The direct and spillover effects of large-scale affirmative action at an elite Brazilian university

Cecilia Machado[#]
Germán Reyes*
Evan Riehl[+]

May 2023

ABSTRACT. We examine the effects of an affirmative action policy at an elite Brazilian university that reserved 45 percent of admission slots for Black and low-income students. We find that marginally-admitted students who enrolled through the affirmative action tracks experienced a 14 percent increase in early-career earnings. But the adoption of affirmative action also caused a large decrease in earnings for the university's most highly-ranked students. We present evidence that the negative spillover effects on highly-ranked students' earnings were driven by both a reduction in human capital accumulation and a decline in the value of networking.

Top universities are under growing pressure to increase the racial and socioeconomic diversity of their student bodies. In an influential paper, Chetty et al. (2020) argue that large-scale income-based affirmative action at selective colleges could significantly increase intergenerational mobility in the United States. Consistent with this, Bleemer (2022) finds that race-based affirmative action in the University of California system increased the earnings of underrepresented minority students, and he argues that the policy improved allocative efficiency because students who were displaced to other schools were not worse off.

These arguments about the benefits of affirmative action rely on the assumption that selective universities can increase diversity without reducing their value added.[1] Yet the value of attending a top college may depend on its student body composition. At schools with high-achieving students, professors can teach courses at an advanced level (Duflo et al., 2011), and individuals may learn from peers with similar academic preparation (Carrell et al., 2013; Arcidiacono et al., 2016). At schools with wealthy students, individuals may gain access to peer and alumni job networks in high-paying sectors of the economy (Zimmerman, 2019; Michelman et al., 2022). Employers' recruiting and hiring decisions may depend on the expected ability of a school's student body (MacLeod et al., 2017; Weinstein, 2018). If these mechanisms are important, then policies that significantly increase the scale of affirmative action can have negative spillover effects on the outcomes of *all* students. There is little compelling evidence on the existence and magnitude of such spillovers because it is difficult to isolate variation in the composition of a college's student body.

In this paper, we examine the direct and spillover effects of large-scale affirmative action at Rio de Janeiro State University (UERJ), which is one of the most prestigious universities in Brazil. UERJ is consistently ranked among the top 15 universities nationally, and in some years, over 100,000 students take UERJ's entrance exam to compete for roughly 5,000 admission slots. Thus in terms of national prestige and the difficulty of gaining admission, UERJ is comparable to elite private colleges in the United States.

UERJ was one of the first universities in Brazil to adopt affirmative action, and it did so at a large scale. Historically, white students from private high schools were disproportionately likely to gain admission through UERJ's entrance exam. In 2004, UERJ began reserving 45 percent of slots in each major for Black and public high school students from low-income families. This policy caused a sudden and dramatic increase in the racial and socioeconomic diversity of UERJ's student body.

---

[1] Chetty et al. (2020) write: "[W]e also assume that [our] estimated causal effects do not change under our counterfactual student reallocations, in particular ignoring potential changes in value-added that may arise from having a different group of students (peer effects)" (p. 1626). Similarly, Bleemer (2022)'s claim that affirmative action improved allocative efficiency relies on the untested assumption that the policy did not reduce the returns of inframarginal white and Asian students.

We collected data on the schooling and labor market outcomes of students who applied to UERJ before and after the adoption of affirmative action (AA). Our base dataset contains the entrance exam scores and admission outcomes of all UERJ applicants in 1995–2001 (pre-AA) and 2004–2011 (post-AA). We link these data to UERJ enrollment/graduation records, and to Brazil's national employer-employee dataset for the years 2003–2019.

Our analysis exploits a unique feature of UERJ's policy that created variation in exposure to affirmative action across majors. Admission to UERJ is major-specific, and while the fraction of slots reserved for affirmative action was the same in each major, the *take-up* of these slots varied significantly. In UERJ's most prestigious programs, affirmative action students made up 45 percent of the incoming class because the number of applicants typically exceeded the reserved quotas. The quotas often went unfilled in less-selective programs, and UERJ would fill open seats from the general applicant pool. Thus the share of enrollees who were from an affirmative action track was as low as 10–20 percent in some programs.

We use two empirical strategies to identify the effects of affirmative action on both its intended beneficiaries and on other students at UERJ. In majors with high take-up of affirmative action, we use a regression discontinuity (RD) design that compares applicants above and below the admission score cutoffs (Hoekstra, 2009; Kirkebøen et al., 2016). Our RD design identifies the returns to attending UERJ for applicants who were marginally-admitted through each track, and, in particular, through the affirmative action tracks.

Our second strategy exploits variation in affirmative action take-up to identify the policy's spillover effects on other UERJ students. We use a simple difference-in-differences (DD) design that estimates changes in outcomes between the pre- and post-AA cohorts, and across majors with higher and lower take-up rates. In this analysis, we focus on a sample of *top enrollees* whose entrance exam scores were high enough to gain admission to UERJ regardless of whether affirmative action existed in their cohort. For these highly-ranked students, our DD design identifies the effects of a 19 percentage point increase in the share of their classmates who were from an affirmative action track.

We have two main findings. First, for students who were marginally-admitted through affirmative action, enrolling in UERJ led to a 14 percent increase in hourly wages measured 6–9 years after application. We find no effects of UERJ enrollment on the likelihood of earning a college or postgraduate degree, but affirmative action enrollees were significantly more likely to obtain jobs at high-paying firms affiliated with UERJ alumni. This suggests that their early-career earnings gains were primarily driven by networking mechanisms, broadly defined. We find evidence that both the earnings and networking benefits decreased as affirmative action students' careers progressed, but our later-career results are less powered.

Second, the adoption of affirmative action lowered the earnings of UERJ's highly-ranked students. Our DD estimates show that top enrollees' hourly wages decreased by 14 percent

in majors with high affirmative action take-up relative to those with lower take-up, and this decline persisted up through the end of our data range. Notably, we also find declines in earnings for highly-ranked underrepresented minority students who could have gained admission to UERJ even in absence of the affirmative action. We do not find significant changes in the demographic characteristics and admission scores of top enrollees in more- vs. less-affected majors, although our point estimates suggest that compositional effects could play a small role in our earnings results. Instead, we find evidence that the negative spillover effects on earnings were driven by both networking and learning mechanisms. The adoption of affirmative action reduced the likelihood that top enrollees obtained jobs at high-paying firms affiliated with UERJ peers and alumni, and it also reduced the performance of UERJ's top students on Brazil's national college exit exam.

Taken together, our findings show that elite universities face a tradeoff between promoting upward mobility for disadvantaged students and maintaining sources of their value added that stem from admitting high-achieving and wealthy students. Further, our results suggest that disadvantaged students with the highest admission scores may have been better off if UERJ's affirmative action policy had been smaller in scale.

Our early-career results for affirmative action students are consistent with other research that finds that disadvantaged students benefit from attending selective universities. There is a large literature on affirmative action in university admissions, but there is limited evidence on the earnings impacts of these policies (Arcidiacono et al., 2015).[2] Our results are consistent with other evidence of earnings gains for disadvantaged students who were given admission preference at selective colleges (Bertrand et al., 2010; Bleemer, 2022), including Francis-Tan and Tannuri-Pianto (2018)'s finding of earnings benefits for male students admitted through reserved quotas at the University of Brasília. Related work finds earnings gains for low-income or minority students who were marginally-admitted to U.S. public university systems (Zimmerman, 2014; Smith et al., 2020; Bleemer, 2021). Our estimate of a 14 percent early-career earnings return for affirmative action students is substantially smaller than analogous estimates from many of these papers.[3] This is likely due to the fact that these papers often find effects on both earnings and bachelor's degree attainment, whereas we find no effects on educational attainment for UERJ's affirmative action applicants.

---

[2] Other research on affirmative action looks primarily at its effects on student body diversity or graduation rates (Cortes, 2010; Backes, 2012; Hinrichs, 2012; Kapor, 2015; Arcidiacono et al., 2016; Bagde et al., 2016). This is true of most work on affirmative action in Brazil (Francis and Tannuri-Pianto, 2012; Ribeiro, 2016; Estevan et al., 2019; Vieira and Arends-Kuenning, 2019; Ribeiro and Estevan, 2021; Mello, 2022).

[3] For example, Zimmerman (2014) finds that *admission* to the Florida State University system increased individuals' earnings by 22 percent, and admission increased the likelihood of enrolling by roughly 50 percent. Bleemer (2022) finds that an affirmative action ban decreased enrollment in selective University of California colleges by eight percentage points for minority applicants, and earnings fell by 0.05 log points. These estimates imply returns to selective college enrollment of roughly 44–87 percent.

Our findings are new in showing that affirmative action can benefit disadvantaged students through networking channels. Zimmerman (2019) and Michelman et al. (2022) show that networking is an important mechanism for the long-run earnings benefits of attending an elite university, but both papers find that these benefits accrue primarily to students from advantaged backgrounds. Our data is unique in allowing us to examine early-career employment in a broad set of firms. Our results suggest that students admitted through affirmative action can also benefit from access to high-wage firms affiliated with alumni, at least at the start of their careers. We also find similar effects for marginal enrollees from the general application pool, suggesting that networking may be an important mechanism in many papers on the returns to college selectivity (e.g., Dale and Krueger, 2002).[4]

Lastly, our paper presents novel evidence on the spillover effects of large-scale affirmative action. Several papers examine the efficiency effects of admission policies that benefit disadvantaged students by comparing the earnings returns for students who were "pulled in" and "pushed out" by these policies. The evidence from this work is mixed; Bleemer (2022) finds efficiency gains, Bertrand et al. (2010) and Riehl (2023) find efficiency losses, and Black et al. (2023) find limited earnings effects in either group.[5] But a full evaluation of the efficacy of affirmative action must also take into account spillover effects on untargeted students (Durlauf, 2008). A few papers examine how a university's racial or socioeconomic diversity affects the earnings of other students (Daniel et al., 2001; Arcidiacono and Vigdor, 2010; Hinrichs, 2011), but this work relies on strong selection-on-observables assumptions. We find negative earnings spillovers under weaker assumptions, and we present evidence on both learning and networking mechanisms for these effects. The existence of spillovers means that the true effects of large-scale admission reforms can differ significantly from those estimated using existing student/college allocations, as in, for example, Chetty et al. (2020)'s "need-affirmative" counterfactual admission policy.

1. Context and data

1.1. **UERJ and higher education in Brazil.** Our setting is an elite public university in Brazil called Rio de Janeiro State University, or UERJ (*Universidade do Estado do Rio de Janeiro*). It is one of the oldest and most prestigious universities in Brazil; UERJ ranked

---

[4] There is a large literature on the earnings returns to attending selective colleges and/or majors (Hoekstra, 2009; Saavedra, 2009; Hastings et al., 2013; Kirkebøen et al., 2016; Canaan and Mouganie, 2018; Hoxby, 2018; Anelli, 2020; Sekhri, 2020; Ng and Riehl, 2022). These papers typically do not have the data necessary to examine both job networks and earnings. Our paper adds to a small body of work on network formation in college and its effects on job finding (Marmaros and Sacerdote, 2002; Mayer and Puller, 2008; Zhu, 2023).

[5] Otero et al. (2021) find evidence of efficiency gains from the nationwide adoption of affirmative action policies in Brazil, but the authors can only examine effects on college persistence given the timing of their data. A related literature examines student/college match effects in graduation and earnings outcomes (Andrews et al., 2016; Arcidiacono et al., 2016; Dillon and Smith, 2020; Mountjoy and Hickman, 2020).

11$^{\text{th}}$ nationally in a 2012 ranking by the newspaper *Folha*. UERJ is part of Brazil's system of *state universities*, which are funded and managed by the governments of each state. Brazil also has *federal universities* that are controlled by the federal government. State and federal universities are highly-regarded and tuition-free, and thus admissions are highly competitive. At UERJ, the number of applicants is often 10–20 times greater than the number of slots. Yet these schools are a small share of the Brazilian higher education system; there are more than 2,000 colleges in the country, and roughly two-thirds of students attend a private institution.

UERJ typically offers 40–50 undergraduate majors each year in a variety of fields. Students apply to specific programs. Admission is determined by a two-round entrance exam that the university administers near the end of each year. The first round consists of a qualifying exam that is common to all applicants. Students who pass the qualifying exam take field exams in several subjects that depend on the major they are applying to. Admissions are based on a weighted average of an applicant's field exam scores. The highest-scoring applicants are admitted up to a cutoff that is determined by the program's capacity.

1.2. **Data.** Our analysis matches two administrative datasets from UERJ to Brazil's national employer-employee records. Our base dataset is a list of all individuals who applied to UERJ in the years 1995–2001 and 2004–2011 (UERJ, 2020a).[6] We focus on applicants who passed the first-round exam, which is the relevant sample of potential admits for our analyses. We observe the program individuals applied to, their overall admission score, and their admission outcome. In some cohorts, we also observe demographic characteristics and subject scores on the entrance exams.[7] Our second dataset covers all students who enrolled in UERJ from 1995–2011 (UERJ, 2020b). In this dataset we observe the student's program, enrollment date, status as of 2020 (graduated, dropped out, or still enrolled), and final year.

Lastly, we use the 2003–2019 years of Brazil's employer-employee dataset, the *Relação Anual de Informações Sociais*, or RAIS (RAIS, 2021). This dataset is assembled by the Ministry of Labor and covers the universe of formal-sector jobs in Brazil. Worker variables include demographics, educational attainment, occupation, hours worked, and monthly earnings. It also includes information on the firm's number of employees, industry, and location.

We link the UERJ and RAIS datasets using individuals' national ID numbers (*Cadastro de Pessoas Físicas*). For individuals with missing ID numbers, we use names and birthdates to link the datasets. Appendix B.2 provides details on the merge process and match rates.

1.3. **Affirmative action at UERJ.** Historically, white students from private high schools were overrepresented at state and federal universities, while Black and public high school students were underrepresented. This was mainly due to the fact that advantaged students

---

[6] UERJ does not have application records for the 2002–2003 cohorts.

[7] Appendix B.1 provides details on our variable definitions and the cohorts in which each variable is available.

typically earned higher scores on the schools' entrance exams.[8] The lack of diversity was a source of contention because these elite universities are publicly-funded and free to attend.

UERJ was one of the first Brazilian universities to address this disparity through affirmative action. In 2003, the state government of Rio de Janeiro passed a law that required UERJ to reserve seats for students from underrepresented groups. Only two other top public universities had affirmative action at the time, and both were located in other states (Júnior and Daflon, 2014). Other universities adopted race- and/or income-based quotas in subsequent years (Ferman and Assunção, 2005; Vieira and Arends-Kuenning, 2019), and a 2012 national law mandated affirmative action at all federal universities. But UERJ was the only elite university in Rio de Janeiro with affirmative action for much of the 2000s.

UERJ implemented the policy by reserving 45 percent of seats in each program for low-income applicants from disadvantaged groups. Historically there was one admission track for each major, but in 2004 UERJ added three affirmative action tracks per major.[9] 20 percent of slots in each major were reserved for public high school applicants. Another 20 percent of slots were reserved for Black applicants.[10] Lastly, 5 percent of slots were reserved for students from other disadvantaged groups (e.g., disabled and indigenous applicants). To be eligible for an affirmative action track, applicants also had to be from a low-income family, and they were required to submit tax records to verify their income.[11] Applicants who did not meet these criteria could apply through the general track, which governed admissions for the remaining seats. Within each track, admissions were based solely on field exam scores.

Although the fraction of reserved slots was the same in each major, the *take-up* of these slots varied significantly. Figure 1 plots the share of affirmative action enrollees in the 2004–2011 cohorts (y-axis) against a measure of each program's selectivity (x-axis). In highly-selective programs like Law and Medicine, the number of applicants to each track typically exceeded the reserved quota, so affirmative action students made up 45 percent of the incoming class. In less selective programs like Math and Teaching, the number of affirmative action applicants was frequently less than the quota, and UERJ filled open slots from the general track. Thus the share of affirmative action enrollees was as low as 10–20

---

[8] Other factors contributed to the underrepresentation of Black and public school students, e.g., exam access and information about the admission process (Hoxby and Avery, 2013; Machado and Szerman, 2021).

[9] UERJ first introduced affirmative action in the 2003 cohort following the state law. In this cohort, there were only two admission tracks—low-income and general—and each track reserved some seats for Black applicants. The quota system described in the text was in place for all cohorts in 2004–2011.

[10] In Brazil, race is commonly classified in five groups: *branco* (white), *pardo* (brown), *preto* (Black), *amarelo* (yellow), and indigenous. UERJ's race-based quota was reserved for individuals who self-identified as Black; this occasionally differs from their racial identity reported in the entrance exam or RAIS data.

[11] In 2004, for example, applicants had to be from families whose per capita income was below R$300 per month (Zoninsein and Júnior, 2008), which was roughly 40 percent of the national GDP per capita.

percent in these programs. The low take-up is attributable to the lower desirability of these programs and the fact that UERJ had strict criteria for affirmative action eligibility.[12]

UERJ's admission policy gave a large implicit preference to affirmative action students. Figure 2 plots the distribution of admission scores for 2004–2011 applicants in the Black, public school, and general tracks, where scores are standardized to be mean zero and standard deviation (SD) one among *all* applicants to a given program/cohort. Vertical lines displays the mean cutoff score in each track, which is the mean of the standardized scores for the last students admitted to each program/cohort. The mean cutoff was −0.5 in the public school track, −0.6 in the Black track, and +0.9 in the general track. Thus the last admitted students in the affirmative action tracks typically scored 1.5 standard deviations below the last admitted student in the general track.

1.4. **Samples.** We create two samples to analyze the impacts of affirmative action at UERJ. In Sections 2–3, we use a regression discontinuity (RD) design that compares the outcomes of admitted and rejected applicants. In Sections 4–5, we use a difference-in-differences (DD) design that compares the outcomes of enrollees in programs with higher and lower take-up rates of affirmative action. We construct our samples to be suitable for these two analyses. Appendix B.4 provides details on our sample construction.

