failures 10th Sem. | 3.52 | 5.31 | -1.8 | 4853 | 4.42 | 4.93 | -0.51 | 2548 |
| | (4.46) | (5.61) | [0.00] | | (4.96) | (5.31) | [0.01] | |
| Absence Failures 10th Sem. | 2.52 | 2.7 | -0.18 | 4853 | 2.78 | 2.6 | 0.18 | 2548 |
| | (4.03) | (4.31) | [0.14] | | (4.28) | (4.26) | [0.3] | |
| Failures 10th Sem. | 6.04 | 8.01 | -1.97 | 4853 | 7.2 | 7.53 | -0.33 | 2548 |
| | (6.8) | (7.61) | [0.00] | | (7.3) | (7.4) | [0.26] | |
| Graduation on Time | 0.6 | 0.58 | 0.02 | 4854 | 0.6 | 0.59 | 0.01 | 2548 |
| | (0.49) | (0.49) | [0.09] | | (0.49) | (0.49) | [0.62] | |
| Dropout | 0.32 | 0.32 | 0.03 | 4854 | 0.31 | 0.31 | 0.01 | 2548 |
| | (0.47) | (0.47) | [0.99] | | (0.46) | (0.46) | [0.98] | |
| Employment After Enrollment (1-8 Years) | 0.68 | 0.66 | 0.02 | 4399 | 0.69 | 0.67 | 0.03 | 2331 |
| | (0.47) | (0.47) | [0.11] | | (0.46) | (0.47) | [0.17] | |
| Employment After Enrollment (5-8 Years) | 0.63 | 0.56 | 0.07 | 4399 | 0.64 | 0.57 | 0.07 | 2331 |
| | (0.48) | (0.5) | [0.00] | | (0.48) | (0.5) | [0.00] | |
| Wages After Enrollment (1-8 Years) | 12300.32 | 9313.54 | 2986.79 | 4399 | 10369.96 | 10566.22 | -196.26 | 2331 |
| | (25510.51) | (24523.38) | [0.00] | | (20745.74) | (26437.68) | [0.85] | |
| Wages After Enrollment (5-8 Years) | 8942.61 | 6440.19 | 2502.42 | 4399 | 7604.25 | 7199.86 | 404.39 | 2331 |
| | (15879.22) | (14805.24) | [0.00] | | (13100.7) | (15780.24) | [0.52] | |
| Wages After Enrollment (1-8 Years) Not Considering Zeros | 19742.02 | 16137.46 | 3604.55 | 2642 | 16591.94 | 17938 | -1346.06 | 1407 |
| | (29962.88) | (30531.59) | [0.00] | | (24200.12) | (32477.56) | [0.4] | |
| Wages After Enrollment (5-8 Years) Not Considering Zero | 15234.3 | 12703.3 | 2531 | 2410 | 13004.37 | 13793.1 | -788.73 | 1276 |
| | (18269.24) | (18786.26) | [0.00] | | (14945.84) | (19654.51) | [0.43] | |

*Notes*: This table reports the average statistics for affirmative action students dividing them based on enrollment in the first or second classes. The first class is the one starting in March, and the second class is the one starting in September. Column 1 shows the average for each of the characteristics for all affirmative action students who enrolled in the first class. Column 2 shows the average for students who enroll in the second class. We then regress each of the characteristics on a dummy indicating if the student enrolled in the first class, column 3 reports the estimate for that dummy. Column 4 shows the number of observations for the characteristic. Columns 5 to 8 are analogous to columns 1 to 4, but include only affirmative action students around the entry exam score cutoff. The bandwidth used for all characteristics to define students around the cutoff is the one used for the 10th semester GPA 10th semester GPA estimation of equation (1). Using the optimal bandwidth of the estimation for the first semester or first-year GPA do not change the results. Standard deviations are reported in parenthesis while standard errors are presented in square brackets. Employment after enrollment is an indicator of having formal employment between 2013 and 2020. Wages after enrollment is the total inflation-corrected earnings between six to thirteen years after enrollment, including zeroes for individuals without income in any period. These two measures follow Roux and Riehl. (2022). The difference in the number of observations between the academic and labor outcomes is due to students that enrolled in more than one major in the sample. For example, student "A" can enroll in economics major in 2007 and in a law major in 2010. The change can be because the student decided to do a second major after finishing the first one, or because of decisions regarding major switching. When a student appears more than one time, we consider the first registry only to match with the labor market outcomes.

## Table 2: Impacts of Class Allocation on Academic Outcomes For Affirmative Action Students

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | 1st Sem. GPA | 1st Year GPA | 10th Sem. GPA | Dropout | 10th Sem. Failures | Graduation on time |
| Everyone | -0.104 | -0.159 | -0.191 | 0.021 | 1.124 | -0.003 |
| | (0.118) | (0.101) | (0.13) | (0.04) | (0.784) | (0.042) |
| Technology Majors | -0.272 | -0.434*** | -0.367*** | 0.099 | 4.335*** | -0.129 |
| | (0.209) | (0.144) | (0.114) | (0.207) | (1.226) | (0.101) |
| Health Majors | -0.263 | -0.261* | -0.414*** | 0.09 | 4.146*** | -0.059 |
| | (0.174) | (0.145) | (0.147) | (0.057) | (1.116) | (0.063) |
| Social Sciences Majors | 0.037 | -0.081 | 0.034 | -0.076* | -0.994 | 0.092 |
| | (0.204) | (0.187) | (0.152) | (0.046) | (1.267) | (0.059) |
| Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Major FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Gender | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Notes*: This table reports the estimated impacts of being allocated to the first class on six outcomes. The hour-weighted average grade in the first semester ("1st Sem GPA"), the hour-weighted average grade in the first year ("1st Year GPA"), the hour-weighted average grade in the first 10 semesters ("10th Sem GPA") – we use this definition instead of the final course GPA to avoid selection into graduation, the total number of failed courses in the first 10 semesters ("Failures"), a dummy which is equal to 1 if the student eventually evaded the major ("Dropout"), and a dummy equal to 1 if the student graduated on time ("Graduation on time"). The estimates refer to the RDD coefficient $\beta_2$ from equation (1). The optimal bandwidth is estimated according to Calonico et al. (2014). The unit of observation is a student-semester-major, and the analysis includes the cohorts of students enrolling at UFBA in the years 2006-2012. The running variable is the entry exam score and the cutoff is the score of the last student enrolled in the first class. Standard errors clustered at the major level are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

## Table 3: Impacts of Class Allocation on Academic Outcomes For Affirmative Action Students

| | Technology | | | | Health | | | | Social Sciences | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| | Introductory Calculus | Linear Algebra | Descriptive I Drawing I | Physics I | MicroBiology | Anatomy | Anatomy II | Histology II | Introduction to Law | Anthropology | Introduction to Philosophy | Political Science |
| Course Grade | -1.254*** | 0.273 | -0.483 | 0.497 | -0.834* | -0.514** | 0.569 | -0.383 | -0.247 | 0.043 | 0.272 | -0.157 |
| | (0.307) | (0.34) | (0.328) | (0.418) | (0.443) | (0.254) | (0.591) | (0.516) | (0.29) | (0.383) | (0.205) | (0.284) |
| Prob(Failure) | 0.651*** | -0.169 | 0.033 | 0.192 | 0.446** | 0.112 | $--^{+}$ | 0.052 | -0.321** | -0.062 | -0.244** | 0.021 |
| | (0.137) | (0.139) | (0.071) | (0.172) | (0.218) | (0.074) | – | (0.072) | (0.153) | (0.063) | (0.095) | (0.084) |
| Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Major FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Gender | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Notes*: This table reports the estimated impacts of being allocated to the first class on two outcomes. Course standardized grades ("Course Grade") and the probability of failing the course ("Prob(Failure)"). The estimates refer to the RDD coefficient $\beta_2$ from equation (1). The optimal bandwidth is estimated according to Calonico et al. (2014). The unit of observation is a student-semester-major, and the analysis includes the cohorts of students enrolling at UFBA in the years 2006-2012. The running variable is the entry exam score and the cutoff is the score of the last student enrolled in the first class. Standard errors clustered at the major level are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level. "+" The number of students that fail in anatomy is lower than 1%, which does not allow running the regression.

**Table 4: Class Allocation Effects On Labor Market Outcomes For Affirmative Action Students**

| | Has a Job (1) | Number of Years With Formal Employment (2) | Total Wage (3) | Log (Mean Wage) (4) |
|---|---|---|---|---|
| Everyone | 0.049 | -0.237 | 336.935 | -0.086 |
| | (0.055) | (0.52) | (1766.833) | (0.098) |
| Technology Majors | -0.064 | -0.005 | 2237.521 | -0.241 |
| | (0.09) | (1.087) | (2191.388) | (0.282) |
| Health Majors | 0.08 | -0.206 | -80.471 | -0.037 |
| | (0.059) | (0.701) | (1232.928) | (0.178) |
| Social Sciences Majors | 0.096 | -0.432 | 2736.15 | -0.109 |
| | (0.1) | (0.775) | (3544.677) | (0.115) |
| Cohort FE | ✓ | ✓ | ✓ | ✓ |
| Major FE | ✓ | ✓ | ✓ | ✓ |
| Gender | ✓ | ✓ | ✓ | ✓ |

*Notes*: This table reports the estimated impacts of being allocated to the first class on four outcomes. A dummy equal to 1 if the student had formal employment between the first and eight years after enrollment ("Has a Job"), the number of years in which the student had formal employment between first and eight years after enrollment ("Number of Years With Formal Employment"), the total inflation-corrected earnings between first and eight years after enrollment ("Total Wage"), which includes zeroes for unemployed individuals, and the log of mean wages between first and eight years after enrollment ("Log (Mean Wage)"), which does not include zeroes for the unemployed. The estimates refer to the RDD coefficient $\beta_2$ from equation (1). The optimal bandwidth is estimated according to Calonico et al. (2014). The unit of observation is a student-semester-major, and the analysis includes the cohorts of students enrolling at UFBA in the years 2006-2012. The running variable is the entry exam score, and the cutoff is the score of the last student enrolled in the first class. Standard errors clustered at the major level are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

## Table 5: Impacts of Class Allocation for AA Students by Difference in Peer Quality

| | (1) 1st Sem. GPA | (2) 1st Year GPA | (3) Final GPA | (4) Dropout | (5) Failures | (6) Graduation on time | (7) Has a Job | (8) Log (Wage) |
|---|---|---|---|---|---|---|---|---|
| **Panel A: Effects for AA Students (Below Median)** | | | | | | | | |
| Everyone | -0.123 | -0.247* | -0.177 | -0.036 | 1.409* | 0.034 | 0.063 | -0.011 |
| | (0.132) | (0.127) | (0.148) | (0.052) | (0.733) | (0.058) | (0.052) | (0.101) |
| Technology Majors | -0.135 | 0.022 | 0.089 | -0.121 | 3.744*** | 0.202*** | -0.022 | 0.131 |
| | (0.469) | (0.209) | (0.157) | (0.297) | (1.184) | (0.061) | (0.152) | (0.121) |
| Health Majors | -0.175 | -0.296* | -0.54*** | 0.103 | 1.995 | -0.081 | 0.147** | 0.021 |
| | (0.138) | (0.151) | (0.161) | (0.068) | (1.388) | (0.077) | (0.071) | (0.21) |
| Social Sciences Majors | 0.018 | -0.084 | 0.021 | -0.15** | 1.827*** | 0.17** | 0.034 | -0.115 |
| | (0.239) | (0.201) | (0.166) | (0.065) | (0.662) | (0.078) | (0.104) | (0.12) |
| **Panel B: Effects for AA Students (Above Median)** | | | | | | | | |
| Everyone | -0.017 | -0.077 | -0.092 | 0.045 | 1.084 | -0.026 | 0.037 | -0.285 |
| | (0.149) | (0.114) | (0.134) | (0.053) | (1.209) | (0.065) | (0.067) | (0.19) |
| Technology Majors | -0.141 | -0.364 | -0.659** | 0.397* | 6.277*** | -0.365** | -0.178 | -0.395 |
| | (0.478) | (0.262) | (0.272) | (0.212) | (1.872) | (0.175) | (0.205) | (0.303) |
| Health Majors | -0.332 | -0.267 | -0.31 | 0.056 | 4.822*** | -0.03 | 0.004 | -0.012 |
| | (0.221) | (0.183) | (0.201) | (0.083) | (1.405) | (0.093) | (0.104) | (0.198) |
| Social Sciences Majors | 0.2 | 0.073 | 0.217 | -0.026 | -5.975 | 0.02 | 0.121 | -0.136 |
| | (0.284) | (0.302) | (0.366) | (0.061) | (3.809) | (0.047) | (0.152) | (0.191) |
| Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Major FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Gender | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

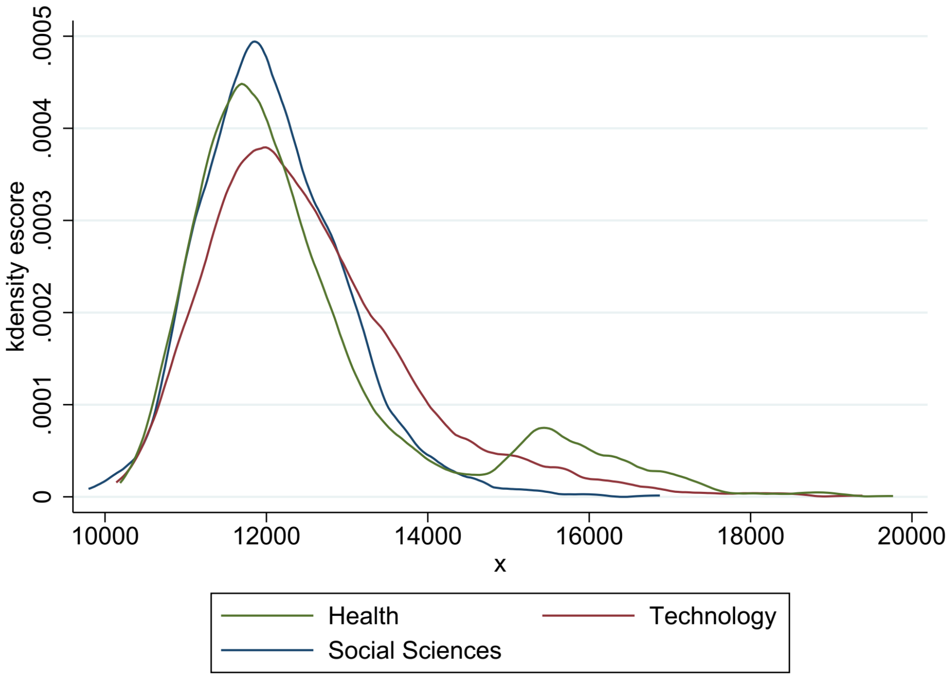
*Notes*: This table reports the estimated impacts of being allocated to the first class on the eight main outcomes when splitting the cohorts for which the difference in peer quality was above or below the median. The hour-weighted average grade in the first semester ("1st Sem GPA"), the hour-weighted average grade in the first year ("1st Year GPA"), the hour-weighted average grade in the first 10 semesters ("10th Sem GPA") – we use this definition instead of the final course GPA to avoid selection into graduation, the total number of failed courses in the first 10 semesters ("Failures"), a dummy which is equal to 1 if the student eventually evaded the major ("Dropout"), a dummy equal to 1 if the student graduated on time ("Graduation on time"), a dummy equal to 1 if the student had formal employment one to eight years after enrollment, ("Has a Job"), the log of mean wages between one to eight years after enrollment ("Log (Mean Wage)"). The estimates refer to the RDD coefficient $\beta_2$ from equation (1). The optimal bandwidth is estimated according to Calonico et al. (2014). Peer quality is constructed as in Ribas et al. (2020). It represents the quality of the peers in the major $m$, entry-year $t$, and class $j$ based on the average of the entry exam score. The unit of observation is a student-semester-major, and the analysis includes the cohorts of students enrolling at UFBA in the years 2006-2012. The running variable is the entry exam score and the cutoff is the score of the last student enrolled in the first class. Standard errors clustered at the major level are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

## Table 6: Impacts of class allocation on academic outcomes of regular students

| | (1) 1st Sem. GPA | (2) 1st Year GPA | (3) 10th Sem. GPA | (4) Dropout | (5) 10th Sem. Failures | (6) Graduation on time |
|---|---|---|---|---|---|---|
| Everyone | 0.217* | 0.226* | 0.218* | -0.052** | 0.266 | 0.074** |
| | (0.114) | (0.123) | (0.117) | (0.024) | (0.516) | (0.031) |
| Technology Majors | -0.003 | -0.12 | 0.071 | 0.001 | 1.884 | 0.094 |
| | (0.12) | (0.113) | (0.128) | (0.069) | (1.239) | (0.062) |
| Health Majors | 0.039 | 0.007 | -0.005 | 0.043 | 0.511 | -0.085 |
| | (0.103) | (0.118) | (0.097) | (0.049) | (0.631) | (0.053) |
| Social Sciences Majors | 0.606** | 0.696** | 0.52* | -0.155*** | -1.174 | 0.208*** |
| | (0.265) | (0.277) | (0.311) | (0.058) | (0.988) | (0.066) |
| Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Major FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Gender | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