Our RD sample includes programs in which we can identify the effects of admission to UERJ in the affirmative action tracks. Some UERJ programs had few Black and public school applicants, so in many cohorts there were no rejected students in these tracks. Since we cannot implement our RD design in these cases, we restrict our RD sample to programs where the Black and public school quotas typically filled up. Specifically, our RD sample includes the 24 programs in which 30 percent or more of the 2004–2011 enrollees were from an affirmative action track (i.e., programs above the horizontal line in Figure 1). In these programs, we also exclude any cohort/application-track pair in which there are fewer than five applicants below the admission threshold (see Appendix Tables B2–B4).[13]

Our DD sample includes all programs that UERJ offered both before and after 2004.[14] This includes the 24 programs in our RD sample plus 19 other programs with lower rates of affirmative action take-up. We focus on a sample of students who could have attended UERJ regardless of whether affirmative action existed in their cohort (see Section 4).

---

[12] In fact, the *majority* of Black and public high school students who applied to UERJ were ineligible for affirmative action because they did not meet the low-income requirement. All applicants also had to pass the first-round qualifying exam, for which there was no affirmative action.

[13] Although the quotas almost always filled up in the general track, we restrict to the same programs in our RD sample of general applicants so that it is comparable to the Black and public school samples. Our RD sample excludes all applicants in the disabled/indigenous track, as these quotas were rarely filled.

[14] UERJ re-organized a few large programs into sub-programs during our sample period. In these cases we combine sub-programs into one program in our DD analysis. See Appendix Tables B2–B4.

Table 1 shows summary statistics for our RD and DD samples. Panel A includes programs in both samples, and Panel B includes programs that are only in our DD sample. Our RD sample includes all of UERJ's health programs as well as many humanities and social science majors (e.g., History, Journalism, and Social Work). Our DD sample includes many teacher-training programs in different subjects, but it also includes Economics, Math, and several engineering majors. We display statistics separately for applicants in pre-AA cohorts (column A) and to the four tracks in the post-AA cohorts (columns B–E). Affirmative action applicants were disadvantaged relative to general applicants as measured by mother's education and family income. They were also older on average and more likely to be female.

2. RD SPECIFICATION

2.1. **Regression model.** We use a two-stage least squares (2SLS) RD model to estimate the returns to enrolling in UERJ:

$$(1) \qquad E_{ip} = \theta D_{ip} + \alpha x_{ip} + \psi D_{ip} x_{ip} + \gamma_p + \varepsilon_{ip} \quad \text{if } |x_{ip}| \le h^Y$$

$$(2) \qquad Y_{ip} = \beta E_{ip} + \tilde{\alpha} x_{ip} + \tilde{\psi} D_{ip} x_{ip} + \tilde{\gamma}_p + \tilde{\varepsilon}_{ip} \quad \text{if } |x_{ip}| \le h^Y.$$

Y_{ip} is an outcome for individual i who applied to UERJ in application pool p. Application pools are defined by a program, cohort, and admission track. The endogenous treatment variable, E_{ip}, is an indicator that equals one if the applicant enrolled in the UERJ program and cohort that they applied to. We instrument for UERJ enrollment with D_{ip}, which is an indicator for having an admission score above the final cutoff for application pool p.

We use a local linear specification to estimate returns for applicants on the margin of admission. Equations (1)–(2) include an interaction between D_{ip} and the running variable, x_{ip}, which is individual i's admission score in application pool p. We normalize x_{ip} to equal zero for the last admitted student, and so that it has SD one in the population of all applicants to a given program/cohort. Our regressions include the subset of applicants whose admission scores are within h^Y standard deviations of the admission threshold. Our benchmark results use the Calonico et al. (2014) bandwidth computed separately for each outcome Y; Appendix Tables A4–A6 show that our main results are robust to different bandwidths. We include fixed effects for each application pool, γ_p, and we cluster standard errors at the individual level, as some individuals in our sample applied to UERJ more than once.

We estimate equations (1)–(2) separately for three groups: applicants in the pre-AA cohorts (1995–2001), general track applicants in the post-AA cohorts (2004–2011), and affirmative action applicants (pooled across the Black and public school tracks to increase power). The estimates for affirmative action applicants shed light on the direct effects of UERJ's policy on its targeted beneficiaries. The estimates for pre-AA and post-AA general applicants provide suggestive evidence on how the policy changed the returns to attending

UERJ for untargeted students, but this evidence is not conclusive because the policy also implicitly raised the location of the admission thresholds in the general applicant pools.

2.2. **Identification assumptions and balance tests.** Our identification relies on the standard RD, instrumental variable, and local average treatment effect (LATE) assumptions.

The main RD assumption requires that applicants' admission scores are effectively randomly assigned near the admission cutoffs. There is little scope for applicants to manipulate their exam scores, but a potential issue is that UERJ engages in waitlist admissions to fill the available seats. UERJ sends initial admission offers in January, and admits have several weeks to accept or decline their offer. UERJ then fills open seats through multiple rounds of waitlist admission offers to applicants with the next highest scores. Our instrument and running variable, D_{ip} and x_{ip}, are defined by the final threshold in each application pool. This creates the possibility of non-random sorting around the cutoff; the last admitted student may be particularly likely to accept an admission offer, and this tendency may be correlated with potential outcomes. Appendix B.3 describes UERJ's admission process in detail.

We find no evidence that the RD assumption is violated for affirmative action applicants using the standard balance tests. Appendix Table A1 presents RD estimates from regressions that use applicant characteristics as dependent variables (e.g., age, gender, race, and qualifying exam scores). We cannot reject the hypothesis that these coefficients are jointly equal to zero ($p = 0.88$). We find similar results combining these characteristics into an index of predicted wages (Appendix Figure A1). There is no evidence of a discontinuity in the density of admission scores using the McCrary (2008) test (Appendix Figure A2). These results corroborate our prior that the waitlist is unlikely to be an issue in the Black and public school tracks because most applicants accepted their admission offer.

We also find covariate balance for general applicants, but the McCrary test reveals a statistically significant decrease in the admission score densities at both the pre- and post-AA general track thresholds. UERJ's yield was lower in the general track, so there was more scope for non-random sorting from waitlist admissions. Thus our RD results for general applicants should be interpreted with some caution. Reassuringly, we find similar results in "donut hole" regressions that drop applicants near the threshold (Appendix Tables A4–A6).

We also make the standard instrumental variable and LATE assumptions (Angrist et al., 1996). Instrument relevance is satisfied because the UERJ enrollment rate increases sharply at the admission threshold (Table 2, Panel A). The exclusion restriction requires that our instrument affects outcomes only through the channel of enrolling in UERJ. This could be violated if, for example, admission to UERJ caused individuals to apply to other schools. We cannot rule out this possibility, but we believe our results are primarily attributable to UERJ enrollment, particularly in the affirmative action tracks where the first-stage coefficient is

large. The monotonicity assumption requires that there are no applicants who would attend UERJ if and only if they were *below* the cutoff, which is plausible in our setting.

Under these assumptions, the β coefficient from equation (2) can be interpreted as the average causal effect of attending UERJ for marginally-admitted compliers. These compliers are students who would have enrolled if and only if they scored above the cutoff. This estimand measures the returns to UERJ enrollment relative to the educational choices that students would have made if they were not admitted, which is relevant for evaluating the efficacy of UERJ's affirmative action policy for reducing earnings inequality.

3. Effects of affirmative action on marginal admits

3.1. **Graduation and earnings.** We begin our RD analysis by examining the effects of UERJ enrollment on graduation rates and post-college earnings. Table 2 presents RD results for pre-AA general applicants (columns A–B), post-AA general applicants (columns C–D), and affirmative action applicants (columns E–F). Columns (A), (C), and (E) show means of each dependent variable for applicants who scored just below the threshold (within 0.1 SD). Columns (B), (D), and (F) display RD coefficients. Panel A presents first-stage coefficients, θ, from equation (1). Panels B–C show 2SLS RD coefficients, β, from equations (1)–(2) for graduation and earnings outcomes. In Panel B, we measure outcomes 6–9 years after individuals' applied to UERJ to capture their initial jobs after (potential) graduation. To examine longer-run effects, Panel C examines outcomes measured 10–13 years after application.[15] We also present RD graphs for our main outcomes in Figure 3; these graphs show the reduced-form effects of admission to UERJ by plotting means of each outcome in 0.1 SD bins of the standardized admission score.

Panel A of Table 2 shows that crossing the admission threshold increased the likelihood that affirmative action applicants enrolled in UERJ by 69 percentage points (column F). The first stage for affirmative action applicants is more than twice as large as that for general applicants (columns B and D). UERJ's yield was high in the Black and public school tracks because most other universities in Rio did not have affirmative action during 2004–2011. In the general track, marginal admits would typically have been competitive for admission to other top colleges in the area (see Section 3.2).

Our first finding is that marginal enrollees in the affirmative action tracks were only slightly less likely to graduate from UERJ than those in the general track. The first row of Panel B shows the effects of enrolling in a UERJ program on the likelihood of graduating from that program by nine years later. 64 percent of marginal affirmative action enrollees graduated

---

[15] All of our RD regressions include one observation per applicant. For earnings outcomes, we use the applicant's mean real earnings over the periods of 6–9 or 10–13 years after application. For binary outcomes, we use the maximum over each period, so our estimates reflect ever having a job with those characteristics. Most UERJ students who graduate do so in 4–6 years (see Appendix Figure A3).

by this time, as compared with 68–71 percent of marginal enrollees in the general tracks. The similarity of these graduation rates is striking since marginal affirmative action admits scored 1.5 standard deviations below those in the general track on the entrance exam.

UERJ enrollment did not significantly affect the likelihood that individuals worked in the formal sector. Our measures of formal employment are indicators for appearing in the RAIS data at any time in 6–9 or 10–13 years after application. In the affirmative action tracks, the formal employment rates for marginally-rejected applicants are above 70 percent in both time periods (column E), and the 2SLS RD coefficients are close to zero (column F). For pre-AA applicants, we find a positive and statistically significant effect on early-career formal employment (Panel B), but this effect does not persist into the later time period (Panel C).

Importantly, we find that affirmative action students experienced an increase in early-career earnings from attending UERJ. For marginal affirmative action compliers, UERJ enrollment caused a 14 percent increase in mean hourly wages measured 6–9 years after application. The gain in early-career monthly earnings was $110 (in 2019 U.S. dollars), which is similar as a percentage of mean below-threshold earnings. Panel C of Figure 3 shows visual evidence of a discontinuity in the early-career hourly wages of marginally-admitted affirmative action applicants (black triangles). The RD coefficient for monthly earnings is roughly one-fifth of the earnings gap between marginally-rejected general and affirmative action applicants ($1,391 vs. $817). Thus UERJ's affirmative action policy meaningfully reduced early-career earnings inequality among applicants on the margin of admission.

However, we find some evidence that the initial earnings gain for affirmative action students declined as their careers progressed. Panel C of Table 2 shows that the effect of UERJ enrollment on affirmative action students' hourly wages declined to 0.024 log points measured 10–13 years later (see also Panel D of Figure 3). The later-career wage coefficient is not statistically different from zero, and we reject equality of the early- and later-career wage coefficients at $p < 0.05$ (Appendix Table A3). The gain in monthly earnings for affirmative action students also declined to $56 in the later period, although this estimate is not statistically distinguishable from the early-career return. Appendix Figure A4 shows that the wage gains for affirmative action students decreased both over time (holding the sample of cohorts fixed) and across cohorts (holding potential experience fixed).

For general track applicants, we find some evidence of a *negative* early-career return to attending UERJ in the years in which the university had affirmative action. We find no significant earnings effects in the pre-AA cohorts (column B of Table 2), but in the 2004–2011 cohorts, UERJ enrollment reduced the early-career hourly wages of general track applicants by eight percent (column D). Similarly, the 2SLS RD estimate for 2004–2011 general applicants' monthly earnings is −153 USD, and this estimate is statistically significant at $p < 0.05$.

The evidence for a negative return is not conclusive because both of these coefficients decline in magnitude in the later time period and are not statistically distinguishable from zero (Panel C). Nonetheless, these results suggest that the returns to attending UERJ for non-AA students may have been lower in the cohorts with affirmative action.

3.2. **College selectivity and major choice.** To interpret our earnings results, it is important to understand the types of college programs that UERJ enrollees would have attended if they were not admitted. While UERJ is an elite school, it exists in a highly-competitive market. The federal university in Rio de Janeiro, UFRJ, is even more selective than UERJ for most programs. UFRJ ranked 3rd in a 2012 national ranking by the newspaper *Folha*, while UERJ ranked 11th. There are three other selective federal universities in the Rio suburbs, and more than five private universities in the city itself (see Appendix Table A8). UERJ applicants in the general and affirmative action tracks differed in the likelihood that they could gain admission to these other colleges because UERJ was the only school in Rio with affirmative action during most of our sample period. Further, applicants to a particular UERJ program may have also chosen a different major if they enrolled in another school.

We examine effects on college and major choice using data from Brazil's higher education census (INEP, 2019), which covers all colleges in the country. We do not have access to ID numbers in this dataset, so we match it to our sample of UERJ applicants using exact day of birth, gender, and year of enrollment. These variables do not uniquely identify individuals, so we define our dependent variables as the *total* number of students at a particular university or major that have the same birthdate, gender, and enrollment year as the UERJ applicant. In addition, we can only include 2009–2011 UERJ applicants in this analysis because individual-level census data does not exist prior to 2009. The fuzzy merge and smaller sample reduce the precision of our RD estimates for this analysis.[16]

With these caveats, we find evidence that UERJ's affirmative action policy allowed Black and public school applicants to attend a more selective college. Panel A of Table 3 displays θ coefficients from our reduced-form RD specification (1), which estimates the effects of *admission* to UERJ. The number of UERJ enrollees in the census data increases by 0.88 at the affirmative action thresholds (column F), which is broadly similar to our first stage estimate of 0.69 in Table 2. We do not find effects on enrollment in UFRJ, other federal universities in Rio, or private universities in the top 100 of the *Folha* ranking. Instead, we find that the number of enrollees in lower-ranked Rio universities falls by roughly 0.5 at the affirmative action thresholds. Although these estimates are imprecise, they match our prior that many affirmative action applicants would not have gained admission to other top universities, and thus likely had less-selective private schools as their fallback option.

_____

[16] Appendix B.5 gives details on our merge with the higher education census.

Admission to UERJ also significantly affected the major choices of affirmative action applicants. In the last two rows of Panel A, our dependent variables measure the total number of enrollees in Rio de Janeiro universities with the same major as the one that the UERJ applicant applied to. The total number of Rio enrollees with the same major as the applicant increases by 0.35 at the affirmative action thresholds using 2-digit major codes, and it increases by 0.46 using 3-digit majors codes. These coefficients are about half of the size of the RD estimate for the total number of UERJ enrollees (0.88), which suggests that roughly half of affirmative action compliers would have chosen a different major if they were not admitted to UERJ. These changes in field of study are an important mechanism for our RD earnings estimates, but such changes are relevant for the evaluation of affirmative action policies in any context where individuals may choose different majors at different schools.[17]

For general applicants, we find that admission to UERJ reduced the likelihood of enrolling in other top federal and private universities in Rio (Table 3, Panel A, column D). Thus most general track compliers would likely have attended other selective universities if they had been rejected. As in the affirmative action tracks, we also find that many general track compliers would have chosen different majors if they were not admitted.

3.3. **Educational attainment.** Another key question for understanding our earnings results is whether UERJ enrollment affected the likelihood that individuals earned *any* college or postgraduate degree. To examine this, we use the RAIS data to define three binary measures of educational attainment: 1) a college degree during the period of 6–9 years after UERJ application; 2) a college degree by 2019; and 3) a postgraduate degree by 2019. Panel B of Table 3 shows 2SLS RD estimates for these three outcomes using regression samples that include all applicants who appear in the RAIS data.[18]

We find no effects on overall educational attainment for both affirmative action and general applicants. Most notably, UERJ enrollment did not significantly affect the likelihood that affirmative action applicants earned a college or postgraduate degree (column F). 71 percent of marginally-rejected Black and public school applicants earned a college degree by 2019 (column E), which is a high rate even by the standards of developed countries. This reflects the fact that UERJ's affirmative action applicants were very high-achieving, even though they were disadvantaged relative to general UERJ applicants. In the general tracks, the rates of college degree attainment were even higher (columns A and C), and we also find no effects on educational attainment (columns B and D).

---

[17] Our estimates of the rate of major switching are comparable to those in Ng and Riehl (2022), who examine admission to selective STEM programs at a university in Colombia.

[18] We find no evidence that the observable characteristics of UERJ applicants who appear in the RAIS change discontinuously at the admission thresholds (Appendix Table A1).

3.4. **Employment with UERJ alumni.** As a final potential mechanism, we consider the effects of networking with UERJ peers and alumni. Elite university networks can improve students' access to high-paying jobs through many channels (Rivera, 2016), including on-campus recruiting (Weinstein, 2022), referrals (Calvo-Armengol and Jackson, 2004), and school reputation (MacLeod and Urquiola, 2015).

To test for these network mechanisms, we use the RAIS data to define outcome variables that indicate when UERJ applicants obtained jobs at firms that are affiliated with other UERJ alumni. Specifically, consider a UERJ applicant i who applied to major m. We define applicant i as obtaining a job at an *alumni firm* if their firm employed another individual j who graduated from major m (the "alum"). Our simplest network outcome is an indicator equal to one if the applicant's firm ever hired another alum, but we define different versions of this variable based on the alum's characteristics, the timing of their employment, and the concentration of alumni at the firm. We use major-specific networks because students in the same program often take classes together and work in similar labor markets.[19]

Attending UERJ significantly increased the likelihood that affirmative action students obtained jobs at firms affiliated with other UERJ alumni. Panel A of Table 4 shows that marginal affirmative action enrollees were 13.7 percentage points more likely to work at a firm affiliated with any UERJ alum in the period of 6–9 years after application (see also Panel E of Figure 3). This is a 29 percent increase from the mean rate of employment at alumni firms for marginally-rejected applicants (47.7 percent). Affirmative action applicants were more likely to work with both general and AA alumni when they attended UERJ, and the two effects are similar in magnitude (second and third rows of Panel A). Similarly, attending UERJ increased the *proportion* of UERJ alumni at affirmative action applicants' firms by 0.85 percent, or 8.5 alumni per 1000 workers at the firm (fourth row of Panel A). We also find large and significant effects on early-career employment at alumni firms for general track applicants (columns B and D).