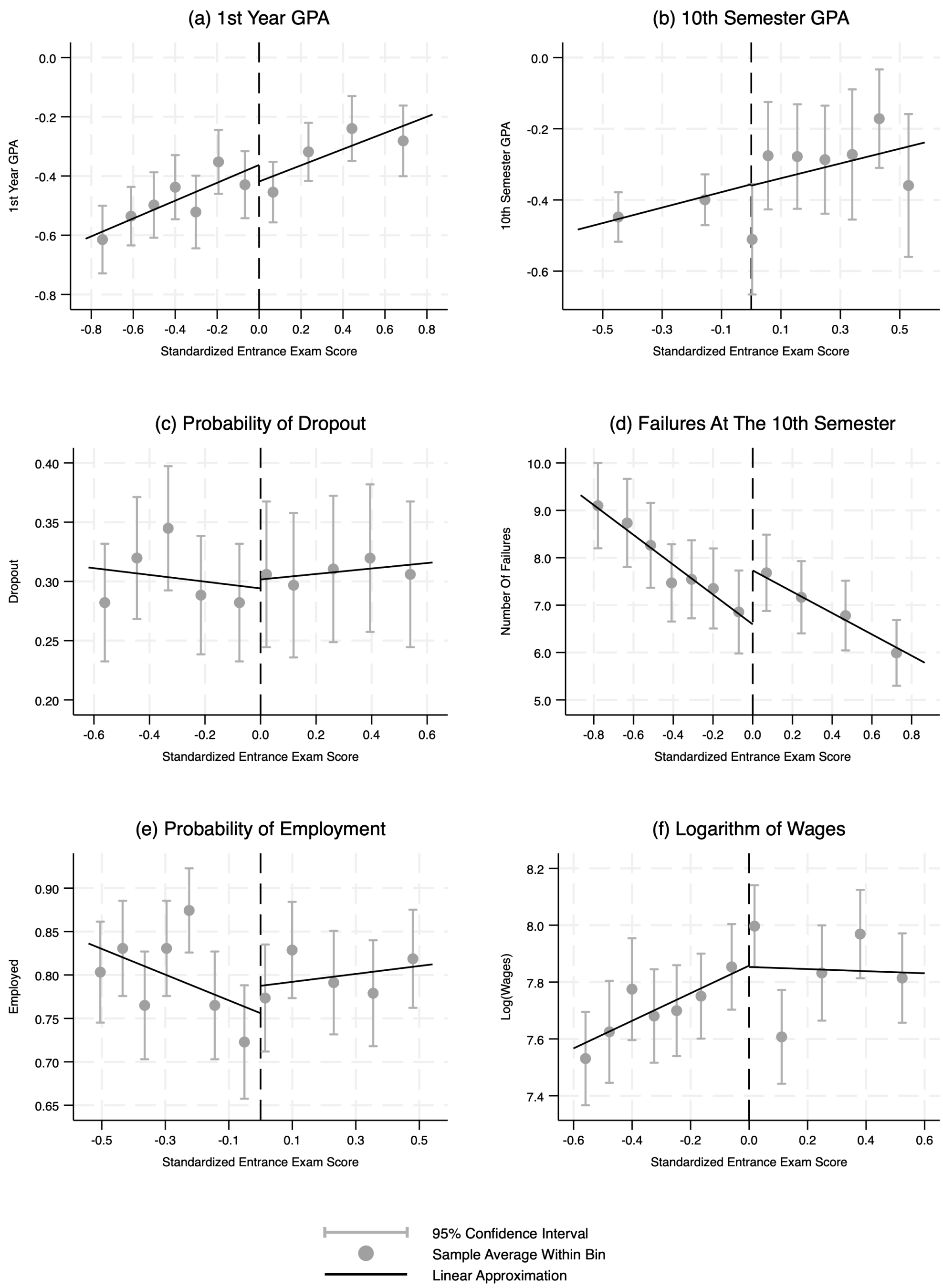
*Notes*: This table reports the estimated impacts of being allocated to the first class on six outcomes. The hour-weighted average grade in the first semester ("1st Sem GPA"), the hour-weighted average grade in the first year ("1st Year GPA"), the hour-weighted average grade in the first 10 semesters ("10th Sem GPA") – we use this definition instead of the final course GPA to avoid selection into graduation, the total number of failed courses in the first 10 semesters ("Failures"), a dummy which is equal to 1 if the student eventually evaded the major ("Dropout"), and a dummy equal to 1 if the student graduated on time ("Graduation on time"). The estimates refer to the RDD coefficient $\beta_2$ from equation (1). The optimal bandwidth is estimated according to Calonico et al. (2014). The unit of observation is a student-semester-major, and the analysis includes the cohorts of students enrolling at UFBA in the years 2006-2012. The running variable is the entry exam score and the cutoff is the score of the last student enrolled in the first class. Standard errors clustered at the major level are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

## Fig. 1: Distribution of The Entrance Exam Score for Affirmative Action Students




*Notes.* This figure shows the distribution of the entrance exam score for affirmative action students. We present the distribution for the main broad fields of study, technology, health and social sciences.

**Fig. 2: Affirmative Action Students' Outcomes Along The Standardized Entrance Exam Score**




*Notes*: This figure displays the linear approximation for (a) 1st year GPA, (b) 10th semester GPA, (c) probability of dropout, (d) cumulative number of failed courses at the 10th semester, (e) probability of formal employment between 2013 and 2020, and (f) logarithm of mean wages between 2013 and 2020 for all affirmative action students coming from estimating (1). The bandwidth used in each graph is estimated according to Calonico et al. (2014). For each graph the observations are binned based on the standardized entry exam score so that each bin has the same number of observations, and the 95% confidence interval is calculated based on the standard deviation for each bin.

**ONLINE APPENDIX - NOT FOR PUBLICATION**

Do disadvantaged students benefit from attending classes with more skilled colleagues? Evidence from a top university in Brazil

Rodrigo Oliveira, Henrique Motte, and Alei Santos

December 2, 2024

# A Appendix: Tables and Figures

**Table A.1: Average Rank of The best- and worst-ranked student by AA status and starting semester**

| | (1) AA Students | (2) Regular Students |
|---|---|---|
| Worst-ranked Student in The First Semester | 0 | 0.38 |
| Best-ranked student in The Second Semester | 0.44 | 0.97 |
| Observations | 163 | 163 |

*Notes*: This table reports the average rank for the worst-ranked student in the first class and the best-ranked student in the second class for AA and regular students. Rank equal to 1 means a student that is the best-ranked student of the class. Rank equal to 0 means a student that is the worst-ranked student of the class. The class is composed by both affirmative action and regular students. Column 1 shows the average rank for the worst-ranked in the first class and for the best-ranked student in the second class considering affirmative action students, while Column 2 is analogous to column 1, but considering regular students.

**Table A.2: Average share of AA students in the first and second class**

| Major | % A.A. first class | % A.A. second class |
|---|---|---|
| Civil Engineering | 46.45 | 44.82 |
| Mechanical Engineering | 45.97 | 45.31 |
| Chemical Engineering | 46.66 | 42.82 |
| Computer Science | 45.95 | 35.47 |
| Nursing | 46.34 | 46.09 |
| Pharmacy | 47.78 | 45.11 |
| Medicine | 45.55 | 45.89 |
| Veterinary Science | 46.92 | 45.86 |
| Nutrition | 46.75 | 46.11 |
| Dentistry | 46.51 | 45.42 |
| Phonoaudiology | 48.52 | 42.50 |
| Physiotherapy | 46.75 | 44.30 |
| Biotechnology | 46.75 | 42.89 |
| Accounting | 45.40 | 42.85 |
| Social Sciences | 54.71 | 43.83 |
| Journalism | 43.83 | 44.61 |
| Cultural Production | 48.98 | 40.85 |
| Law - Day Shift | 46.60 | 44.10 |
| Pedagogy | 47.53 | 35.96 |
| Psychology | 47.39 | 43.12 |
| Executive Assistant | 48.70 | 38.96 |
| Business | 46.15 | 42.57 |
| Social Service | 47.73 | 44.43 |
| Law - Night Shift | 47.03 | 44.53 |

*Notes*: This table presents the average share(%) of AA students in the first and second class per major.

## Table A.3: Average Class Size by Semester and Major and Class Size Difference Between First and Second Semester by Major

| Majors | 20101 (Enrolled) | 20102 (Enrolled) | 20101 (Expected) | 20102 (Expected) | Average Difference (2006 - 2012) |
|---|---|---|---|---|---|
| **Social Sciences** | | | | | |
| Management | 78 | 62 | 80 | 75 | 8.5 |
| Cultural Production | 27 | 25 | 30 | 30 | 1.4 |
| Law (Night Course) | 97 | 89 | 100 | 100 | 5.3 |
| Law (Morning Course) | 101 | 92 | 100 | 100 | 0.2 |
| Finance | 54 | 51 | 55 | 55 | 2.8 |
| Social Science | 41 | 44 | 60 | 60 | 0.2 |
| Communication & Journalism | 27 | 26 | 30 | 30 | 0.3 |
| Management Assistant | 34 | 36 | 40 | 40 | -0.8 |
| Pedagogy | 44 | 39 | 80 | 40 | 2.5 |
| Social Service | 42 | 43 | 45 | 45 | 0.6 |
| **Technology** | | | | | |
| Biotechnology | 29 | 29 | 30 | 30 | -0.2 |
| Computer Science | 41 | 42 | 45 | 45 | 0.4 |
| Civil Engineering | 88 | 73 | 90 | 90 | 3.5 |
| Mechanical Engineering | 40 | 39 | 45 | 45 | 0.7 |
| Chemistry Engineering | 46 | 42 | 45 | 45 | -0.1 |
| **Health** | | | | | |
| Nurse | 50 | 49 | 50 | 50 | 1.6 |
| Pharmacy | 68 | 63 | 70 | 70 | -0.5 |
| Physiotherapy | 43 | 43 | 45 | 45 | -1.0 |
| Speech Therapist | 28 | 30 | 30 | 30 | -0.5 |
| Medicine | 78 | 73 | 80 | 80 | 2.7 |
| Vet Medicine | 73 | 68 | 75 | 75 | 1.4 |
| Nutrition | 46 | 45 | 50 | 50 | -0.6 |
| Dentist | 58 | 55 | 60 | 60 | -0.1 |
| Biology | 29 | 33 | 45 | 45 | 0.1 |
| **All Majors** | | | | | **1.1** |

*Notes*: This table presents the expected number of enrolled students and the actual number of enrolled students for each major in the sample for the first and second semesters of 2010. Column 1 and 2 presents the number of students in UFBA's database per major and semester in 2010. Column 2 presents the expected number of students per major and class in 2010 according to documents available at UFBA's webpage. Column 5 presents the average difference between class sizes in the first and second semester for each course and for all majors between 2006 and 2012, the period we used to estimate the model.

## Table A.4: Impacts of Class Allocation on Academic Outcomes For Affirmative Action Students dropping the last approved AA student of the first class per major and year

| | (1) 1st Sem. GPA | (2) 1st Year GPA | (3) 10th Sem. GPA | (4) Dropout | (5) 10th Sem. Failures | (6) Graduation on time |
|---|---|---|---|---|---|---|
| Everyone | -0.043 | -0.107 | -0.103 | -0.007 | 0.818 | 0.03 |
| | (0.142) | (0.133) | (0.158) | (0.046) | (0.953) | (0.055) |
| Technology Majors | -0.266 | -0.338* | -0.155 | 0.09 | 4.829*** | -0.152 |
| | (0.236) | (0.183) | (0.138) | (0.209) | (1.769) | (0.101) |
| Health Majors | -0.216 | -0.087 | -0.331* | 0.029 | 2.549* | 0.013 |
| | (0.161) | (0.154) | (0.174) | (0.065) | (1.464) | (0.08) |
| Social Sciences Majors | 0.158 | -0.075 | 0.05 | -0.066 | -0.784 | 0.106 |
| | (0.272) | (0.24) | (0.192) | (0.054) | (1.594) | (0.073) |
| Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Major FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Gender | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Notes*: This table reports the estimated impacts of being allocated to the first class on six outcomes. The hour-weighted average grade in the first semester ("1st Sem GPA"), the hour-weighted average grade in the first year ("1st Year GPA"), the hour-weighted average grade in the first 10 semesters ("10th Sem GPA") – we use this definition instead of the final course GPA to avoid selection into graduation, the total number of failed courses in the first 10 semesters ("Failures"), a dummy which is equal to 1 if the student eventually evaded the major ("Dropout"), and a dummy equal to 1 if the student graduated on time ("Graduation on time"). In this specification, we dropped the last approved candidate of the first class per major and year as Duryea et al. (2023). The estimates refer to the RDD coefficient $\beta_2$ from equation (1). The optimal bandwidth is estimated according to Calonico et al. (2014). The unit of observation is a student-semester-major, and the analysis includes the cohorts of students enrolling at UFBA in the years 2006-2012. The running variable is the entry exam score and the cutoff is the score of the last student enrolled in the first class. Standard errors clustered at the major level are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

**Table A.5: Class Allocation Effects On Labor Market Outcomes For Affirmative Action Students dropping the last approved AA student of the first class per major and year**

| | Has a Job | Number of Years With Formal Employment | Total Wage | Log (Mean Wage) |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Everyone | 0.067 | -0.322 | -1290.135 | -0.176 |
| | (0.066) | (0.565) | (1851.739) | (0.11) |
| Technology Majors | -0.04 | 0.195 | 913.967 | -0.551** |
| | (0.11) | (0.959) | (3488.747) | (0.254) |
| Health Majors | 0.175* | 0.205 | -191.223 | 0.047 |
| | (0.097) | (0.952) | (1156.407) | (0.227) |
| Social Sciences Majors | 0.019 | -1.346 | -2701.951 | -0.276** |
| | (0.101) | (1.019) | (2958.102) | (0.118) |
| Cohort FE | ✓ | ✓ | ✓ | ✓ |
| Major FE | ✓ | ✓ | ✓ | ✓ |
| Gender | ✓ | ✓ | ✓ | ✓ |

*Notes*: This table reports the estimated impacts of being allocated to the first class on four outcomes. A dummy equal to 1 if the student had formal employment between the first and eight years after enrollment ("Has a Job"), the number of years in which the student had formal employment between first and eight years after enrollment ("Number of Years With Formal Employment"), the total inflation-corrected earnings between first and eight years after enrollment ("Total Wage"), the log of mean wages between first and eight years after enrollment ("Log (Mean Wage)"). In this specification, we dropped the last approved candidate of the first class per major and year as Duryea et al. (2023). The estimates refer to the RDD coefficient $\beta_2$ from equation (1). The optimal bandwidth is estimated according to Calonico et al. (2014). The unit of observation is a student-semester-major, and the analysis includes the cohorts of students enrolling at UFBA in the years 2006-2012. The running variable is the entry exam score, and the cutoff is the score of the last student enrolled in the first class. Standard errors clustered at the major level are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

**Table A.6: Impacts of Class Allocation on Academic Outcomes For Affirmative Action Students inputting zeroes when students fail by attendance**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | 1st Sem. GPA | 1st Year GPA | 10th Sem. GPA | Dropout | 10th Sem. Failures | Graduation on time |
| Everyone | -0.023 | -0.077 | -0.102 | 0.021 | 1.124 | -0.003 |
| | (0.118) | (0.104) | (0.11) | (0.04) | (0.784) | (0.042) |
| Technology Majors | -0.393** | -0.38* | -0.264* | 0.099 | 4.335*** | -0.129 |
| | (0.168) | (0.211) | (0.152) | (0.207) | (1.226) | (0.101) |
| Health Majors | -0.282 | -0.288* | -0.374** | 0.09 | 4.146*** | -0.059 |
| | (0.179) | (0.16) | (0.167) | (0.057) | (1.116) | (0.063) |
| Social Sciences Majors | 0.292 | 0.2 | 0.235* | -0.076* | -0.994 | 0.092 |
| | (0.209) | (0.184) | (0.132) | (0.046) | (1.267) | (0.059) |
| Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Major FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Gender | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Notes*: This table reports the estimated impacts of being allocated to the first class on six outcomes. The hour-weighted average grade in the first semester ("1st Sem GPA"), the hour-weighted average grade in the first year ("1st Year GPA"), the hour-weighted average grade in the first 10 semesters ("10th Sem GPA") – we use this definition instead of the final course GPA to avoid selection into graduation, the total number of failed courses in the first 10 semesters ("Failures"), a dummy which is equal to 1 if the student eventually evaded the major ("Dropout"), and a dummy equal to 1 if the student graduated on time ("Graduation on time"). In this specification, students receive a grade of zero when they fail by attendance. The estimates refer to the RDD coefficient $\beta_2$ from equation (1). The optimal bandwidth is estimated according to Calonico et al. (2014). The unit of observation is a student-semester-major, and the analysis includes the cohorts of students enrolling at UFBA in the years 2006-2012. The running variable is the entry exam score and the cutoff is the score of the last student enrolled in the first class. Standard errors clustered at the major level are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

## Table A.7: Impacts of Class Allocation on First-Semester Courses Grades For Affirmative Action Students Without Professor Fixed-Effects

| | Technology | | | | Health | | | | Social Sciences | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| | Introductory Calculus | Linear Algebra | Descriptive I Drawing I | Physics I | MicroBiology | Anatomy | Anatomy II | Histology II | Introduction to Law | Anthropology | Introduction to Philosophy | Political Science |
| Course Grade | -1.332*** | -0.555 | -0.437 | -0.139 | -0.511 | -0.45 | 0.648 | -0.363 | -0.295 | 0.505 | -0.121 | 0.035 |
| | (0.307) | (0.455) | (0.42) | (0.416) | (0.458) | (0.306) | (0.571) | (0.498) | (0.278) | (0.551) | (0.312) | (0.313) |
| Prob(Failure) | 0.576*** | 0.222 | 0.097 | 0.19 | 0.367 | 0.089 | 0.000*** | 0.056 | -0.278* | -0.124 | -0.09 | -0.002 |
| | (0.111) | (0.147) | (0.108) | (0.181) | (0.229) | (0.068) | (00) | (0.067) | (0.146) | (0.08) | (0.071) | (0.077) |
| Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Major FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Gender | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