Figure 4 presents evidence that the results in Table 4 are driven at least in part by networking mechanisms. Since major choices changed at the UERJ admission thresholds (Table 3), it is possible that the alumni firm employment effects reflect major-specific human capital accumulation rather than networking. To examine this possibility, Figure 4 displays heterogeneity in RD estimates for the number of UERJ alumni per 1000 workers at the

---

[19] All of our network outcomes are leave-individual-out; even if an applicant completed a UERJ degree, these variables equal one only if there is *another* alum affiliated with that firm. We define firms at the establishment level (organization × branch) using 14-digit CNPJ codes. Our variable definitions allow applicants to be both beneficiaries and benefactors of UERJ's alumni network; for example, an applicant could work at an alumni firm if they got a job from the alum's referral, or they themselves referred the alum to the firm. Our outcomes are similar in spirit to Zimmerman (2019)'s measures of co-leadership rates, but we consider firm-level employment rather than leadership positions.

applicant's firms (pooling across all applicant groups). We find larger estimates in cases where one would expect networking to be more important. The RD coefficients are larger for employment in small private firms than in large public firms.[20] Similarly, the employment effects are largest for alumni from the same cohort as the applicant, and for alumni who work at the firm at the same time as the applicant. This variation is consistent with referral and recruiting mechanisms, and it is hard to reconcile with major-specific human capital mechanisms. Appendix Table A10 shows that enrolling in UERJ increased access to firms with UERJ alumni even *within* groups of firms that operate in the same location and industry. In other words, the presence of UERJ alumni at a firm is a strong predictor of an applicant's employment outcome even among firms in the same narrowly-defined labor market.

Mean wages at firms affiliated with UERJ alumni were 0.44 log points higher than those at other firms in our sample (Appendix Table A12), suggesting that affirmative action enrollees benefited from increased access to these firms.[21] Consistent with this, we find that attending UERJ increased the mean wage at affirmative action compliers' early-career firms by 0.11 log points (last row of Panel A, column F), which is only slightly smaller than the magnitude of their early-career earnings gain (0.13 log points). Notably, UERJ enrollment *reduced* the average wage at post-AA general applicants firms' by 0.10 log points (column D), consistent with the negative earnings effects for these applicants in Table 2.

Yet Panel B of Table 4 shows that the benefits of accessing UERJ's alumni network decreased as individuals' careers progressed. This panel shows RD estimates for the same alumni firm outcomes as in Panel A, but we measure employment 10–13 years after UERJ application. For all outcomes and all applicant groups, the effects of UERJ enrollment are smaller in this later period, and many are not statistically different from zero. This suggests that alumni networks are most important for initial job placement, and that their influence declines as individuals form new networks in the labor market.

3.5. **Discussion.** Our graduation results show that most affirmative action students succeeded academically at UERJ. Related work argues that affirmative action may cause disadvantaged students to drop out or switch majors—particularly in STEM fields—because it places them in schools where they are less-prepared than their classmates (Arcidiacono et al., 2016). UERJ graduation rates are high by Brazilian standards, and most programs in our RD sample are in non-STEM fields (Table 1, Panel A). Thus relative academic preparation

---

[20] Networking is likely more important at small private firms because most public firms in Brazil use exams to hire workers (Mocanu, 2022).

[21] Appendix Table A9 provides examples of alumni firms. Firms with the most UERJ alumni were large public entities like Rio de Janeiro City Hall and the State Secretary of Education. Firms with the highest alumni concentration include financial organizations like Accenture and the Brazilian Development Bank, as well as branches of the multinational petroleum company Petrobras.

may be less important for degree completion in our setting.[22] On the other hand, we do not find that affirmative action increased the likelihood that disadvantaged students earned a college degree, as other work has found (Bleemer, 2022). One possibility is that any negative effects of mismatch in academic preparation were offset by positive effects of UERJ's greater resources, and the net effect on degree attainment was zero.

Our findings suggest that the early-career earnings gains for affirmative action students were driven partly by networking mechanisms. UERJ's affirmative action policy increased disadvantaged students' access to higher-paying firms affiliated with its alumni (Table 4). Taken together with the OLS wage premium for alumni firms (0.44 log points), the RD estimate for alumni firm employment (14pp) can explain nearly half of the early-career wage gains for affirmative action students (0.13 log points). Our paper differs somewhat from Zimmerman (2019) and Michelman et al. (2022), which find that the benefits of networking at elite universities accrue *only* to students from advantaged backgrounds. This difference can be explained by the fact that our data allows us to examine early-career networking in a broader set of firms, but our results are unclear about whether these initial networking benefits are persistent. Further, our estimate of the early-career return for affirmative action students is substantially smaller than that in Bleemer (2022) (see footnote 3), which may be because we find no effects on college degree attainment.

For general track enrollees, we find some evidence that the early-career return to attending UERJ declined from the pre-AA to post-AA cohorts (Table 2, Panel B). This suggests that there may have been negative spillover effects of UERJ's affirmative action policy on other students. In the post-AA cohorts, marginally-admitted general applicants were less likely to work at high-paying firms, and they were more likely to work with UERJ's affirmative action alumni (Table 4). This suggests that affirmative action may have reduced the value of networking at UERJ because the new disadvantaged students tended to obtain lower-paying jobs. But our RD analysis cannot conclusively identify any spillover effects of affirmative action because the policy also affected the characteristics of marginally-admitted general applicants.[23] To present more compelling evidence on the potential spillover effects of UERJ's policy, we now turn to our second empirical strategy.

4. DD specification

4.1. **Top enrollee sample.** To estimate the effects of affirmative action on other UERJ students, we construct a sample of *top enrollees* who could have attended UERJ regardless

---

[22] Appendix Table A7 shows that the early-career earnings gains for affirmative action students was driven by UERJ's health and business programs, which also have high graduation rates. In UERJ's STEM programs, affirmative action students graduated at much lower rates, and we find no evidence of positive returns.

[23] Appendix Table A2 shows that post-AA general track compliers were more likely to be non-white and younger than pre-AA compliers, although the magnitude of these differences is relatively modest.

of whether affirmative action existed in their cohort. For each major m, we define N_m to be the minimum number of students who enrolled through the *general* track in any cohort in 1995–2011.[24] Our top enrollee sample includes the N_m enrollees with the highest admission scores in each cohort. This is a balanced panel at the major level that includes N_m enrollees from every cohort. This sample roughly corresponds to the top half of students in each program/cohort, as 55 percent of slots were reserved for general applicants.[25]

4.2. **Regression model.** For identification, we exploit variation in the take-up of affirmative action across UERJ's majors, as depicted in Figure 1. We use this variation in a difference-in-differences (DD) specification:

$$(3) \qquad Y_{imc} = \gamma_m + \gamma_{cf(m)} + \pi[\text{ExposureToAA}_m \times \text{Post}_c] + \varepsilon_{imc}.$$

Y_{imc} is an outcome for individual i who enrolled in major m and cohort c. Our variable of interest is the interaction between a major's exposure to affirmative action and a dummy for post-AA cohorts ($\text{ExposureToAA}_m \times \text{Post}_c$). Our benchmark results use a binary measure of exposure that equals one if the share of affirmative action enrollees in the 2004–2011 cohorts was 30 percent or higher (the horizontal line in Figure 1). We include major and cohort fixed effects, and cluster standard errors at the major level.

We estimate equation (3) in our sample of top enrollees to examine the effects of UERJ's affirmative action policy on untargeted students. In this case, the π coefficient measures how affirmative action changed top enrollees' outcomes in more-affected majors relative to less-affected majors. We refer to these π coefficients as "spillover" effects because they reflect the impacts of affirmative action students' enrollment on top enrollees' outcomes (Arcidiacono and Vigdor, 2010). We also present DD coefficients for the small subset of top enrollees who are from underrepresented minority (URM) groups; these estimates provide evidence on whether affirmative action affected the outcomes of URM students who would have been admitted to UERJ even in absence of the policy.

Panel A of Table 5 shows that our DD specification identifies the effects of a 19 percentage point increase in the fraction of enrollees in an individual's program/cohort who were admitted through affirmative action. This is a large effect on diversity relative to the scale of affirmative action at many elite US universities, but it is similar to the magnitude of Chetty et al. (2020)'s "need-affirmative" counterfactual admission policy.

4.3. **Identification assumptions.** Our key identification assumption is that the outcomes of enrollees in majors with more and less exposure would have followed parallel trends in

---

[24] In other words, we compute the number of general track enrollees in each major m and cohort c, N_{mc}, and then define $N_m = \min_{c \in \{1995,\dots,2011\}} N_{mc}$. On average, N_m is 53 percent of the total number of enrollees.
[25] Our top enrollee sample includes any affirmative action students who were among the top N_m in their cohort. But in practice, 93 percent of top enrollees were admitted through the general track (Table 1).

the absence of affirmative action. A potential concern is that Brazil experienced a recession in the mid-2010s, which may have had heterogeneous impacts across UERJ's majors. To address this, we interact the cohort dummies in equation (3), γ_c, with fixed effects for the field of study groups, $f(m)$, shown in Table 1 (business, health, humanities, natural sciences, and social sciences). This restricts identification to comparisons between majors in the same field, which were likely to be similarly affected by macroeconomic conditions.

Appendix Figure A6 provides support for parallel trends by showing that mean wages evolved similarly in industries that typically hired students from UERJ majors with more and less exposure to affirmative action. For this figure, we first compute the mean hourly wage in each industry × year pair using all workers in the RAIS data. We then compute a weighted average of these industry × year means for each UERJ major, where the weights are the share of pre-AA top enrollees who were employed in each industry. We find that industry mean wages trended similarly between majors with more and less exposure to affirmative action across all years of our data. In the years in which post-AA graduates entered the labor market (2009–2019), the change in industry mean wages between more- and less-affected majors is small and statistically insignificant (-0.02 log points). This suggests that our earnings results below are not driven by different growth rates across industries or by heterogeneous impacts of the mid-2010s recession. Below we also present event study and robustness results to test the validity of our identification strategy.

5. SPILLOVER EFFECTS OF AFFIRMATIVE ACTION

5.1. **Characteristics of UERJ enrollees.** We begin our DD analysis by asking whether affirmative action affected the composition of UERJ's top enrollees. Research finds that families prefer schools with high-achieving peers (Abdulkadiroğlu et al., 2020), and thus UERJ's policy may have induced some students to attend other colleges. To test for these effects, Table 5 uses the demographic characteristics and entrance exam scores of UERJ enrollees as dependent variables in our DD specification (3). Column (A) shows the dependent variable mean for top enrollees in the pre-AA cohorts (1995–2001). Our main results are the DD coefficients, π, for top enrollees in column (B). Column (C) shows DD estimates for URM top enrollees, which we define as top enrollees who identify as Black or indigenous in the RAIS. Column (D) shows DD estimates for students who are not top enrollees.

We do not find significant effects of exposure to affirmative action on observable characteristics of top enrollees. Column (B) in Table 5 shows that the DD coefficients for top enrollees' age, gender, and race are small and statistically insignificant (Panel B). We find no effects on top enrollees' field exam scores or their overall admission score (Panel C).[26] In Panel D, we combine the characteristics and admission scores into a single index using the

---

[26] We standardize scores to mean zero and SD one in the population of all UERJ enrollees in a given cohort.

predicted values from a log wage regression. The DD estimates for these predicted wages are small and insignificant, and they are similar whether or not we restrict to enrollees who appear in the RAIS data. Thus, the composition of top enrollees in more- and less-affected majors did not diverge significantly with the adoption of affirmative action.[27]

A possible explanation for this finding is that prospective students may not have known that the take-up of affirmative action would differ across UERJ's majors. Students were surely aware of the admission policy, but our DD analysis nets out school-level changes in the characteristics of top enrollees. Before enrolling, it may have been hard to know that the share of affirmative action students would be, for example, 15 percentage points lower in Economics than in Business on average. Thus affirmative action may have deterred some students from enrolling in any major at UERJ, but Table 5 shows that compositional changes are unlikely to explain a significant portion of our DD results for labor market outcomes.

By contrast, affirmative action had a large impact on the demographics and test scores of non-top enrollees (column D of Table 5). In majors with greater exposure to affirmative action, the student bodies became more racially diverse, older, and lower-ability as measured by entrance exam scores. This reflects the admission preference given to affirmative action students and the intended effects on diversity.

5.2. **Labor market outcomes.** Our main finding is that greater exposure to affirmative action reduced the post-college earnings of top enrollees. Table 6 has the same structure as Table 5, but it presents DD estimates for graduation and labor market outcomes measured 6–9 years after application. We find that UERJ's policy reduced the mean hourly wage of top enrollees by 14 percent in more-affected majors relative to less-affected majors (column B, Panel B). The DD estimate is similar in magnitude using mean monthly earnings as the dependent variable (-170 USD). Panel A of Figure 5 shows an event-study version of this result. The log hourly wage coefficient for top enrollees (red line) drops sharply between the last pre-AA cohort (2001) and the first post-AA cohort (2004). The wage coefficient declines over subsequent cohorts, reaching -0.20 log points by the 2011 cohort. These negative effects persist when we measure outcomes 10–13 years after application, and the magnitudes of the early- and later-career earnings coefficients are similar (see Appendix Table A14).

The negative earnings effects were largely driven by a reduction in top enrollees' firm quality as measured by firm average wages. The DD estimate for log firm mean hourly wage is -0.095 for top enrollees (Panel B of Table 6), which is 70 percent of the magnitude of the individual wage coefficient. The event-study coefficients for firm average wage also decline sharply in the first cohort with affirmative action (Panel B of Figure 5). We find no effect of

---

[27] We also find no evidence of differential changes in the composition of top enrollees using using SES indices based on individuals' first and last names (Appendix Table A13).

exposure to affirmative action on the graduation rates of top enrollees (Panel A of Table 6), suggesting that the earnings effect is not driven by changes in educational attainment. The DD estimate for employment in any formal sector job is negative and marginally significant (-0.027), but it is relatively small compared to the mean formal employment rate (0.74).

Appendix Table A16 shows that our results for top enrollees are robust to a variety of specification checks. Our earnings estimates do not change significantly if we restrict our regression sample to pre-recession years, or if we include program-specific linear trends estimated in the pre-AA cohorts. Controlling for student demographics and entrance exam scores only slightly reduces the magnitudes of our coefficients, consistent with the small compositional effects identified in Table 5. We continue to find negative effects when we compare programs in the same quartile of selectivity (defined by the x-axis in Figure 1) and when we exclude field of study controls. Lastly, our results are similar when we use a continuous measure of our treatment variable, ExposureToAA$_m$, which is the share of all 2004–2011 enrollees who were from an affirmative action track (the y-axis in Figure 1).

Notably, we also find that the adoption of affirmative action reduced the earnings of top enrollees from URM groups (column C of Table 6). These estimates are imprecise because our top enrollee sample includes only about 1,600 Black and indigenous students across all programs and cohorts, but nonetheless we find negative and significant point estimates for both individual and firm average wages (Panel B). Similarly, we find large declines in earnings for non-top enrollees in more- vs. less-exposed majors (column D). The DD estimate for non-top enrollees' hourly wages is larger in magnitude than the predicted wage effect based on individual characteristics (-0.212 vs. -0.154). Thus spillover effects may have also reduced the wages of affirmative action students, although this evidence is only suggestive.

5.3. **Networking mechanisms.** To shed light on mechanisms for these spillover effects, we first consider whether affirmative action affected the types of jobs that UERJ students obtained through networking. Affirmative action may have caused some employers to forgo recruiting at UERJ or making offers to UERJ graduates because it reduced the *expected* ability of a UERJ student (MacLeod et al., 2017; Weinstein, 2018). In addition, Black and public school students tended to obtain lower-paying jobs than general track students (Table 2), and thus the value of referrals or information sharing from classmates likely declined in majors with a large proportion of affirmative action students.

To test for these mechanisms, Panels C–D of Table 6 use dependent variables that measure employment in firms affiliated with UERJ alumni from different cohorts and application tracks.[28] In Panel C, the outcome variables are indicators for employment at firms that hired

---

[28] These alumni firm variables are similar to those in our RD analysis except we define them to be non-overlapping. For example, in the second row of Panel C, the dependent variable equals one only if the firm did *not* hire a pre-AA alum. As in Table 4, we require that the applicant and alum are from the same major.

pre-AA alumni from the enrollee's program versus firms that hired *only* post-AA alumni. In Panel D, we examine employment at firms that hired a general track alumni from the enrollee's own cohort versus firms that hired *only* alumni from other cohorts or from the affirmative action tracks. These variables test whether affirmative action changed the types of firms that hired UERJ students (Panel C) and the peer connections that UERJ students used to obtain jobs (Panel D). Appendix Table A12 shows that firms that hired pre-AA and general track alumni paid significantly higher average wages than those that hired only post-AA or affirmative action alumni.

We find that affirmative action reduced top enrollees' employment rates at higher-paying alumni network jobs. Panel C of Table 6 shows that top enrollees' likelihood of employment at firms with pre-AA alumni declined by 5.5 percentage points in more-affected majors relative to less-affected majors (column B). This decline was offset by a 4.9 percentage point increase in the rate of employment at firms that hired only post-AA alumni. Similarly, Panel D shows that the likelihood of employment with same-cohort general track alumni declined by 9.8 percentage points for top enrollees in more- versus less-affected majors. Correspondingly, top enrollees in more-affected majors became relatively more likely to work at firms that hired only general track alumni from another cohort (+4.2pp) or only affirmative action alumni (+4.6pp). These findings show a shift in employment toward firms with lower average wages (Appendix Table A12), which suggests that the negative spillover effects of affirmative action on earnings can partly be explained by a decline in the value of networking.

5.4. **Learning mechanisms.** Affirmative action may also have affected top enrollees' earnings through human capital channels. For example, the average level of academic preparation of UERJ's student body was lower under affirmative action, which may have reduced the benefits of peer interactions or caused professors to teach less advanced material.