*Notes*: This table reports the estimated impacts of being allocated to the first class on two outcomes for twelve. Course standardized grades ("Course Grade") and the probability of failing the course ("Prob(Failure)"). We selected courses with a minimum number of students that allowed us to estimate the optimal bandwidth proposed by Calonico et al. (2014). The estimates refer to the RDD coefficient $\beta_2$ from equation (1). The unit of observation is a student-semester-major, and the analysis includes the cohorts of students enrolling at UFBA in the years 2006-2012. The sample includes students taking the course for the first time and we exclude courses given by temporary professors for this analysis. The running variable is the entry exam score and the cutoff is the score of the last student enrolled in the first class. Standard errors clustered at the major level are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

**Table A.8: Class Allocation Effects On Labor Market Outcomes For Regular Students**

| | Had a Job | # of Jobs | Total Wage | Log(Mean Wage) |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Everyone | 0.012 | -0.029 | 320.47 | 0.108 |
| | (0.036) | (0.283) | (1100.288) | (0.068) |
| Technology Majors | 0.051 | 0.137 | -2585.959 | -0.054 |
| | (0.042) | (0.487) | (2242.903) | (0.091) |
| Health Majors | 0.006 | -0.013 | 239.261 | 0.145 |
| | (0.066) | (0.276) | (1241.048) | (0.121) |
| Social Sciences Majors | 0.026 | -0.015 | 3349.966 | 0.132 |
| | (0.043) | (0.494) | (3116.944) | (0.116) |
| Cohort FE | ✓ | ✓ | ✓ | ✓ |
| Major FE | ✓ | ✓ | ✓ | ✓ |
| Gender | ✓ | ✓ | ✓ | ✓ |

*Notes*: This table reports the estimated impacts of being allocated to the first class on four outcomes. A dummy equal to 1 if the student had formal employment between the first and eight years after enrollment ("Has a Job"), the number of years in which the student had formal employment between first and eight years after enrollment ("Number of Years With Formal Employment"), the total inflation-corrected earnings between first and eight years after enrollment ("Total Wage"), the log of mean wages between first and eight years after enrollment ("Log (Mean Wage)"). The estimates refer to the RDD coefficient $\beta_2$ from equation (1). The optimal bandwidth is estimated according to Calonico et al. (2014). The unit of observation is a student-semester-major, and the analysis includes the cohorts of students enrolling at UFBA in the years 2006-2012. The running variable is the entry exam score, and the cutoff is the score of the last student enrolled in the first class. Standard errors clustered at the major level are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

## Table A.9: Impacts of Class Allocation on Academic Outcomes for Students Who Enrolled in 2003

| | (1) 1st Sem. GPA | (2) 1st Year GPA | (3) Final GPA | (4) Dropout | (5) Failures | (6) Graduation on Time |
|---|---|---|---|---|---|---|
| Everyone | 0.246 | 0.076 | 0.327 | -0.059 | -0.839 | 0.095 |
| | (0.253) | (0.248) | (0.26) | (0.099) | (1.215) | (0.117) |
| Technology Majors | 1.173 | 0.924 | 0.956 | -0.38 | -5.624* | 0.517 |
| | (1.087) | (0.973) | (0.954) | (0.428) | (3.302) | (0.456) |
| Health Majors | -0.371 | -0.264 | 0.25 | -0.047 | 0.28 | 0.072 |
| | (0.266) | (0.303) | (0.372) | (0.107) | (0.842) | (0.134) |
| Social Sciences Majors | 0.489 | 0.031 | 0.239 | 0.01 | -0.177 | 0.073 |
| | (0.404) | (0.511) | (0.454) | (0.208) | (2.954) | (0.26) |
| Cohort FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Major FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Gender | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

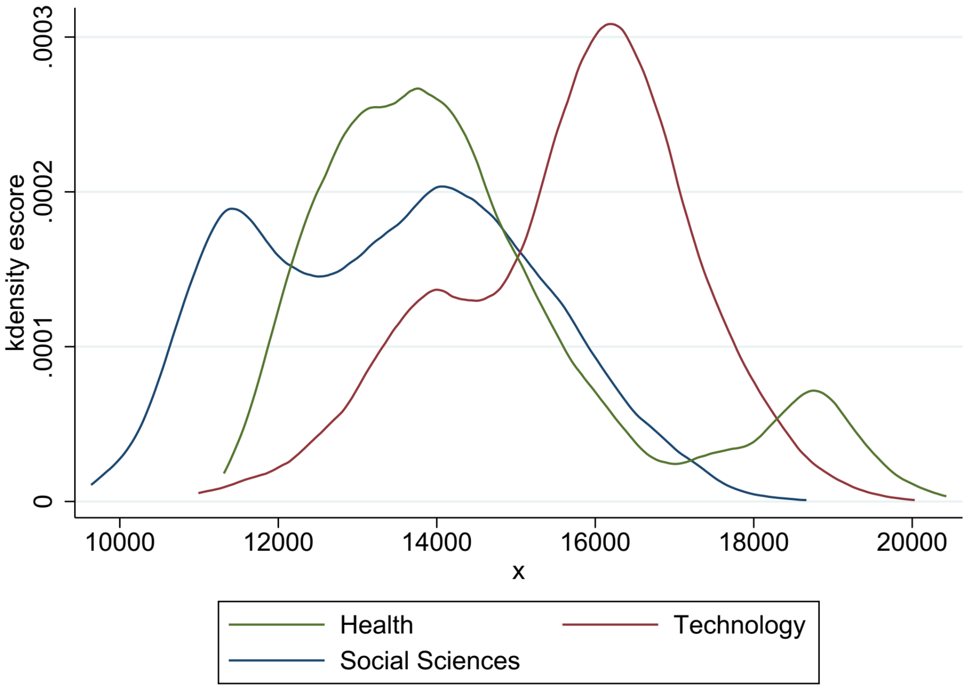
*Notes*: This table reports the estimated impacts of being allocated to the first class on five outcomes. The hour-weighted average grade in the first semester ("1st Sem GPA"), the hour-weighted average grade in the first year ("1st Year GPA"), the hour-weighted average grade in the first 10 semesters ("10th Sem GPA") – we use this definition instead of the final course GPA to avoid selection into graduation, the total number of failed courses in the first 10 semesters ("Failures") and a dummy which is equal to 1 if the student eventaully evaded the major ("Dropout"). The estimates refer to the RDD coefficient $\beta_2$ from equation (1). The optimal bandwidth is estimated according to Calonico et al. (2014). The unit of observation is a student-semester-major, and the analysis includes the cohorts of students enrolling at UFBA in the years 2006-2012. The running variable is the entry exam score and the cutoff is the score of the last student enrolled in the first class. Standard errors clustered at the major level are reported in parentheses. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

**Table A.10: First Semester Grades For First And Second Class Students Controlling for Instructor-by-Course Fixed Effects**

| | (1) Course Grade |
|---|---|
| AA Student | 0.0735*** |
| | (0.0233) |
| Second Semester Course | -0.0719* |
| | (0.0379) |
| Entry Exam Score | 0.000249*** |
| | (0.0000117) |
| AA Student × Second Semester Course | -0.0396 |
| | (0.0300) |
| Observations | 44,670 |
| Cohort FE | ✓ |
| Major FE | ✓ |
| Instructor-by-Course FE | ✓ |

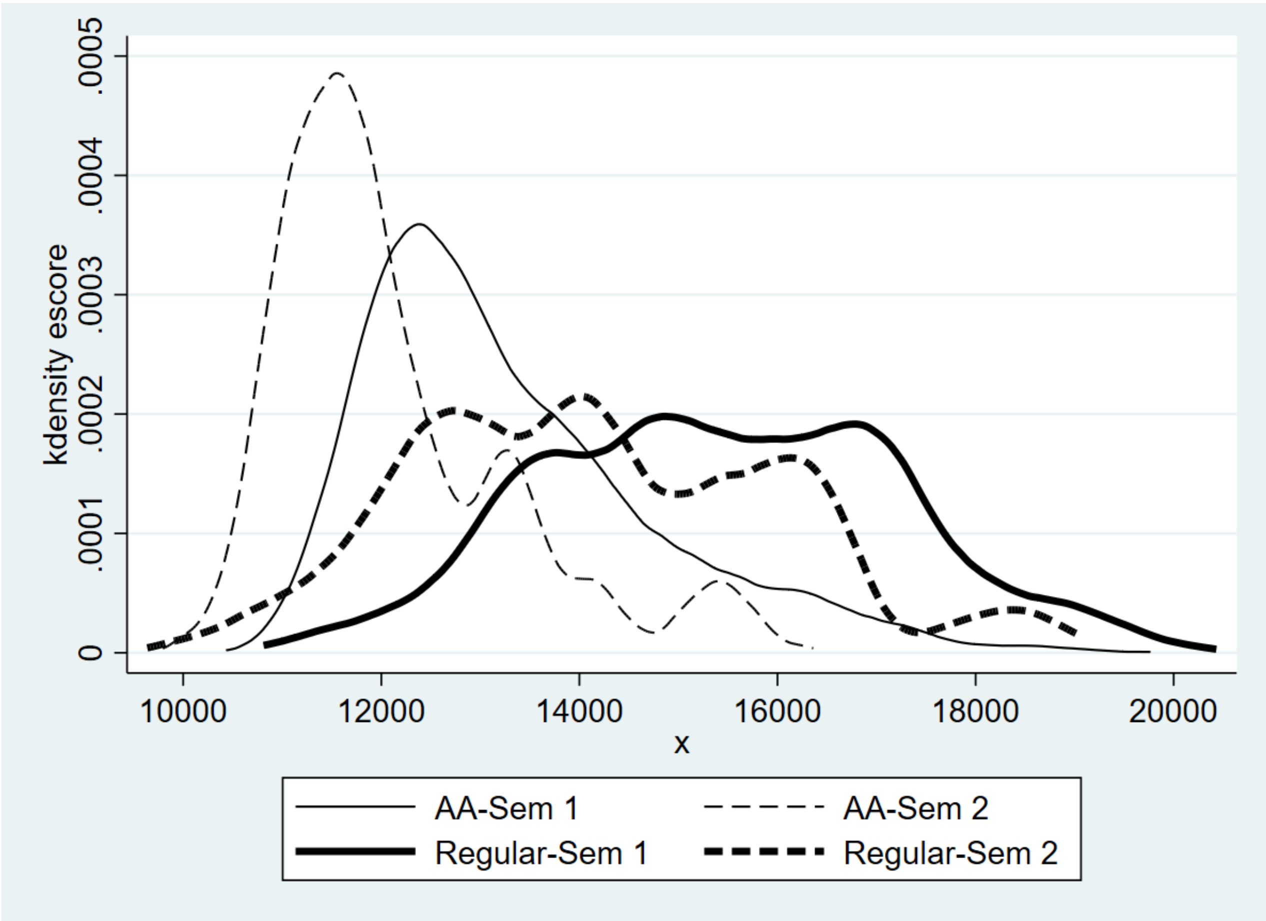
*Notes*: This table reports a comparison of the average grade in first semester courses ("Course Grade") for students who enroll in the firs The estimates in column 1 refer to coefficients $\beta_0$, $\beta_1$, $\beta_2$ and $\beta_3$ from equation (2). In this estimation, we include only first-semester c taught by the same instructor to first and second semester classes and add instructor-by-course fixed effects so that we only compare st first and second classes who took the same course with the same instructor. The unit of observation is a student-course, and the anal cohorts of students enrolling at UFBA in the years 2006-2012. "Entry exam score" refers to the entry exam score as measured by the s the student in the entry exam. Standard errors clustered at the course-professor are reported in parentheses. *** denotes statistical signif level, ** at the 5% level, and * at the 10% level.

**Fig. A.1: Distribution of The Entrance Exam Score for Regular Students**



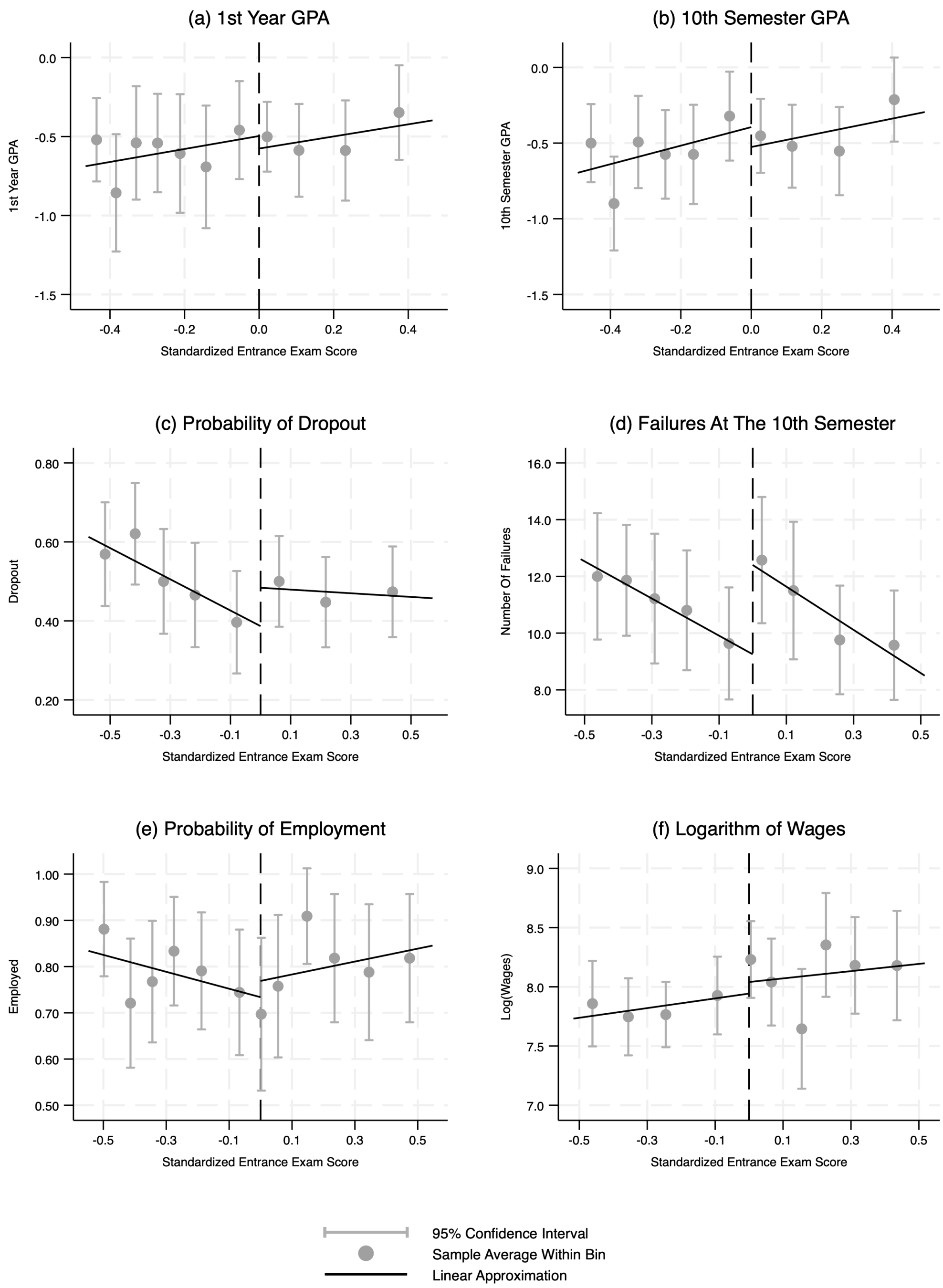

*Notes.* This figure shows the distribution of the entrance exam score for regular students. We present the distribution for the main broad fields of study, technology, health, and social sciences.