To test for learning mechanisms, we use data from Brazil's national college exit exam, the Enade (INEP, 2022). The Enade is a field-specific exam that has been administered every year since 2004, although each field is tested every three years on a staggered schedule. The government assigns ratings to each higher education program based in part on their students' Enade scores, so many universities ask students who are close to graduation to take the exam (Pedrosa et al., 2013). But the Enade is typically low stakes from the student's perspective, as most universities do not require students to pass the exam in order to graduate.

Table 7 examines how affirmative action affected the characteristics and performance of UERJ's Enade exam takers. This table presents difference-in-differences (DD) estimates that compare 2004–2015 exam takers at UERJ to those at other federal and state universities that did not have affirmative action during this period.[29] Column (A) shows the means of each

---

[29] See Appendix Table A17 for details on our Enade sample and the exam fields.

outcome in the 2004–2006 cohorts at UERJ; we define this as the pre-AA period since UERJ students who took the Enade in these years enrolled prior to 2003. Column (B) displays DD estimates for all exam takers, which are the coefficients on an indicator for UERJ interacted with an indicator for the post-AA cohorts (2007–2015).[30] We do not have the Enade data linked to our UERJ records at the individual level, so we cannot estimate this regression in our top enrollee sample. As an alternative, column (C) restricts the sample to white students who attended private high schools, who were not eligible for affirmative action. Column (D) presents results for non-white and/or public high school exam takers.

Panel A of Table 7 shows that affirmative action increased the diversity of UERJ's Enade exam takers, but we do not find compositional changes within the sample of white private school students. The first row shows that the mean number of UERJ exam takers per program/cohort (36 students) did not change significantly with affirmative action (column B), but there were 7.9 fewer white private student students on average (column C), and 9.1 more non-white and/or public school students (column D). We find significant increases in the racial and socioeconomic diversity of the average UERJ exam taker (column B). However, we do not find significant changes in gender, age, mother's education, or household income within the sample of white private school students (column C). When we combine all demographic characteristics into a single index using scores on the general component of the Enade, we find that the relative change in the composition of UERJ's white private school students is close to zero (last row of Panel A).

In Panel B of Table 7, we find that affirmative action decreased the Enade scores of UERJ students, including *within* the sample of white private school students. Enade scores are expressed as the proportion of correct answers, and the overall score is a weighted average of its field-specific and general components. For the average UERJ exam taker, the overall proportion of correct answers declined by 3.8 percentage points with the introduction of affirmative action (column B) from a pre-AA mean of 55 percent (column A). These declines are likely attributable, at least in part, to the policy's effects on student body diversity. Yet the mean overall score of UERJ's white private school students also declined by 2.2 percentage points (column C). This decline is sizable, as it is 15 percent of a standard deviation of

---

[30] Our DD specification for Table 7 is

$$(4) \qquad Y_{mjt} = \gamma_{mj} + \gamma_{mt} + \pi[\text{UERJ}_j \times \text{Post}_t] + \varepsilon_{mjt}.$$

Regressions are at the exam field (m) by institution (j) by year (t) level, with observations weighted by the number of exam takers. We include institution fixed effects, γ_{mj}, and year fixed effects, γ_{mt}, and we interact both with field dummies so that identification comes only from within-field comparisons. The coefficient of interest, π, is on an indicator for UERJ interacted with an indicator for the 2007–2015 cohorts, $\text{UERJ}_j \times \text{Post}_t$. We cluster standard errors at the institution level.

the full distribution of Enade scores (14.4pp). Similarly, Figure 6 shows scores declined by about two percentage points at the highest quantiles of UERJ's score distribution.[31]

These findings suggest that affirmative action reduced the learning of UERJ's top students. At high quantiles and in the white private school sample, the declines in Enade performance are not likely to be attributable to changes in the set of students who gained admission under affirmative action. We find no evidence of negative selection in the sample of white private school students, and all else equal, one would expect *positive* selection within this sample because the bar for admission was higher in the cohorts with affirmative action. Thus we take these results as evidence that the negative effects of affirmative action on top enrollees' earnings are partly driven by learning spillovers.[32]

5.5. **Discussion.** Our results show that UERJ's adoption of affirmative action reduced the earnings of its highly-ranked students. Our point estimates imply that a one percentage point increase in the affirmative action share led to a 0.7 percent decrease in the wages of top enrollees. Thus the negative effects on top students' earnings were large in the majors with the highest exposure to affirmative action. These spillover results are also consistent with our RD analysis, which found a negative early-career earnings return to UERJ enrollment for general applicants in the post-AA cohorts (Table 2).

Our analysis of mechanisms suggests that the negative spillover effects were driven by a combination of compositional, networking, and learning mechanisms. Although we do not find statistically significant effects on the characteristics of top enrollees, our point estimate for the log wage index in Panel D of Table 5 (-0.033) is 25 percent of our main estimate on log wages (-0.132). By combining the DD estimates for access to alumni firms (Panels C–D of Table 6) with the OLS wage premia associated with these jobs (Appendix Table A12), network mechanisms can explain 10–17 percent of the overall wage effect. We do not have the Enade exam scores linked to wages, but Reyes (2023) finds that a one percentage point increase in the proportion of correct answers on Brazil's national college entrance exam (ENEM) is associated with a 0.02 log point increase in early-career wages. Assuming that the relationship between correct answers and wages is the same on the Enade exam, the decline in overall scores for white private school students in Panel B of Table 7 (2.2pp) can explain 32 percent of the overall wage effect. Taken together, these compositional, networking, and learning effects can explain two-thirds of the decrease in top enrollees' hourly wages.[33]

---

[31] Figure 6 plots DD coefficients in which the dependent variables are quantiles of Enade scores within each exam field × institution × year cell (rather than mean scores, as in Table 7).

[32] Appendix Table A18 shows that the declines in UERJ's Enade scores are larger in the majors with more exposure to affirmative action, which is consistent with our empirical strategy for top enrollees.

[33] Appendix B.6 provides details on these back-of-the-envelope calculations.

Further, we find that UERJ's adoption of affirmative action reduced the earnings of highly-ranked URM students. URM students who would have gained admission in the absence of the policy were likely affected by the same networking and learning spillovers as other top enrollees, and they may also have faced statistical discrimination from employers when the URM share of the student body increased (Coate and Loury, 1993). Thus our results suggest that highly-ranked URM students may have been better off if UERJ's affirmative action policy had been smaller in scale.

6. Conclusion

This paper presented evidence of a tradeoff between the direct and spillover effects of affirmative action at the Rio de Janeiro State University (UERJ). On the one hand, marginally-admitted Black and low-income students who attended UERJ as a result of affirmative action experienced a 14 percent increase in early-career earnings. This earnings gain was driven not by educational attainment, but rather by increased access to high-paying firms affiliated with UERJ alumni. This suggests that the primary benefit of affirmative action at elite universities may be to help disadvantaged students' gain access to job networks in high-wage sectors of the economy. And yet we found that the earnings and networking benefits for affirmative action students may have decreased as their careers progressed, suggesting that they faced additional barriers to career advancement in the labor market.

On the other hand, UERJ's adoption of affirmative action had negative impacts on the careers of students who were not targeted by the policy, including high-achieving URM students. Our results suggest that a 19 percentage point increase in the share of the student body that was admitted through affirmative action led to a 14 percent decrease in the post-college wages of UERJ's highly-ranked students. This earnings effect was partly attributable to a change in the composition of UERJ's top students, but it was also driven by negative spillover effects on their learning and a decline in the value of networking with peers. These results can explain why elite schools around the world use admission policies that favor high-achieving and wealthy students (Arcidiacono et al., 2022), and why they may be hesitant to unilaterally adopt affirmative action at a large scale.

Our paper shows that elite universities face a tradeoff between serving as engines of upward mobility for disadvantaged students and maintaining sources of their value added that stem from having high-achieving and wealthy student bodies. An important caveat is that we do not examine non-pecuniary benefits of interacting with classmates from diverse backgrounds (e.g., Carrell et al., 2019), which can further justify the adoption of large-scale affirmative action at selective schools.

References

Abdulkadiroğlu, A., P. A. Pathak, J. Schellenberg, and C. R. Walters (2020). Do parents value school effectiveness? *American Economic Review 110*(5), 1502–39.

Andrews, R. J., J. Li, and M. F. Lovenheim (2016). Quantile treatment effects of college quality on earnings. *Journal of Human Resources 51*(1), 200–238.

Anelli, M. (2020). The returns to elite university education: A quasi-experimental analysis. *Journal of the European Economic Association 18*(6), 2824–2868.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association 91*(434), 444–455.

Arcidiacono, P., E. M. Aucejo, and V. J. Hotz (2016). University differences in the graduation of minorities in STEM fields: Evidence from California. *American Economic Review 106*(3), 525–562.

Arcidiacono, P., J. Kinsler, and T. Ransom (2022). Legacy and athlete preferences at Harvard. *Journal of Labor Economics 40*(1), 133–156.

Arcidiacono, P., M. Lovenheim, and M. Zhu (2015). Affirmative action in undergraduate education. *Annual Review of Economics 7*(1), 487–518.

Arcidiacono, P. and J. L. Vigdor (2010). Does the river spill over? estimating the economic returns to attending a racially diverse college. *Economic Inquiry 48*(3), 537–557.

Backes, B. (2012). Do affirmative action bans lower minority college enrollment and attainment? evidence from statewide bans. *Journal of Human Resources 47*(2), 435–455.

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

Bertrand, M., R. Hanna, and S. Mullainathan (2010). Affirmative action in education: Evidence from engineering college admissions in India. *Journal of Public Economics 94*(1), 16–29.

Black, S. E., J. T. Denning, and J. Rothstein (2023). Winners and losers? the effect of gaining and losing access to selective colleges on education and labor market outcomes. *American Economic Journal: Applied Economics 15*(1), 26–67.

Bleemer, Z. (2021). Top percent policies and the return to postsecondary selectivity. Research & Occasional Paper Series: CSHE.1.2021.

Bleemer, Z. (2022). Affirmative action, mismatch, and economic mobility after California's Proposition 209. *The Quarterly Journal of Economics 137*(1), 115–160.

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

Calvo-Armengol, A. and M. O. Jackson (2004). The effects of social networks on employment and inequality. *American Economic Review 94*(3), 426–454.

Canaan, S. and P. Mouganie (2018). Returns to education quality for low-skilled students: Evidence from a discontinuity. *Journal of Labor Economics 36*(2), 395–436.

Carrell, S. E., M. Hoekstra, and J. E. West (2019). The impact of college diversity on behavior toward minorities. *American Economic Journal: Economic Policy 11*(4), 159–82.

Carrell, S. E., B. I. Sacerdote, and J. E. West (2013). From natural variation to optimal policy? the importance of endogenous peer group formation. *Econometrica 81*(3), 855–882.

Chetty, R., J. N. Friedman, E. Saez, N. Turner, and D. Yagan (2020). Income segregation and intergenerational mobility across colleges in the united states. *The Quarterly Journal of Economics 135*(3), 1567–1633.

Coate, S. and G. C. Loury (1993). Will affirmative-action policies eliminate negative stereotypes? *American Economic Review 83*(5), 1220–40.

Cortes, K. E. (2010). Do bans on affirmative action hurt minority students? evidence from the Texas Top 10% Plan. *Economics of Education Review 29*(6), 1110–1124.

Dale, S. B. and A. B. Krueger (2002). Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *The Quarterly Journal of Economics 117*(4), 1491–1527.

Daniel, K., D. A. Black, and J. Smith (2001). Racial differences in the effects of college quality and student body diversity on wages. In *Diversity Challenged: Evidence on the Impact of Affirmative Action*, pp. 221–231. Cambridge, Harvard Publishing Group.

Dillon, E. W. and J. A. Smith (2020). The consequences of academic match between students and colleges. *Journal of Human Resources 55*(3), 767–808.

Duflo, E., P. Dupas, and M. Kremer (2011). Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review 101*(5), 1739–1774.

Durlauf, S. N. (2008). Affirmative action, meritocracy, and efficiency. *Politics, Philosophy & Economics 7*(2), 131–158.

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

Ferman, B. and J. Assunção (2005). Affirmative action in university admissions and high school students' proficiency. *XXVii Encontro Brasileiro de Econometria. Anais*.

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

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.

Hastings, J. S., C. A. Neilson, and S. D. Zimmerman (2013). Are some degrees worth more than others? Evidence from college admission cutoffs in Chile. National Bureau of Economic Research Working Paper 19241.

Hinrichs, P. (2011). The effects of attending a diverse college. *Economics of Education Review 30*(2), 332–341.

Hinrichs, P. (2012). The effects of affirmative action bans on college enrollment, educational attainment, and the demographic composition of universities. *Review of Economics and Statistics 94*(3), 712–722.

Hoekstra, M. (2009). The effect of attending the flagship state university on earnings: A discontinuity-based approach. *The Review of Economics and Statistics 91*(4), 717–724.

Hoxby, C. M. (2018). The productivity of U.S. postsecondary institutions. In C. M. Hoxby and K. Stange (Eds.), *Productivity in Higher Education*. National Bureau of Economic Research.

Hoxby, C. M. and C. Avery (2013). Missing one-offs: The hidden supply of high-achieving, low-income students. *Brookings Papers on Economic Activity.*

INEP (2019). Microdata for Higher Education Census (Censo da Educação Superior). Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Brasília, Brazil. https://www.gov.br/inep/ (accessed June 2019).

INEP (2022). Microdata for Enade exam (Exame Nacional de Desempenho dos Estudantes). Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Brasília, Brazil. https://www.gov.br/inep/ (accessed May 2022).

Júnior, J. F. and V. T. Daflon (2014). Políticas da igualdade racial no ensino superior. *Cadernos do Desenvolvimento Fluminense* (5), 31–44.

Kapor, A. (2015). Distributional effects of race-blind affirmative action. Working Paper.

Kirkebøen, L., E. Leuven, and M. Mogstad (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics 131*(3), 1057–1111.

Machado, C. and C. Szerman (2021). Centralized college admissions and student composition. *Economics of Education Review 85*, 102184.

MacLeod, W. B., E. Riehl, J. E. Saavedra, and M. Urquiola (2017, July). The big sort: College reputation and labor market outcomes. *American Economic Journal: Applied Economics 9*(3), 223–261.

MacLeod, W. B. and M. Urquiola (2015). Reputation and school competition. *American Economic Review 105*(11), 3471–3488.

Marmaros, D. and B. Sacerdote (2002). Peer and social networks in job search. *European economic review 46*(4-5), 870–879.

Mayer, A. and S. L. Puller (2008). The old boy (and girl) network: Social network formation on university campuses. *Journal of Public Economics 92*(1-2), 329–347.

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

Mello, U. (2022). Centralized admissions, affirmative action, and access of low-income students to higher education. *American Economic Journal: Economic Policy 14*(3), 166–97.

Michelman, V., J. Price, and S. D. Zimmerman (2022). Old boys' clubs and upward mobility among the educational elite. *The Quarterly Journal of Economics 137*(2), 845–909.

Mocanu, T. (2022). Designing gender equity: Evidence from hiring practices and committees. Working Paper.

Mountjoy, J. and B. Hickman (2020). The returns to college(s): Estimating value-added and match effects in higher education. University of Chicago, Becker Friedman Institute for Economics Working Paper.

Ng, K. and E. Riehl (2022). The returns to STEM programs for less-prepared students. Working Paper.

Otero, S., N. Barahona, and C. Dobbin (2021). Affirmative action in centralized college admission systems: Evidence from Brazil. Working paper.

Pedrosa, R. H., E. Amaral, and M. Knobel (2013). Assessing higher education learning outcomes in Brazil. *Higher Education Management and Policy 24*(2), 55–71.

RAIS (2021). Relação Anual de Informações Sociais. Ministério do Trabalho e Emprego, Brasília, Brazil. http://www.rais.gov.br/ (accessed May 2021).

Reyes, G. (2023). Cognitive endurance, talent selection, and the labor market returns to human capital. arXiv preprint arXiv:2301.02575.

Ribeiro, A. C. T. (2016). Affirmative action outcomes: evidence from a law school in Brazil. Department of Economics, FEA-USP Working Paper Series No. 2016-43.

Ribeiro, A. T. and F. Estevan (2021). Long-term outcomes of affirmative action: Evidence from a selective law school in Brazil. Working paper.

Riehl, E. (2023). Do less informative college admission exams reduce earnings inequality? evidence from Colombia. *Journal of Labor Economics*. Forthcoming.

Rivera, L. A. (2016). *Pedigree: How Elite Students Get Elite Jobs*. Princeton University Press.

Saavedra, J. E. (2009). The returns to college quality: A regression discontinuity analysis. Harvard University.

Sekhri, S. (2020). Prestige matters: Wage premium and value addition in elite colleges. *American Economic Journal: Applied Economics 12*(3), 207–25.

Smith, J., J. Goodman, and M. Hurwitz (2020). The economic impact of access to public four-year colleges. National Bureau of Economic Research.

UERJ (2020a). Applicants to UERJ. Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil. https://www.uerj.br/ (accessed January 2020).

UERJ (2020b). Enrollees in UERJ. Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil. https://www.uerj.br/ (accessed January 2020).

Vieira, R. S. and M. Arends-Kuenning (2019). Affirmative action in Brazilian universities: Effects on the enrollment of targeted groups. *Economics of Education Review 73*, 101931.

Weinstein, R. (2018). Employer screening costs, recruiting strategies, and labor market outcomes: An equilibrium analysis of on-campus recruiting. *Labour Economics 55*, 282–299.

Weinstein, R. (2022). Firm decisions and variation across universities in access to high-wage jobs: Evidence from employer recruiting. *Journal of Labor Economics 40*(1), 1–46.

Zhu, M. (2023). Beyond the "old boys' network": Social networks and job finding at community colleges. *Journal of Human Resources*. Forthcoming.

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

Zimmerman, S. D. (2019). Elite colleges and upward mobility to top jobs and top incomes. *American Economic Review 109*(1), 1–47.