**Fig. A.2: Vestibular score distributions for affirmative action and regular students in each semester**



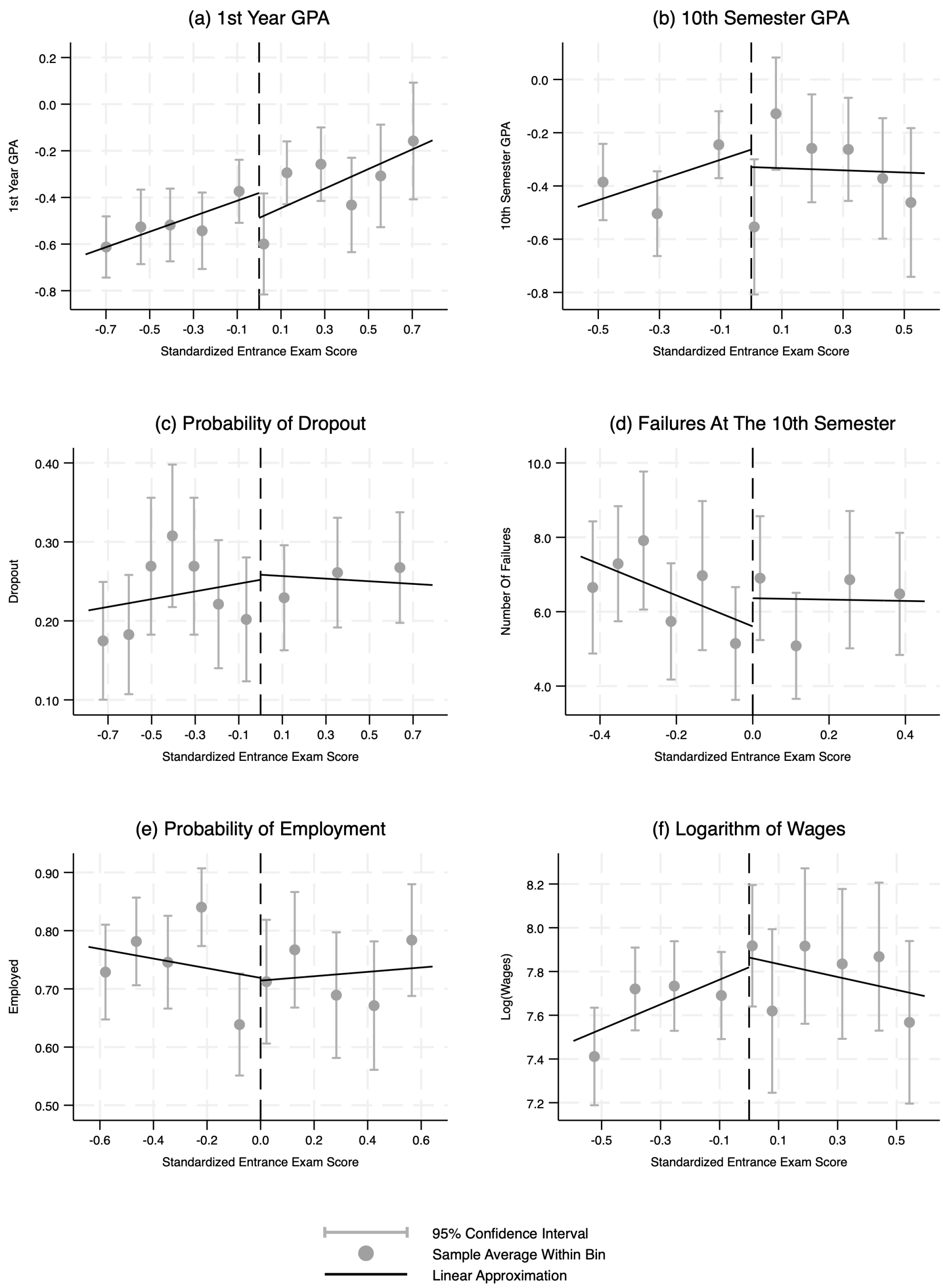

*Notes.* This figure shows the vestibular escore distributions for affirmative action (AA) and regular students, in the first (Sem 1) and second (Sem 2) semesters. The D-statistic of the Kolmogorov-Smirnov test between regular students in the first semester was 0.2912 while for affirmative action students 0.4352. It suggest that regular students between semesters are more similar than affirmative action students.

**Fig. A.3: Affirmative Action Students' Enrolled in Technological Majors Outcomes Along the Standardized Entrance Exam Score**



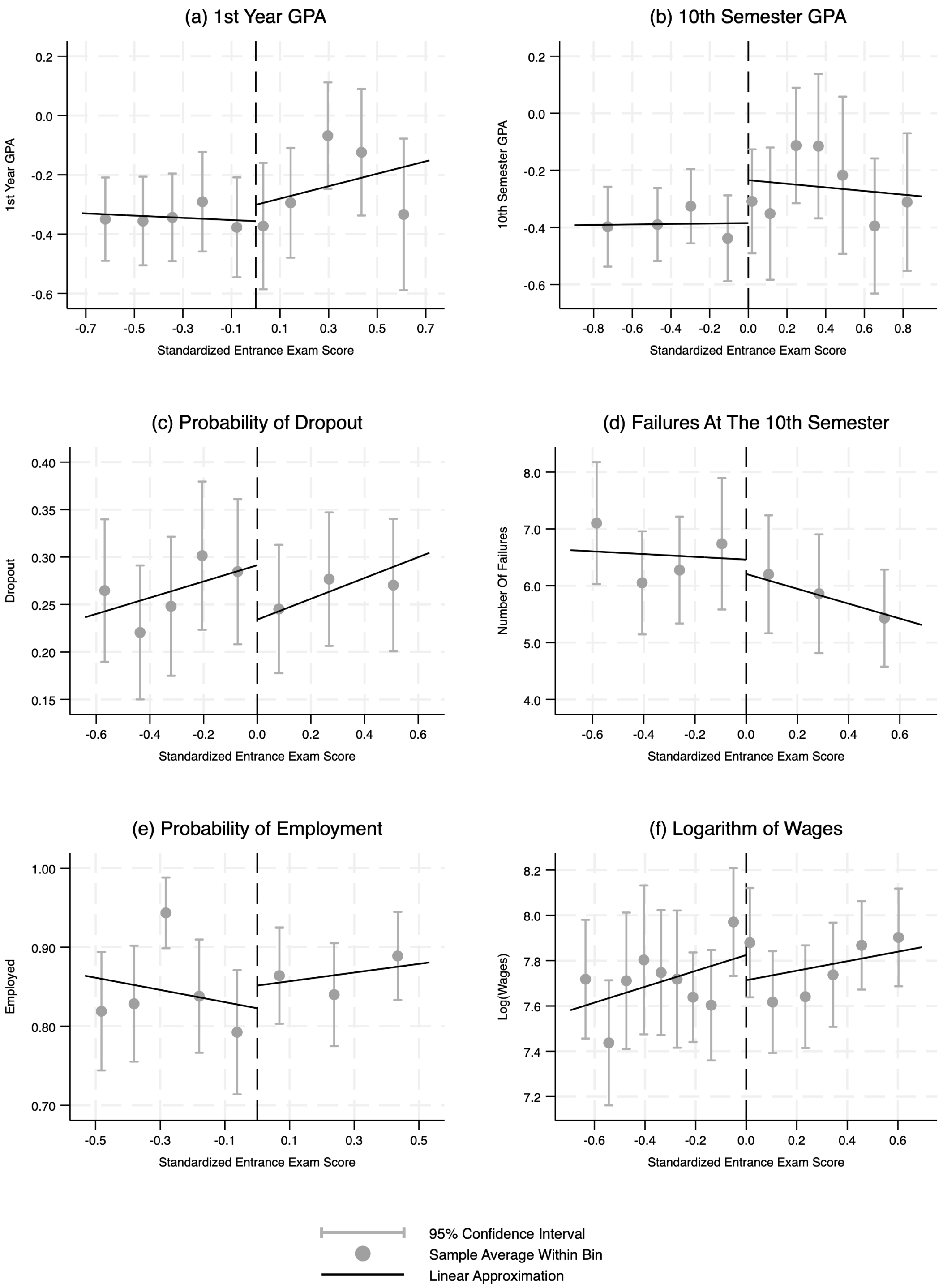

*Notes*: This figure displays the linear approximation for (a) 1st year GPA, (b) 10th semester GPA, (c) probability of dropout, (d) cumulative number of failed courses at the 10th semester, (e) probability of formal employment between 2013 and 2020, and (f) logarithm of mean wages between 2013 and 2020 for affirmative action students enrolled in technology majors coming from estimating (1). The bandwidth used in each graph is estimated according to Calonico et al. (2014). For each graph the observations are binned based on the standardized entry exam score so that each bin has the same number of observations, and the 95% confidence interval is calculated based on the standard deviation for each bin.

**Fig. A.4: Affirmative Action Students' Enrolled in Health Majors Outcomes Along the Standardized Entrance Exam Score**




*Notes*: This figure displays the linear approximation for (a) 1st year GPA, (b) 10th semester GPA, (c) probability of dropout, (d) cumulative number of failed courses at the 10th semester, (e) probability of formal employment between 2013 and 2020, and (f) logarithm of mean wages between 2013 and 2020 for affirmative action students enrolled in health majors coming from estimating (1). The bandwidth used in each graph is estimated according to Calonico et al. (2014). For each graph the observations are binned based on the standardized entry exam score so that each bin has the same number of observations, and the 95% confidence interval is calculated based on the standard deviation for each bin.

**Fig. A.5: Affirmative Action Students' Enrolled in social sciences Majors Outcomes Along the Standardized Entrance Exam Score**




*Notes*: This figure displays the linear approximation for (a) 1st year GPA, (b) 10th semester GPA, (c) probability of dropout, (d) cumulative number of failed courses at the 10th semester, (e) probability of formal employment between 2013 and 2020, and (f) logarithm of mean wages between 2013 and 2020 for affirmative action students enrolled in social science majors coming from estimating (1). The bandwidth used in each graph is estimated according to Calonico et al. (2014). For each graph the observations are binned based on the standardized entry exam score so that each bin has the same number of observations, and the 95% confidence interval is calculated based on the standard deviation for each bin.

# B Appendix: Manipulation test

**Table B.1: Manipulation test for affirmative action and regular students**

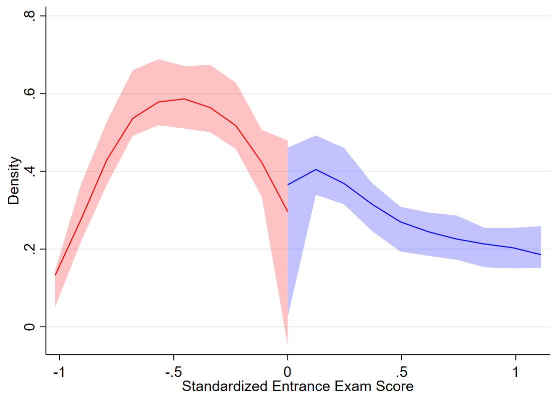
| Kernel Type | All | Technology | Health | Social Sciences |
|---|---|---|---|---|
| Panel A: Affirmative Action students | | | | |
| Triangular | 0.1826 | 0.877 | 0.877 | 0.370 |
| Epanechnikov | 0.244 | 0.937 | 0.355 | 0.685 |
| Panel B: Regular students | | | | |
| Triangular | 0.079 | 0.776 | 0.02 | 0.144 |
| Epanechnikov | 0.1546 | 0.158 | 0.0192 | 0.840 |

*Notes*: This table reports the p-values of the manipulation test proposed by Cattaneo et al. (2020). Panel A report the manipulation test for affirmative action students and Panel B for regular students. We report the results using Triangular and Epanechinkov kernel functions.

## Fig. B.1: Manipulation tests for affirmative action students - Triangular Kernel

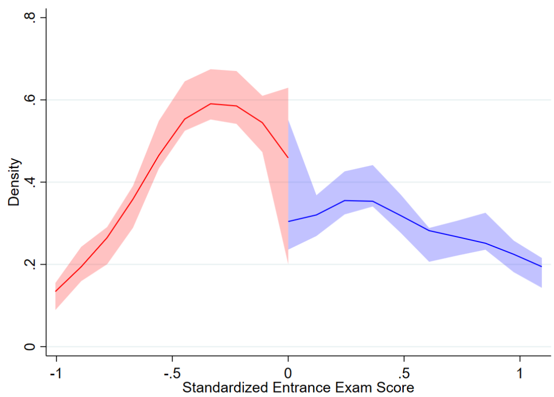
(a) AA


(b) AA - Technology




(c) AA - Health


(d) AA - Social Sciences

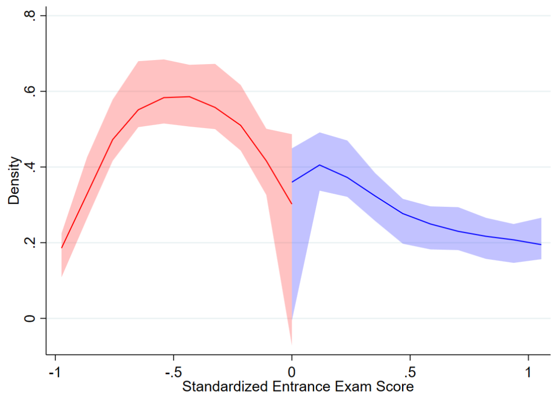



*Notes.* This figure reports the manipulation test proposed by Cattaneo et al. (2020) using a Triangular Kernel for the AA students sample. The shade areas are the 95% confidence intervals.

**Fig. B.2: Manipulation tests for affirmative action students - Epanechnikov Kernel**

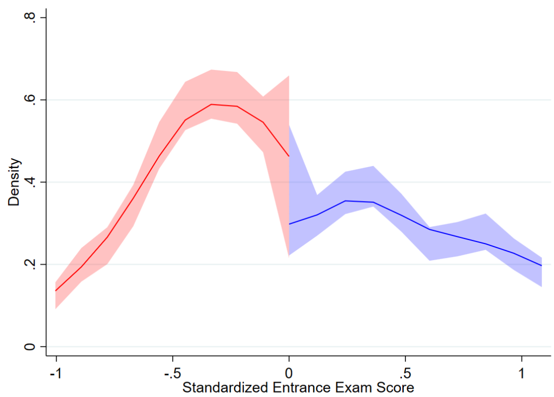
(a) AA


(b) AA - Technology




(c) AA - Health

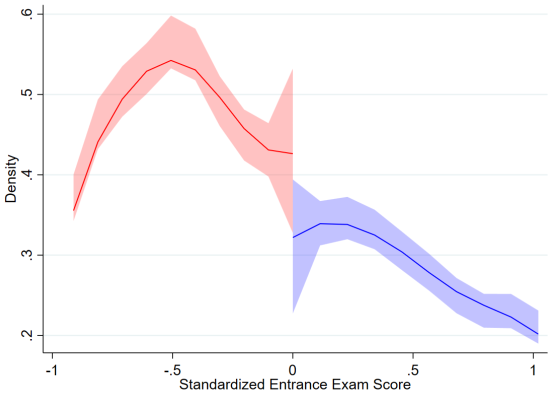

(d) AA - Social Sciences



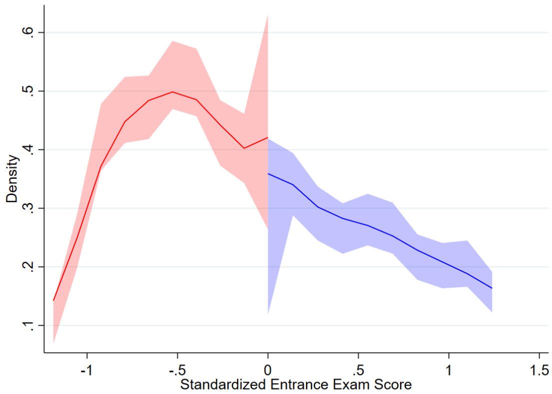

*Notes.* This figure reports the manipulation test proposed by Cattaneo et al. (2020) using an Epanechinkov Kernel for the AA students sample. The shade areas are the 95% confidence intervals.

**Fig. B.3: Manipulation tests for regular students - Triangular Kernel**

(a) Regular

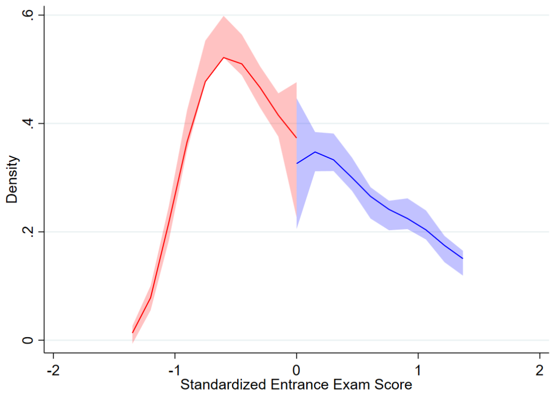



(b) Regular - Technology




(c) Regular - Health

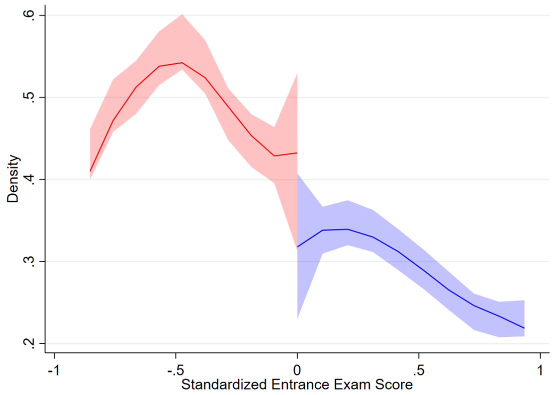

(d) Regular - Social Sciences




*Notes.* This figure reports the manipulation test proposed by Cattaneo et al. (2020) using a Triangular Kernel for the regular students' sample. The shade areas are the 95% confidence intervals.

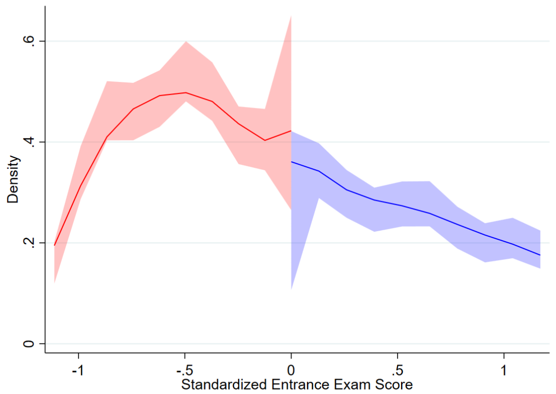
**Fig. B.4: Manipulation tests for regular students - Epanechnikov Kernel**

(a) Regular

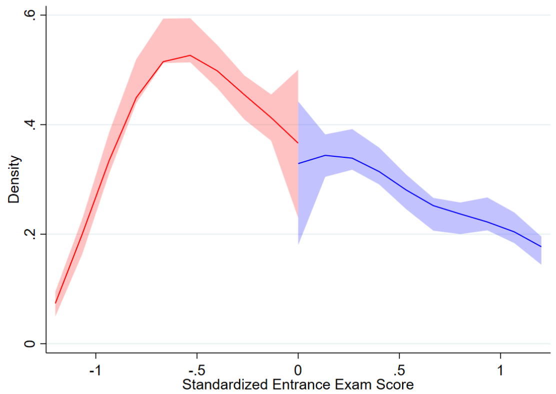



(b) Regular - Technology


(c) Regular - Health




(d) Regular- Social Sciences




*Notes.* This figure reports the manipulation test proposed by Cattaneo et al. (2020) using an Epanechinkov Kernel for the regular students sample. The shade areas are the 95% confidence intervals.