Zoninsein, J. and J. F. Júnior (2008). *Ação afirmativa no ensino superior brasileiro*. Editora UFMG.
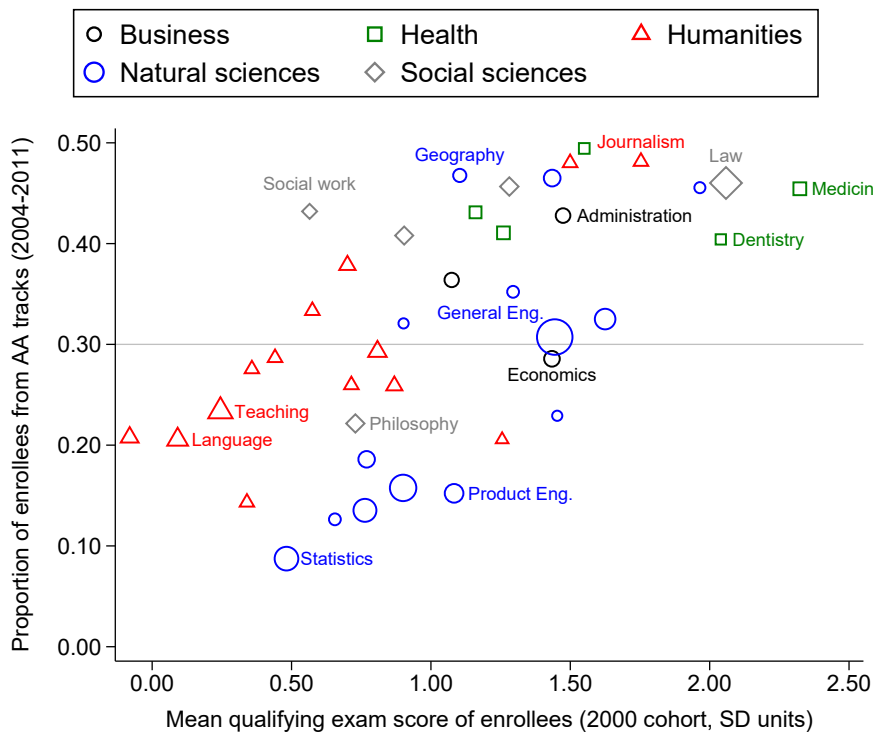
Figure 1. Take-up of affirmative action and program selectivity

Notes: This figure plots exposure to affirmative action (*y*-axis) and selectivity (*x*-axis) for each UERJ program in our sample. The *y*-axis displays the fraction of enrollees in the 2004–2011 cohorts who entered through an affirmative action track. The *x*-axis displays the mean score on the 2000 qualifying exam for students who enrolled in each program. We compute each applicant's average score across subjects in the exam, and standardize the scores to mean zero and SD one in the entire population of qualifying exam takers. The figure does not display two programs in our sample for which we do not have scores in the 2000 qualifying exam (mechanical engineering and production engineering). Marker sizes are proportional to the number of enrollees.
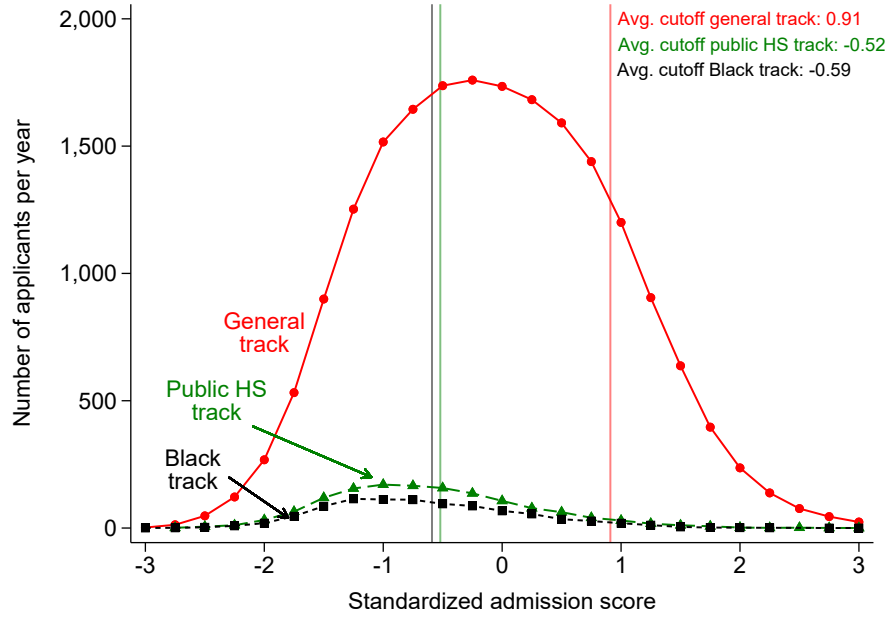
FIGURE 2. Admission score distribution and mean cutoff by application track (2004–2011)

Notes: This figure shows the distribution of standardized admission scores for applicants in each application track. The sample includes the 24 programs in our RD sample (Panel A of Table 1). We standardize scores to be mean zero and SD one in the population of all applicants in the same program/cohort, and plot distributions in 0.25 SD bins of the standardized score. Vertical lines represent the (applicant-weighted) average admission cutoff in each track. The cutoffs are equal to the standardized scores of the last admitted students.
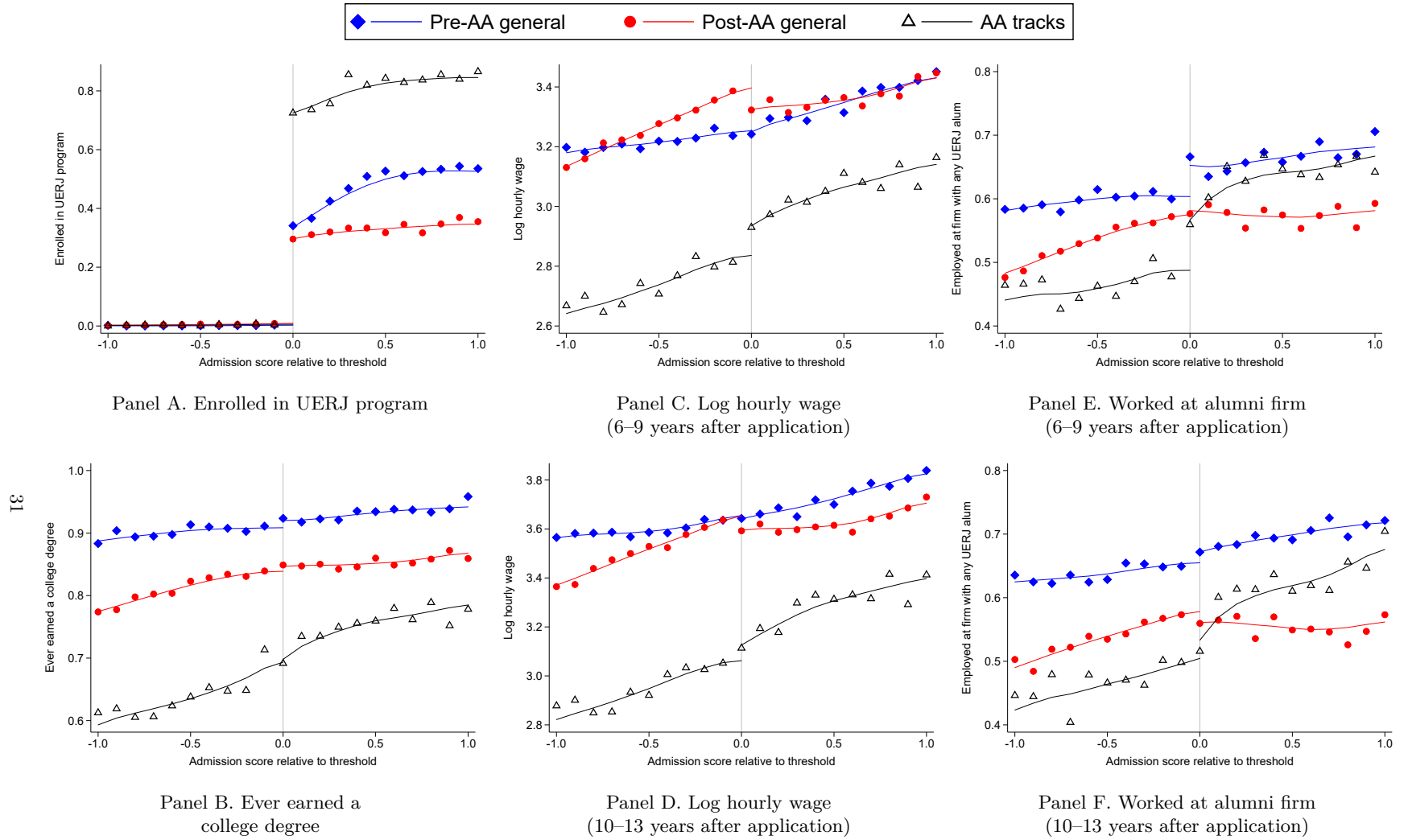
FIGURE 3. RD graphs for UERJ enrollment, earnings, and employment at alumni firms

Notes: This figure presents RD graphs for pre-AA general applicants (blue diamonds), post-AA general applicants (red circles), and Black/public school applicants (black triangles). The x-axis in each panel is an applicant's standardized admission score normalized to zero at the cutoff. The variable on the y-axis of each graph is listed in the panel title. Outcomes in Panels C and E are measured 6–9 years after individuals applied to UERJ. Outcomes in Panels D and F are measured 10–13 years after application. Markers depict means in 0.1 SD bins of the standardized score. Lines are predicted values from local linear regressions estimated separately above and below the threshold with a triangular kernel.
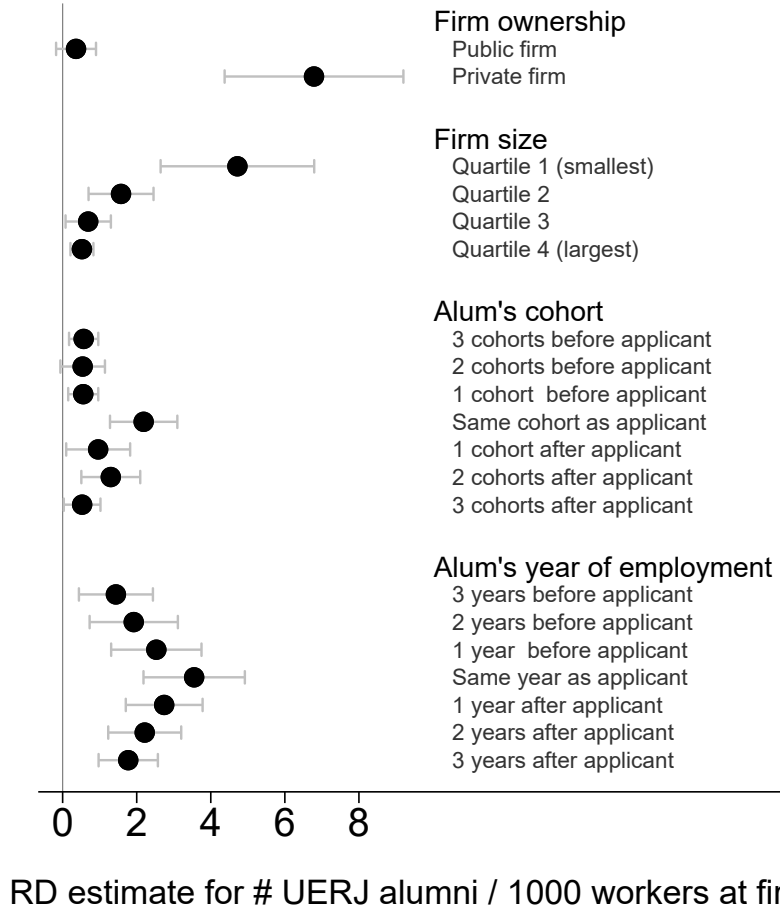
FIGURE 4. Heterogeneity in RD estimates for alumni per 1000 workers at the firm

Notes: This figure displays RD estimates for the effects of UERJ enrollment on the mean number of alumni per 1000 workers at the applicants' firms measured 6–9 years after application. These estimates are analogous to the 2SLS RD coefficients in the fourth row of Table 4 (Panel A), except we estimate equation (2) pooling across all applicant groups. We examine heterogeneity using four types of dependent variables:

- Firm ownership: The firm's mean number of alumni per 1000 workers interacted with dummies for public and private firms;
- Firm size: The firm's mean number of alumni per 1000 workers interacted with dummies for quartiles of firm size (number of workers);
- Alum's cohort: The firm's mean number of alumni per 1000 workers computed separately using alumni who enrolled in UERJ in each cohort from 3 years before to 3 years after the applicant's cohort;
- Alum's year of employment: The firm's mean number of alumni per 1000 workers computed separately using alumni who worked at the firm in each year from 3 years before to 3 years after the applicant.

Markers depict the RD coefficients, β, coefficients from these regressions, and horizontal bars are 95 percent confidence intervals using standard errors clustered at the individual level.

Panel A. Log hourly wage   Panel B. Firm mean hourly wage (log)
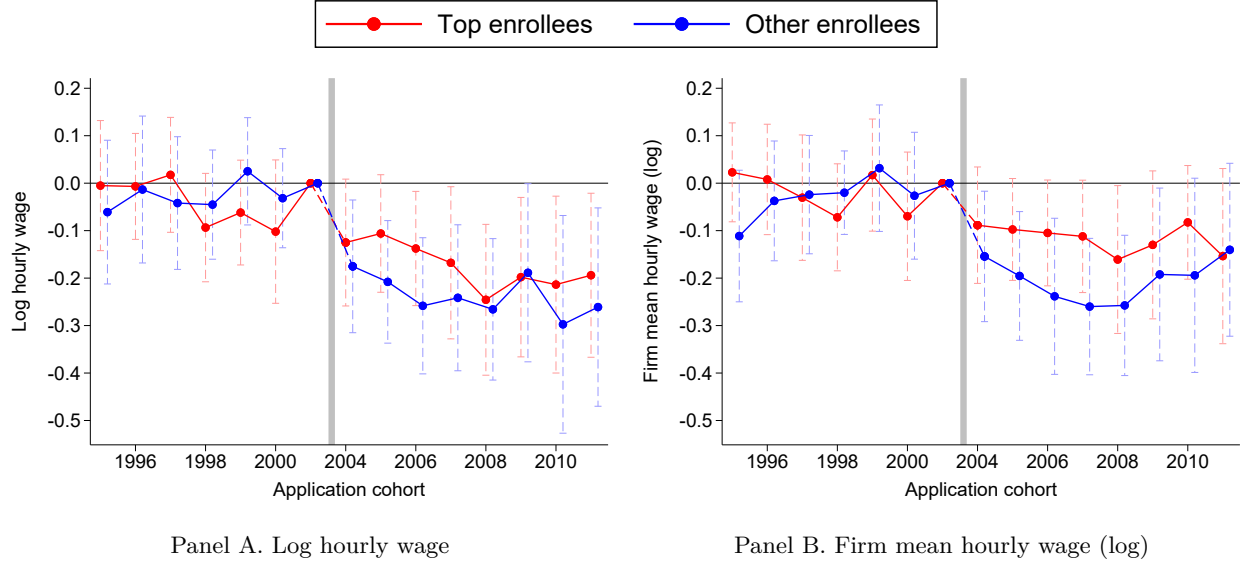
FIGURE 5. Event study estimates for individual and firm mean hourly wages 6–9 years after application

Notes: This figure plots coefficients from an event-study version of our DD regression (3). These are π_c coefficients that we estimate by replacing Post$_c$ with dummies for each cohort (omitting 2001). Dashed lines are 95% confidence intervals using standard errors clustered at the program level. The dependent variables are log hourly wage (Panel A) and firm mean log hourly wage (Panel B), each measured 6–9 years after application. We estimate regressions separately for top enrollees (red markers) and other enrollees (blue markers).
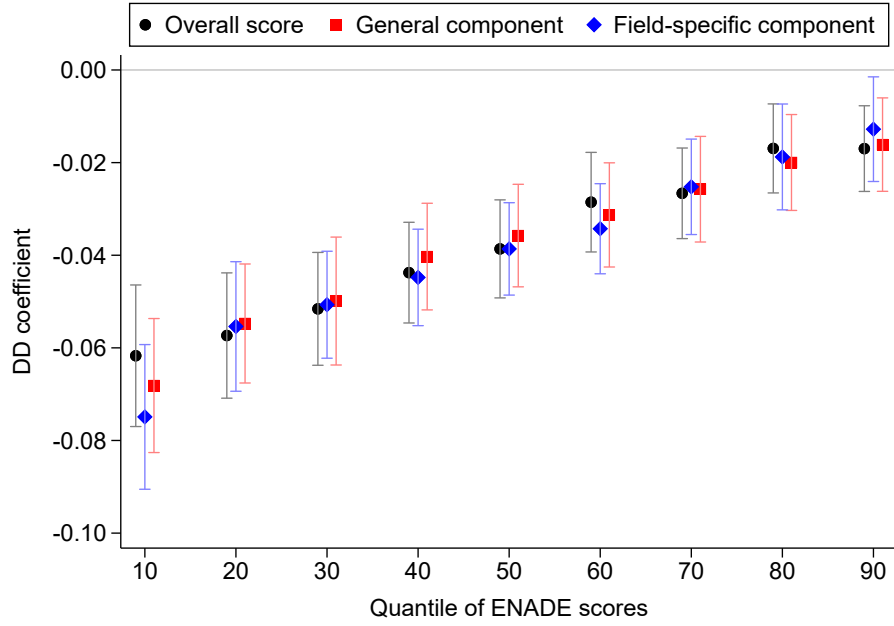
FIGURE 6. Effects of affirmative action at UERJ on quantiles of Enade scores

Notes: This table displays DD estimates of the effect of UERJ's affirmative action policy on quantiles of its graduates Enade exam scores. We estimate the same DD regressions as in Panel B of Table 7, but the dependent variables are quantiles of Enade scores within each institution × program × cohort cell (rather than mean Enade scores). Markers represent the DD coefficients, with the quantile depicted on the x-axis and the coefficient value on the y-axis. Vertical bars are 95 percent confidence intervals using standard errors clustered at the institution level.

TABLE 1. Summary statistics for RD and DD samples

	(A)	(B)	(C)	(D)	(E)

Panel A. Programs in both RD and DD samples

	1995–2001	2004–2011 cohorts			
Sample sizes and characteristics of all applicants	General track	General track	Public high school	Black	Other AA
Total applicants	95,659	159,408	10,996	7,263	318
Applicants in RD sample	93,930	159,383	9,624	5,600	0
Enrollees in DD sample	15,512	11,588	4,465	3,241	211
Top enrollees in DD sample	7,932	8,922	362	178	2
Female	0.50	0.55	0.60	0.60	0.48
Age at applicantion	20.75	20.28	21.88	23.04	24.30
White (UERJ data)		0.64	0.49	0.03	0.35
White (RAIS data)	0.78	0.67	0.57	0.15	0.48
Mother has a high school degree		0.85	0.49	0.56	0.54
HH income > 1.5× min. wage		0.82	0.35	0.35	0.45

Included programs (24 in total):

<u>Business</u>: Accounting, Business Administration.

<u>Health</u>: Biological sciences, Dentistry, Medicine, Nursing, Nutrition.

<u>Humanities</u>: Greek/Latin/Literature, History Ed. (SGO), Journalism, Psychology.

<u>Natural sciences</u>: Chemical Engineering, Chemistry, Computer Science, General Engineering, Geography, Geology, Industrial Design, Mechanical Engineering, Production Engineering.

<u>Social sciences</u>: History, Law, Social Science, Social Work.

Panel B. Programs in DD sample only

	1995–2001	2004–2011 cohorts			
Sample sizes and characteristics of all applicants	General track	General track	Public high school	Black	Other AA
Total applicants	47,633	50,553	4,374	2,118	58
Applicants in RD sample	0	0	0	0	0
Enrollees in DD sample	13,765	14,105	2,469	1,326	38
Top enrollees in DD sample	8,534	9,179	495	253	9
Female	0.56	0.53	0.62	0.63	0.57
Age at applicantion	22.34	21.62	22.54	24.09	26.24
White (UERJ data)		0.59	0.49	0.03	0.32
White (RAIS data)	0.75	0.65	0.60	0.20	0.47
Mother has a high school degree		0.78	0.45	0.52	0.43
HH income > 1.5× min. wage		0.74	0.28	0.30	0.25

Included programs (19 in total):

<u>Business</u>: Economics.

<u>Humanities</u>: Art, Biological Sciences (SGO), English/German/Japanese, Geography Ed. (SGO), Language (SGO), Math Ed. (SGO), Teaching, Teaching (DDC), Physical Ed., Spanish/French/Italian.

<u>Natural sciences</u>: Cartographic Engineering, Math, Mechanical Engineering (NF), Oceanography, Physics, Production engineering (RES), Statistics.

<u>Social sciences</u>: Philosophy.

Notes: This table reports summary statistics for UERJ applicants to programs in our sample. Panel A includes programs that are in our RD and DD samples (24 programs). Panel B includes programs in our DD sample only (19 programs). Programs are at UERJ's main campus in Rio unless denoted with parentheses. Column (A) includes applicants in the pre-AA cohorts. Columns (B)–(E) include applicants to the four admission tracks in the post-AA cohorts. See Appendices B.1 and B.4 for details on variable definitions and the sample.

Table 2. RD estimates of the effects of UERJ enrollment on graduation and earnings

Dependent variable	(A) 1995–2001 general track Mean below	(B) RD coef	(C) 2004–2011 general track Mean below	(D) RD coef	(E) 2004–2011 AA tracks Mean below	(F) RD coef
Panel A. First stage						
Enrolled in UERJ program	0.003	0.313*** (0.010)	0.008	0.292*** (0.006)	0.004	0.689*** (0.014)
N	3,234	17,519	4,012	47,838	543	6,121
Panel B. Graduation and earnings 6–9 years after application (2SLS)						
Graduated from UERJ program	0.002	0.711*** (0.017)	0.003	0.677*** (0.013)	0.004	0.640*** (0.018)
Formal employment	0.627	0.064** (0.029)	0.672	−0.031 (0.027)	0.729	−0.002 (0.026)
Log hourly wage	3.237	−0.003 (0.050)	3.387	−0.079 (0.049)	2.813	0.132*** (0.044)
Monthly earnings (2019 USD)	1,356.069	0.295 (75.313)	1,390.819	−153.473** (77.290)	816.821	110.230** (49.546)
N (formal employment)	3,234	37,794	4,012	55,030	543	8,147
N (log hourly wage)	2,027	24,564	2,694	32,972	394	6,100
Panel C. Graduation and earnings 10–13 years after application (2SLS)						
Graduated from UERJ program	0.002	0.718*** (0.017)	0.003	0.693*** (0.014)	0.003	0.661*** (0.021)
Formal employment	0.693	0.032 (0.027)	0.686	−0.026 (0.031)	0.714	0.037 (0.039)
Log hourly wage	3.636	0.005 (0.054)	3.637	0.005 (0.058)	3.052	0.024 (0.063)
Monthly earnings (2019 USD)	2,005.191	−84.946 (94.587)	1,757.947	−99.418 (109.084)	1,041.942	56.577 (75.202)
N (formal employment)	3,234	39,134	2,974	41,138	388	4,320
N (log hourly wage)	2,237	24,273	2,021	26,407	273	3,746

Notes: This table presents RD estimates for the effects of UERJ enrollment on graduation, formal employment, and earnings. Columns (A), (C), and (E) show means of each dependent variable for applicants in each group who scored $(−0.1, 0)$ SDs below the cutoff. In Panel A, columns (B), (D), and (F) show reduced-form RD coefficients, θ, from equation (1), which measure the effects of UERJ admission on UERJ enrollment. In Panels B and C, these columns show 2SLS RD coefficients, β, from equation (2), which measure the effects of UERJ enrollment on graduation and labor market outcomes. Panel B measures outcomes 6–9 years after application, and Panel C measures outcomes 10–13 years after application. Parentheses contain standard errors clustered at the individual level. We use the Calonico et al. (2014) bandwidth for each outcome; sample sizes refer to the dependent variables indicated in the parentheses after N.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 3. RD estimates for enrollment in other universities and degree attainment

	(A)	(B)	(C)	(D)	(E)	(F)
	1995–2001 general track		2004–2011 general track		2004–2011 AA tracks	
Dependent variable	Mean below	RD coef	Mean below	RD coef	Mean below	RD coef

Panel A. Enrollment in Rio de Janeiro universities (reduced form, 2009–2011 cohorts only)

# enrolled in UERJ			1.465	0.271*** (0.037)	1.051	0.880*** (0.088)
# enrolled in UFRJ			3.369	−0.147*** (0.057)	2.381	−0.111 (0.137)
# enrolled in other federal universities			4.407	−0.165** (0.083)	3.181	0.041 (0.168)
# enrolled in a top-100 private university			5.154	−0.176** (0.077)	4.312	0.147 (0.164)
# enrolled in other private universities			5.110	−0.041 (0.062)	5.181	−0.457** (0.229)
# enrolled in same progam area (2-dig)			3.448	0.120** (0.059)	2.647	0.351** (0.138)
# enrolled in same progam area (3-dig)			1.661	0.192*** (0.040)	1.367	0.459*** (0.082)
N (# enrolled in UERJ)			1,553	19,895	215	2,757

Panel B. Educational attainment measured in RAIS (2SLS)

Any college degree, 6-9 years later	0.731	0.044 (0.032)	0.785	0.006 (0.029)	0.636	−0.002 (0.038)
Ever earned a college degree	0.911	0.012 (0.017)	0.839	0.026 (0.025)	0.713	0.010 (0.033)
Ever earned a graduate degree	0.107	−0.004 (0.020)	0.069	−0.017 (0.017)	0.051	−0.006 (0.015)
N (ever college degree)	2,417	32,718	2,925	36,617	415	5,978

Notes: This table presents RD estimates for enrollment in Rio de Janeiro universities and educational attainment.

In Panel A, columns (B), (D), and (F) show reduced-form RD coefficients, θ, from equation (1). The dependent variables are the *total* number of enrollees in a given group of universities or field of study who share the applicant's birthdate, gender, and enrollment year. We measure these totals in Brazil's higher education census (see Appendix B.5). We categorize universities into four groups based on ownership and selectivity: 1) The federal university in the municipality of Rio de Janeiro (UFRJ); 2) The three other federal universities in the suburbs of Rio de Janeiro (UFF, UFRRJ, UNIRIO); 3) Private universities in the municipality of Rio de Janeiro that ranked in the top 100 of the 2012 *Folha* ranking (PUC-Rio, UNESA); and 4) Other private universities in the municipality of Rio de Janeiro (UGF, UVA, UCAM, Universo, UCB). We categorize fields of study using 2- and 3-digit census major codes. Regressions include only 2009–2011 UERJ applicants, and they include gender and age dummies to increase precision.

In Panel B, columns (B), (D), and (F) show 2SLS RD coefficients, β, from equation (2). Regressions include all UERJ cohorts, and the dependent variables are indicators for educational attainment measured in the RAIS data.

In both panels, columns (A), (C), and (E) show means of each dependent variable for applicants in each group who scored $(-0.1, 0)$ SDs below the cutoff. Parentheses contain standard errors clustered at the individual level. We use the Calonico et al. (2014) bandwidth for each outcome; sample sizes refer to the dependent variables indicated in the parentheses after N.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 4. RD estimates for employment at alumni firms

	(A)	(B)	(C)	(D)	(E)	(F)
	1995–2001 general track		2004–2011 general track		2004–2011 AA tracks	
Dependent variable	Mean below	RD coef	Mean below	RD coef	Mean below	RD coef

Panel A. Employment 6–9 years after application (2SLS)

Employed at firm with any UERJ alum	0.600	0.118*** (0.033)	0.572	0.070** (0.034)	0.477	0.137*** (0.038)
Employed at firm with any general track alum	0.579	0.129*** (0.035)	0.540	0.076** (0.034)	0.437	0.106*** (0.035)
Employed at firm with any AA track alum	0.375	0.036 (0.031)	0.402	0.076** (0.033)	0.386	0.129*** (0.037)
# UERJ alumni per 1000 workers at firm	6.561	4.652*** (1.624)	7.120	9.738*** (2.583)	3.334	8.523** (3.580)
Firm mean wage (log)	3.303	0.018 (0.043)	3.475	−0.095* (0.053)	3.073	0.106* (0.062)
N (firm mean wage)	2,024	30,345	2,681	31,087	394	4,306

Panel B. Employment 10–13 years after application (2SLS)

Employed at firm with any UERJ alum	0.649	0.059* (0.033)	0.573	0.028 (0.037)	0.498	0.080* (0.044)
Employed at firm with any general track alum	0.629	0.060 (0.037)	0.550	0.014 (0.038)	0.451	0.086** (0.039)
Employed at firm with any AA track alum	0.411	0.032 (0.034)	0.384	0.056 (0.036)	0.383	0.082* (0.044)
# UERJ alumni per 1000 workers at firm	5.873	3.648*** (1.078)	5.620	5.954** (3.025)	3.224	−0.092 (2.767)
Firm mean wage (log)	3.572	0.093* (0.053)	3.581	−0.053 (0.062)	3.223	0.049 (0.071)
N (firm mean wage)	2,236	24,701	2,010	21,071	275	3,133

Notes: This table presents RD estimates for the effects of UERJ enrollment on employment at alumni firms. Columns (A), (C), and (E) show means of each dependent variable for applicants in each group who scored $(−0.1, 0)$ SDs below the cutoff. Columns (B), (D), and (F) show 2SLS RD coefficients, β, from equation (2), which measure the effects of UERJ enrollment on employment outcomes. Panel A measures outcomes 6–9 years after application, and Panel B measures outcomes 10–13 years after application. In the first three rows of each panel, the dependent variables are indicators for employment at any firm during this time period with: 1) any UERJ alum; 2) any general track alum; and 3) any AA track alum. In the fourth row, the dependent variable is the number of alumni per 1000 workers at the applicant's firms (averaged over the time period). We define the firm's number of workers using its mean size over all years of our data. In the last row of each panel, the dependent variable is the natural log of the mean hourly wage at the applicant's firms (averaged over the time period). Parentheses contain standard errors clustered at the individual level. We use the Calonico et al. (2014) bandwidth for each outcome; sample sizes refer to the dependent variables indicated in the parentheses after N.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 5. DD estimates of the effects of AA exposure on student body composition

	(A)	(B)	(C)	(D)
	Pre-AA mean	DD coefficients		
Dependent variable	All top enrollees	All top enrollees	URM top enrollees	Non-top enrollees
Panel A. Exposure to affirmative action				
Prop. of classmates from AA tracks	0.000	0.189*** (0.017)	0.208*** (0.021)	0.192*** (0.018)
Panel B. Demographic characteristics				
Age at application	21.921	0.191 (0.312)	0.794 (0.899)	0.666*** (0.229)
Female	0.501	0.032 (0.022)	0.102 (0.078)	0.038* (0.021)
White	0.810	0.013 (0.018)		−0.121*** (0.025)
Brown	0.156	0.000 (0.012)		0.043** (0.017)
Black	0.025	−0.005 (0.010)	−0.009 (0.037)	0.077*** (0.012)
Panel C. Admission exam scores (standardized in population of all enrollees)				
Field exam writing score	0.178	−0.045 (0.043)	0.031 (0.129)	−0.246*** (0.046)
Mean field exam subject score	0.151	−0.029 (0.064)	0.039 (0.121)	−0.182** (0.084)
Admission score	0.270	−0.080 (0.112)	0.034 (0.160)	−0.498*** (0.143)
Panel D. Predicted log wage based on characteristics and scores				
Predicted log wage	3.298	−0.023 (0.029)	−0.011 (0.049)	−0.161*** (0.043)
Predicted log wage (if in RAIS)	3.251	−0.033 (0.028)	−0.009 (0.053)	−0.154*** (0.043)
N (enrollees)	16,466	35,866	1,631	30,854

Notes: This table displays DD estimates of the effect of affirmative action exposure on student characteristics. Column (A) shows the mean of each dependent variable for all top enrollees in the 1995–2001 cohorts. Columns (B)–(D) display estimates of π from equation (3) for all top enrollees, URM top enrollees, and non-top enrollees. The dependent variables are:

- Panel A. The proportion of enrollees in an individual's program/cohort who were from an affirmative action track.
- Panel B. Demographic characteristics of enrollees. For the early cohorts of our data, we use gender/race measured in the RAIS data (see Appendix Table A1).
- Panel C. Applicants' field exam scores and their overall admission score. We normalize scores to be mean 0/SD 1 in the population of all UERJ enrollees in UERJ in a given cohort. The field score regressions include cohort dummies interacted with dummies for the set of subject tests that the applicant took (which vary by major).
- Panel D. The predicted value from a regression of log hourly wage (6–9 years after application) on each of the variables in Panels B–C.

Parentheses contain standard errors clustered at the program level.
   * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 6. DD estimates for graduation, employment, and earnings 6–9 years after application

| | (A) | (B) | (C) | (D) |
| | Pre-AA mean | DD coefficients | | |
Dependent variable	All top enrollees	All top enrollees	URM top enrollees	Non-top enrollees
Panel A. Graduation and formal employment				
Graduated from UERJ program	0.556	0.013	0.013	0.006
		(0.021)	(0.063)	(0.028)
Formal employment	0.734	−0.027*	0.076	−0.012
		(0.015)	(0.054)	(0.015)
Panel B. Earnings				
Log hourly wage	3.245	−0.132***	−0.220**	−0.212***
		(0.045)	(0.107)	(0.062)
Monthly earnings (2019 USD)	1,380.558	−169.838***	−28.036	−272.989***
		(53.057)	(100.006)	(89.500)
Firm mean hourly wage (log)	3.316	−0.095**	−0.334***	−0.183***
		(0.035)	(0.113)	(0.051)
Panel C. Employment at firms with pre- and post-AA alumni				
Pre-AA alumni	0.602	−0.055**	−0.009	−0.044
		(0.023)	(0.081)	(0.033)
Only post-AA alumni	0.067	0.049**	0.019	0.036
		(0.023)	(0.056)	(0.023)
Panel D. Alumni firm employment by application track and cohort				
General track alumni from same cohort	0.451	−0.098***	−0.118*	−0.072**
		(0.021)	(0.068)	(0.028)
General track alumni from diff. cohort	0.233	0.042**	0.138	0.004
		(0.016)	(0.086)	(0.017)
Only AA alumni from same cohort	0.000	0.036***	0.035**	0.051***
		(0.009)	(0.014)	(0.007)
Only AA alumni from diff. cohort	0.012	0.010**	−0.017	0.014***
		(0.005)	(0.022)	(0.004)
N (enrollees)	16,466	35,866	1,631	30,854
N (wage observations)	12,062	26,445	1,323	22,975

Notes: This table displays DD estimates of the effect of affirmative action exposure on graduation, earnings, and employment at alumni firms measured 6–9 years after application. Column (A) shows the mean of each dependent variable for all top enrollees in the 1995–2001 cohorts. Columns (B)–(D) display estimates of π from equation (3) for all top enrollees, URM top enrollees, and non-top enrollees. The dependent variables are defined similarly to those in Tables 2 and 4. In Panel C, we define alumni firms using graduates from the pre- and post-AA cohorts. In Panel D, we define alumni firms using the alum's cohort (same or different than the applicant's cohort) and application track (general or AA). We define the outcomes in Panels C–D to be non-overlapping, i.e., the variables in the lower rows of each panel equal one only if the firm did not hire alumni who meet the criteria for the higher rows. Parentheses contain standard errors clustered at the program level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE 7. DD estimates for UERJ's Enade exam taker characteristics and scores

	(A)	(B)	(C)	(D)
	UERJ pre-AA mean	DD coefficients		
Dependent variable	All students	All students	White private HS students	Other students

Panel A. Characteristics of Enade exam takers

# exam takers	36.086	4.322	−7.926***	9.142***
		(2.664)	(1.398)	(1.777)
White	0.716	−0.132***		−0.066***
		(0.009)		(0.009)
Private HS student	0.570	−0.131***		−0.040***
		(0.015)		(0.009)
Female	0.526	−0.005	−0.009	−0.016
		(0.011)	(0.012)	(0.011)
Age	26.520	0.661**	0.042	0.700*
		(0.261)	(0.209)	(0.361)
Mother has a HS degree	0.692	−0.086***	0.016	−0.081***
		(0.012)	(0.012)	(0.012)
HH income / minimum wage	7.724	−1.577***	0.292	−1.873***
		(0.098)	(0.210)	(0.117)
Predicted general score	0.563	−0.008***	0.001**	−0.007***
		(0.001)	(0.001)	(0.001)

Panel B. Enade scores (proportion correct answers)

Overall score	0.553	−0.038***	−0.022**	−0.051***
		(0.005)	(0.008)	(0.006)
Field-specific component	0.519	−0.037***	−0.021**	−0.048***
		(0.005)	(0.009)	(0.006)
General component	0.657	−0.041***	−0.026***	−0.059***
		(0.005)	(0.007)	(0.007)
N (programs × years)	36	1,664	1,664	1,664
N (exam takers)	1,059	61,112	16,851	37,992

Notes: This table displays DD estimates of the effect of affirmative action on the characteristics (Panel A) and scores (Panel B) of UERJ's Enade exam takers. The sample includes 2004–2015 Enade exam takers from UERJ and other federal and state universities that did not implement affirmative action until 2012 or later. See Appendix Table A17 for details on our Enade sample and the exam fields. Column (A) shows dependent variable means for UERJ exam takers in 2004–2006. The other columns show π coefficients from the DD regression:

$$Y_{mjt} = \gamma_{mj} + \gamma_{mt} + \pi[\text{UERJ}_j \times \text{Post}_t] + \varepsilon_{mjt}.$$

Regressions are at the exam field (m) by institution (j) by year (t) level, with observations weighted by the number of exam takers (except in the first row of Panel A, where we weight by the number of 2004–2006 exam takers in each mj cell.) We include dummies for field × institution, field × year, and the interaction between indicators for UERJ and the 2007–2015 cohorts ($\text{UERJ}_j \times \text{Post}_t$). Columns (B)–(D) include all students, white private high school students, and non-white and/or public school students. "Predicted general score" is the predicted value from a regression of general component scores on age and dummies for gender, race, private high school, mother's education, father's education, and family income. Parentheses contain standard errors clustered at the institution level.

    * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix — For Online Publication

A. Appendix figures and tables

Panel A. All applicants   Panel B. Applicants who appear in the RAIS
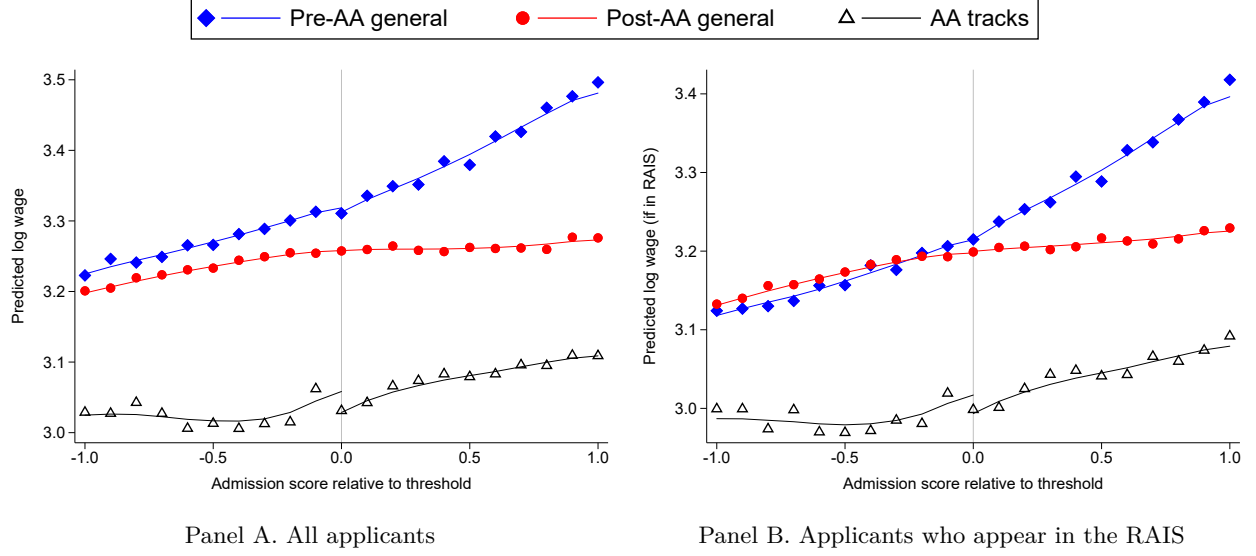
FIGURE A1. Predicted log wage based on applicant characteristics

Notes: This figure presents RD graphs for pre-AA general applicants (blue diamonds), post-AA general applicants (red circles), and Black/public school applicants (black triangles). The x-axis in each panel is an applicant's standardized admission score normalized to zero at the cutoff. The dependent variable on the y-axis is the predicted value from a regression of log hourly wage (6–9 years after application) on student characteristics (age, gender, race, mother's education, family income, score on the writing component of the field exam, and qualifying exam score). Markers depict means in 0.1 SD bins of the standardized score. Lines are predicted values from local linear regressions estimated separately above and below the threshold with a triangular kernel.
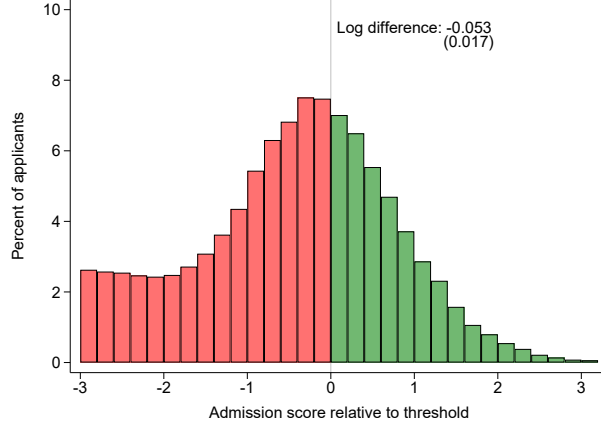
Panel A. General track (1995–2001)
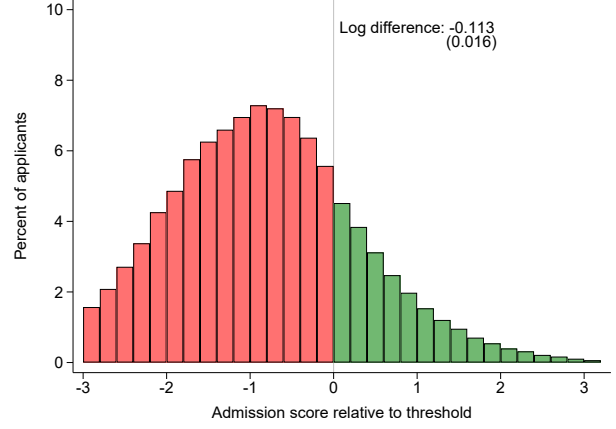
Panel B. General track (2004–2011)
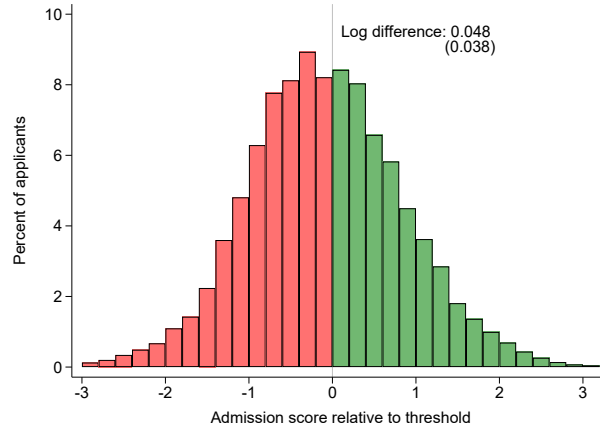
Panel C. AA tracks (2004–2011)

FIGURE A2. Density of admission scores relative to the threshold

Notes: This figure shows the density of admission scores relative to the cutoff. The x-axis is a student's admission score normalized to zero at the cutoff of the relevant application pool. The y-axis shows the percent of applicants within 0.20 SD unit bins of the admission score. We restrict the figure to only display normalized scores within three SD of the cutoff. We also exclude applicants whose score defines the cutoff.

Panel A shows the distribution of admission scores for pre-AA general applicants, Panel B for post-AA general applicants, and Panel C for Black/public school applicants.

Each figure displays the estimated the log difference in height at the threshold using the McCrary (2008) density test. The standard error is shown in parentheses.
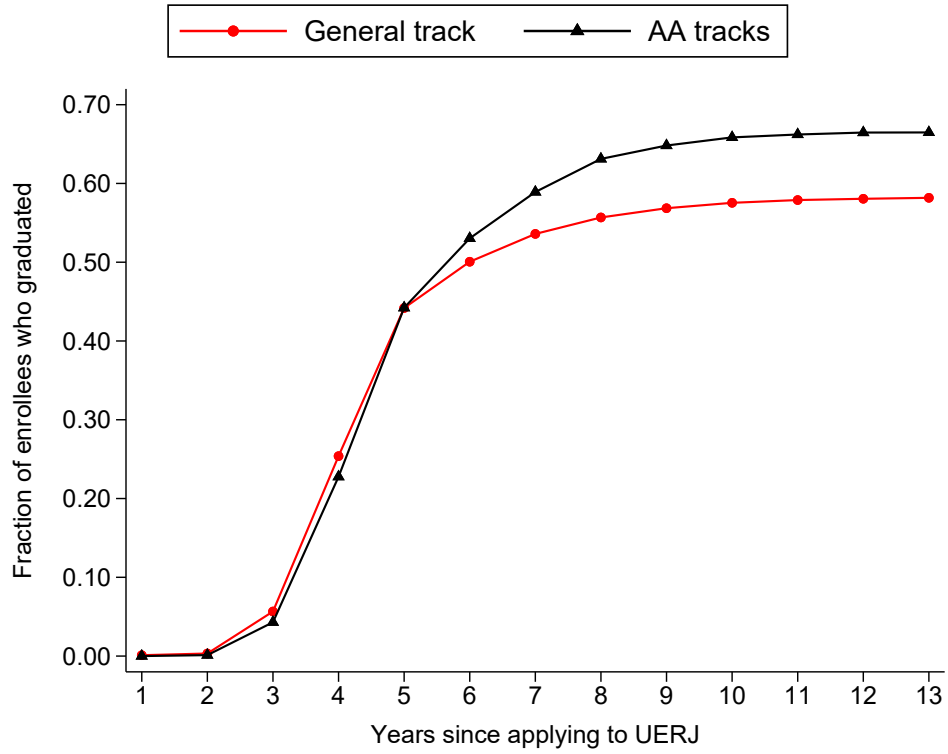
FIGURE A3. UERJ graduation rate by year since application

Notes: This figure show the empirical cumulative distribution function of the graduation rate of students in programs in our RD sample (Panel A of Table 1). We plot separately the graduation rate of general track enrollees (red line) and Black/public school enrollees (black line).

Panel A. By years since application
(1997–2006 cohorts)

Panel B. By cohort
(Wages measured 6–9 years later)
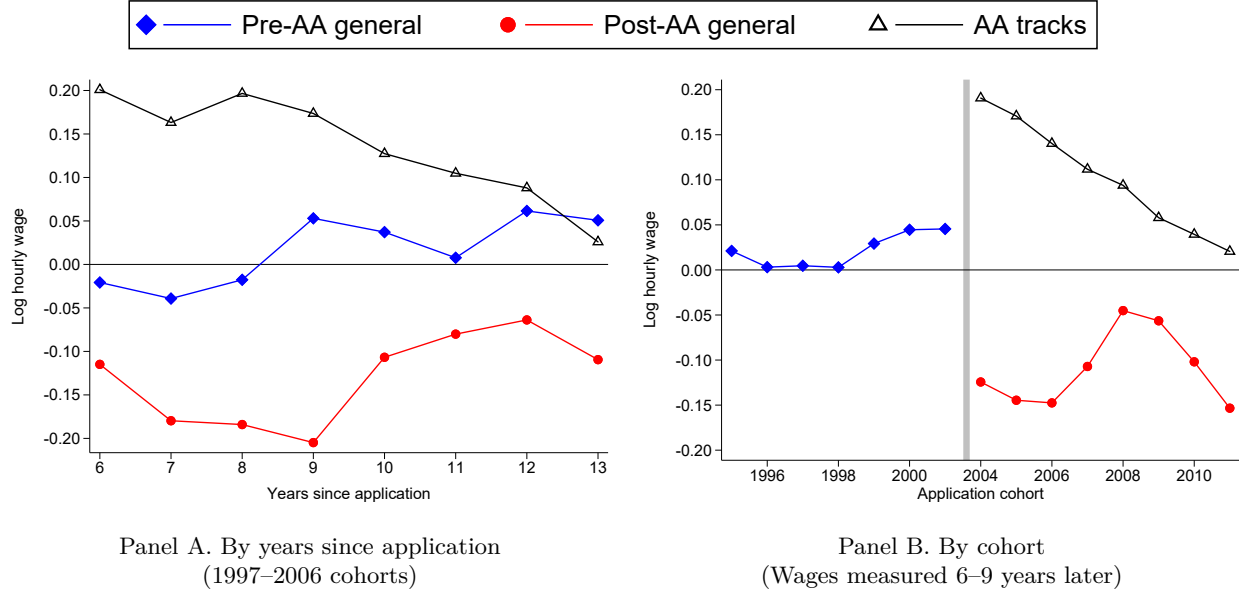
FIGURE A4. RD estimates for log hourly wages by years since application and cohort

Notes: This figure displays 2SLS RD coefficients, β, for pre-AA general applicants (blue diamonds), post-AA general applicants (red circles), and Black/public school applicants (black triangles).

Panel A plots β coefficients for log hourly wages measured at different years since individuals applied to UERJ. To smooth estimates, we use the applicant's three-year average wage as the dependent variable in each regression (years $t-1$, t, and $t+1$). We include only 1997–2006 cohorts since we observe their outcomes in each of 6–13 years later.

Panel B plots β coefficients for log hourly wages estimated in different application cohorts. To smooth estimates, we include three adjacent cohorts for each regression (cohorts $t-1$, t, and $t+1$). All regressions use mean log hourly wage measured 6–9 years after application as the dependent variable.
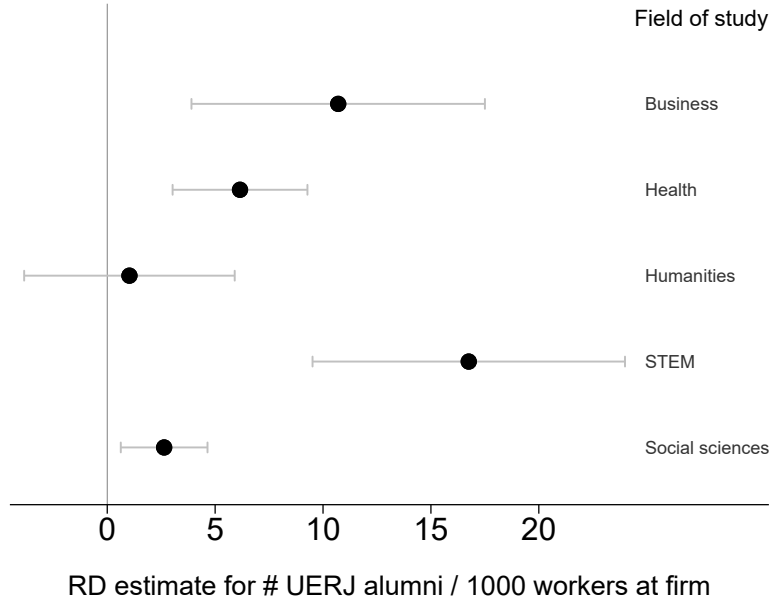
FIGURE A5. Field of study heterogeneity in RD estimates for alumni per 1000 workers at the firm

Notes: This figure examines heterogeneity by field of study in the effects of UERJ enrollment on employment at alumni firms. The figure displays RD estimates for the effects of UERJ enrollment on the mean number of alumni per 1000 workers at the applicants' firms measured 6–9 years after application. These estimates are analogous to the 2SLS RD coefficients in the fourth row of Table 4 (Panel A), except we estimate equation (2) pooling across all applicant groups. We estimate these regressions separately for the five field of study groups listed in Panel A of Table 1. Markers depict the RD coefficients, β, coefficients from these regressions, and horizontal bars are 95 percent confidence intervals using standard errors clustered at the individual level.
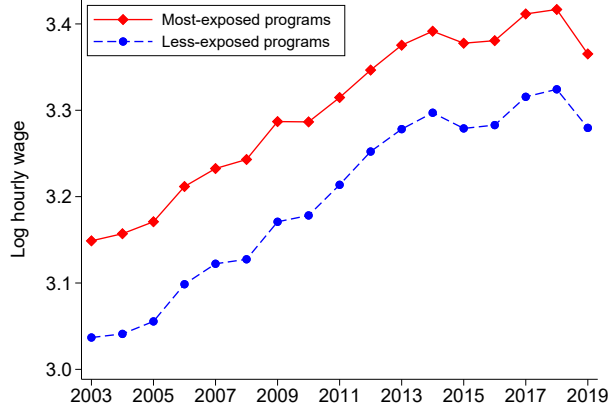
Panel A. Industry mean wage in
majors w/ more and less AA exposure
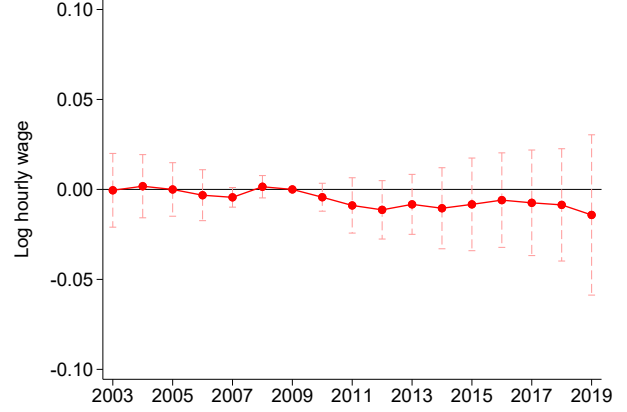
Panel B. Event study of industry mean wage

FIGURE A6. Wage trends in industries that employed pre-AA top enrollees

Notes: This figure shows how hourly wages trended over time in industries that employed top enrollees from the pre-AA cohorts. We begin by computing the mean hourly in each (5-digit) industry j and year t using all workers in the RAIS data; we denote this industry $\times$ year mean wage by $\bar{w}_{jt}$. We then take our sample of top UERJ enrollees in the pre-AA cohorts (1995–2001), and compute the share of individuals in each major m who were employed in industry j measured 6–9 years after UERJ application; we denote these shares s_{mj}, with $\sum_j s_{mj} = 1$ for each m. Lastly, we define $\bar{w}_{mt} = \sum_j s_{mj} * \bar{w}_{jt}$ as the industry mean wage for UERJ major m in year t, where this mean is computed using the pre-AA employment shares in each industry.

In Panel A, we plot the average value of $\bar{w}_{mt}$ in majors with more- (red line) and less- (blue line) exposure to affirmative action in each year from $t = 2003$ to 2019. More-exposed programs are those in which the share of affirmative action enrollees in the 2004–2011 cohorts was 30 percent or higher (Panel A of Table 1). Less-exposed programs are those in which the share of affirmative action enrollees in the 2004–2011 cohorts was below 30 percent (Panel B of Table 1).

Panel B plots π_t coefficients from the following event study regression:

$$\bar{w}_{mt} = \gamma_m + \gamma_{tf(m)} + \pi_t \text{ExposureToAA}_m + \varepsilon_{mt},$$

where γ_m are program fixed effects, $\gamma_{tf(m)}$ are year $\times$ field of study fixed effects, and π_t are coefficients on the interaction between year dummies (omitting 2009) and an indicator for more-exposed programs, ExposureToAA_m. Observations in this event study are weighted by the number of pre-AA top enrollees in each major m. Vertical dashed lines are 95 percent confidence intervals with standard errors clustered at the program level.

Table A1. RD balance tests

Dependent variable (cohorts observed)	(A) 1995–2001 general track Mean below	(B) RD coef	(C) 2004–2011 general track Mean below	(D) RD coef	(E) 2004–2011 AA tracks Mean below	(F) RD coef
Panel A. Applicant characteristics						
Female (2004–2011 cohorts)			0.530	0.001 (0.008)	0.602	0.001 (0.021)
Female (measured in RAIS) (1995–2011 cohorts)	0.486	0.008 (0.010)	0.518	0.012 (0.009)	0.600	0.001 (0.023)
White (2007–2011 cohorts)			0.696	0.002 (0.010)	0.343	−0.022 (0.028)
White (measured in RAIS) (1995–2011 cohorts)	0.790	0.011 (0.009)	0.714	0.003 (0.008)	0.436	−0.027 (0.020)
Brown (2007–2011 cohorts)			0.212	−0.009 (0.009)	0.313	0.023 (0.030)
Brown (measured in RAIS) (1995–2011 cohorts)	0.173	−0.008 (0.010)	0.216	−0.010 (0.007)	0.321	0.028 (0.022)
Age at application (1995–2011 cohorts)	20.608	0.181** (0.081)	20.043	0.037 (0.072)	22.306	−0.446* (0.262)
Mother has HS degree (2007–2011 cohorts)			0.901	−0.002 (0.007)	0.534	0.001 (0.033)
HH income > 1.5x min. wage (2007–2011 cohorts)			0.886	−0.007 (0.007)	0.341	0.008 (0.026)
Writing score (SD units) (1995–2001, 2007–2011 cohorts)	0.174	0.020 (0.015)	0.477	0.011 (0.016)	−0.202	0.011 (0.052)
Qualifying exam score (SD units) (1995–2001 cohorts)	−0.148	−0.009 (0.008)				
Joint balance test (p value)		0.110		0.411		0.875
Panel B. Predicted log wage based on applicant characteristics						
Predicted log wage (1995–2011 cohorts)	3.313	−0.004 (0.006)	3.254	0.001 (0.003)	3.062	−0.003 (0.010)
Predicted log wage (if in RAIS) (1995–2011 cohorts)	3.206	0.004 (0.006)	3.193	0.003 (0.004)	3.019	−0.006 (0.011)
N (predicted wage)	3,234	27,610	4,012	45,731	543	6,410
N (predicted wage if in RAIS)	2,027	17,027	2,694	30,315	394	4,303

Notes: This table presents RD balance tests. Columns (A), (C), and (E) show means of each dependent variable for applicants in each group who scored $(-0.1, 0)$ SDs below the cutoff. Columns (B), (D), and (F) display reduced-form RD coefficients, θ, from equation (1), using the dependent variable listed in the row header.

The last row in Panel A reports the p values from F tests that the coefficients on all covariates are jointly equal to zero. Parentheses contain standard errors clustered at the individual level. We use the Calonico et al. (2014) bandwidth for each outcome; sample sizes refer to the dependent variables indicated in the parentheses after N.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A2. RD complier characteristics

	(A)	(B)	(C)	(D)	(E)	(F)
	Complier characteristics			Differences in complier characteristics		
Dependent variable (cohorts observed)	Pre-AA general track	Post-AA general track	AA tracks	(B)−(A)	(C)−(A)	(C)−(B)

Panel A. Applicant characteristics

Female (measured in RAIS)	0.507***	0.523***	0.616***	0.016	0.109***	0.093***
(1995–2011 cohorts)	(0.014)	(0.014)	(0.019)	(0.020)	(0.023)	(0.024)
White (measured in RAIS)	0.795***	0.727***	0.382***	−0.068***	−0.412***	−0.344***
(1995–2011 cohorts)	(0.012)	(0.012)	(0.017)	(0.017)	(0.021)	(0.021)
Brown (measured in RAIS)	0.180***	0.204***	0.338***	0.023	0.158***	0.134***
(1995–2011 cohorts)	(0.013)	(0.010)	(0.017)	(0.017)	(0.022)	(0.020)
Age at application	21.485***	20.595***	21.879***	−0.891***	0.394*	1.285***
(1995–2011 cohorts)	(0.129)	(0.118)	(0.191)	(0.175)	(0.230)	(0.224)
Writing score (SD units)	0.359***	0.678***	−0.257***	0.319***	−0.616***	−0.935***
(1995–2001, 2007–2011 cohorts)	(0.024)	(0.028)	(0.043)	(0.037)	(0.049)	(0.051)

Panel B. Predicted log wage based on applicant characteristics

Predicted log wage	3.327***	3.291***	3.027***	−0.036***	−0.300***	−0.264***
(1995–2011 cohorts)	(0.012)	(0.005)	(0.010)	(0.013)	(0.015)	(0.011)
Predicted log wage (if in RAIS)	3.231***	3.228***	2.985***	−0.003	−0.245***	−0.242***
(1995–2011 cohorts)	(0.013)	(0.006)	(0.010)	(0.014)	(0.017)	(0.012)

N (predicted wage)	27,707	45,674	6,111	73,381	33,818	51,785
N (predicted wage if in RAIS)	18,093	30,796	4,277	48,889	22,370	35,073

Notes: This table displays complier characteristics for our 2SLS RD regressions. Columns (A)–(C) display complier characteristics for pre-AA applicants, post-AA general track applicants, and AA applicants, respectively. These complier characteristics are the β coefficients from our 2SLS RD specification (2) in which the dependent variables are the interaction between an indicator for enrolling in UERJ and the applicant characteristic listed in the row header. In Panel B, the dependent variables are the interaction between an indicator for enrolling in UERJ and the applicant's predicted log wage based on their demographic characteristics (as in Panel B of Table A1). Columns (D)–(F) report the differences between the complier characteristics for: (D) post-AA general and pre-AA applicants; (E) AA and pre-AA applicants; and (F) AA and post-AA general applicants.

Parentheses contain standard errors clustered at the individual level. We use the Calonico et al. (2014) bandwidth for each outcome; sample sizes refer to the dependent variables indicated in the parentheses after N.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A3. Effects of UERJ enrollment on job outcomes measured 10–13 years later

	(A)	(B)	(C)	(D)	(E)	(F)
	RD coefficient measured 10–13 years later			*Change* in RD coefficients from 6–9 to 10–13 years later		
Dependent variable	1995–01 general	2004–11 general	AA tracks	1995–01 general	2004–11 general	AA tracks
Panel A. Formal employment and earnings (2SLS)						
Formal employment	0.032	−0.026	0.037	−0.032	0.005	0.038
	(0.027)	(0.031)	(0.039)	(0.021)	(0.028)	(0.036)
Log hourly wage	0.005	0.005	0.024	0.008	0.084	−0.108**
	(0.054)	(0.058)	(0.063)	(0.046)	(0.052)	(0.055)
Monthly earnings (2019 USD)	−84.946	−99.418	56.577	−85.241	54.056	−53.652
	(94.587)	(109.084)	(75.202)	(71.026)	(87.121)	(60.324)
N (log hourly wage)	24,273	26,407	3,746	48,837	59,379	9,846
Panel B. Employment at firms that hired UERJ alumni (2SLS)						
Employed at firm with any UERJ alum	0.059*	0.028	0.080*	−0.059*	−0.042	−0.057
	(0.033)	(0.037)	(0.044)	(0.032)	(0.036)	(0.042)
# UERJ alumni / 1000 workers at firm	3.648***	5.954**	−0.092	−1.004	−3.784	−8.615**
	(1.078)	(3.025)	(2.767)	(1.497)	(2.742)	(3.514)
N (employed at firm with any alum)	25,535	28,217	4,179	50,859	62,344	9,239

Notes: This table presents RD estimates for employment and earnings measured 10–13 years after application. Columns (A)–(C) show 2SLS RD coefficients, β, from equation (2) for each applicant group. Columns (D)–(F) show the difference in the 2SLS RD coefficients between the periods of 6–9 and 10–13 years after application. Parentheses contain standard errors clustered at the individual level. We use the Calonico et al. (2014) bandwidth for each outcome; sample sizes refer to the dependent variables indicated in the parentheses after N.

 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A4. Robustness of RD estimates — General track (1995–2001)

	(A)	(B)	(C)	(D)	(E)
	RD coefficients by bandwidth, h^Y				
	1.0× CCT	0.5× CCT	1.5× CCT	Donut RD	Adding controls

Panel A. First stage

Enrolled in UERJ program	0.313***	0.354***	0.311***	0.253***	0.310***
	(0.010)	(0.014)	(0.008)	(0.021)	(0.010)
N	17,519	8,880	25,848	9,101	17,519

Panel B. Returns to UERJ enrollment 6–9 years later (2SLS)

Graduated from UERJ program	0.711***	0.708***	0.719***	0.714***	0.713***
	(0.017)	(0.022)	(0.015)	(0.037)	(0.017)
Formal employment	0.064**	0.104**	0.072***	0.059	0.016
	(0.029)	(0.042)	(0.022)	(0.043)	(0.019)
Log hourly wage	−0.004	0.012	0.027	−0.004	−0.006
	(0.050)	(0.074)	(0.038)	(0.055)	(0.049)
Monthly earnings (2019 USD)	0.440	103.639	66.956	−30.020	−2.036
	(75.269)	(114.267)	(57.390)	(92.678)	(74.054)
N (employment regression)	37,794	20,162	51,674	29,030	37,794
N (wage regression)	24,567	13,140	33,612	24,481	24,567

Panel C. Returns to UERJ enrollment 10–13 years later (2SLS)

Graduated from UERJ program	0.718***	0.729***	0.729***	0.709***	0.720***
	(0.017)	(0.022)	(0.015)	(0.038)	(0.017)
Formal employment	0.032	0.033	0.042**	−0.003	−0.008
	(0.027)	(0.039)	(0.021)	(0.039)	(0.014)
Log hourly wage	0.005	−0.037	0.018	0.006	0.004
	(0.054)	(0.077)	(0.041)	(0.060)	(0.054)
Monthly earnings (2019 USD)	−84.105	−83.191	30.639	−115.173	−85.290
	(94.586)	(145.009)	(73.007)	(135.201)	(94.117)
N (employment regression)	39,133	21,003	53,108	31,666	39,133
N (wage regression)	24,273	12,851	33,695	24,847	24,273

Notes: This table display RD coefficients using different specifications of our estimating equation. The coefficients are estimated on the sample of general track applicants in the pre-AA cohorts (1995–2001).

Columns (A)–(C) display the estimated RD coefficients using different sample bandwidths. Column (A) reproduces our baseline specification, which uses the Calonico et al. (2014) (CCT) optimal bandwidth for each outcome. In Column (B), we use a bandwidth half the size of the optimal CCT bandwidth. In Column (C), we use a bandwidth twice as large as the CCT bandwidth. In Column (D), we exclude applicants with an admission score within 0.05 SD of the cutoff. In Column (E), we include controls for age, gender, race, mother's educational attainment, family income, score on the writing component of the field exam, and qualifying exam score.

Panel A displays the first-stage effect, which the estimated θ from equation (1). Panels B–C display 2SLS RD coefficients, β, from equation (2). Parentheses contain standard errors clustered at the individual level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(A)	(B)	(C)	(D)	(E)
	RD coefficients by bandwidth, h^Y			Donut RD	Adding controls
	1.0× CCT	0.5× CCT	1.5× CCT		
Panel A. First stage					
Enrolled in UERJ program	0.292***	0.288***	0.295***	0.296***	0.292***
	(0.006)	(0.008)	(0.005)	(0.008)	(0.006)
N	47,838	24,477	69,856	36,957	47,838
Panel B. Returns to UERJ enrollment 6–9 years later (2SLS)					
Graduated from UERJ program	0.677***	0.647***	0.684***	0.715***	0.677***
	(0.013)	(0.017)	(0.011)	(0.016)	(0.013)
Formal employment	−0.030	−0.029	−0.025	−0.061*	0.012
	(0.027)	(0.038)	(0.022)	(0.033)	(0.016)
Log hourly wage	−0.080	−0.123*	−0.038	−0.046	−0.081*
	(0.049)	(0.069)	(0.041)	(0.057)	(0.049)
Monthly earnings (2019 USD)	−163.811**	−223.150**	−75.055	−106.698	−157.841**
	(77.383)	(105.344)	(63.181)	(97.113)	(76.054)
N (employment regression)	55,110	28,308	80,093	45,622	55,110
N (wage regression)	32,966	16,930	47,911	31,844	32,966
Panel C. Returns to UERJ enrollment 10–13 years later (2SLS)					
Graduated from UERJ program	0.693***	0.658***	0.709***	0.736***	0.694***
	(0.014)	(0.019)	(0.012)	(0.019)	(0.014)
Formal employment	−0.026	0.008	−0.007	−0.035	−0.008
	(0.031)	(0.043)	(0.026)	(0.037)	(0.021)
Log hourly wage	0.011	−0.072	−0.022	0.049	0.010
	(0.057)	(0.081)	(0.048)	(0.068)	(0.057)
Monthly earnings (2019 USD)	−102.941	−231.406	−53.098	−33.774	−100.881
	(109.073)	(152.445)	(89.843)	(135.372)	(108.029)
N (employment regression)	41,128	21,285	59,285	34,320	41,128
N (wage regression)	26,540	13,748	38,220	23,707	26,540

Notes: This table display RD coefficients using different specifications of our estimating equation. The table is structured similarly to Table A4, but the coefficients are estimated on the sample of general track applicants in the post-AA cohorts (2004–2011). See notes to Table A4 for details.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

	(A)	(B)	(C)	(D)	(E)
	RD coefficients by bandwidth, h^Y				
	1.0× CCT	0.5× CCT	1.5× CCT	Donut RD	Adding controls
Panel A. First stage					
Enrolled in UERJ program	0.689***	0.708***	0.713***	0.626***	0.689***
	(0.014)	(0.019)	(0.012)	(0.030)	(0.014)
N	6,121	3,262	8,606	3,308	6,121
Panel B. Returns to UERJ enrollment 6–9 years later (2SLS)					
Graduated from UERJ program	0.640***	0.642***	0.643***	0.660***	0.638***
	(0.018)	(0.023)	(0.015)	(0.034)	(0.018)
Formal employment	−0.002	−0.008	−0.013	0.022	−0.010
	(0.026)	(0.037)	(0.021)	(0.038)	(0.017)
Log hourly wage	0.132***	0.123**	0.130***	0.161***	0.125***
	(0.044)	(0.062)	(0.036)	(0.058)	(0.043)
Monthly earnings (2019 USD)	110.230**	114.289*	112.040***	202.213***	108.147**
	(49.523)	(66.820)	(40.360)	(76.211)	(48.984)
N (employment regression)	8,147	4,459	11,011	6,276	8,147
N (wage regression)	6,100	3,311	8,203	5,405	6,100
Panel C. Returns to UERJ enrollment 10–13 years later (2SLS)					
Graduated from UERJ program	0.661***	0.654***	0.670***	0.653***	0.660***
	(0.021)	(0.028)	(0.018)	(0.041)	(0.021)
Formal employment	0.037	0.060	0.011	0.057	0.025
	(0.039)	(0.055)	(0.031)	(0.054)	(0.031)
Log hourly wage	0.025	0.101	0.052	0.019	0.014
	(0.063)	(0.087)	(0.050)	(0.080)	(0.062)
Monthly earnings (2019 USD)	56.577	69.183	66.656	120.434	45.488
	(75.149)	(102.962)	(59.072)	(104.238)	(74.087)
N (employment regression)	4,320	2,280	6,109	3,958	4,320
N (wage regression)	3,748	2,024	5,240	3,693	3,748

Notes: This table display RD coefficients using different specifications of our estimating equation. The table is structured similarly to Table A4, but the coefficients are estimated on the sample of Black/public school applicants in the post-AA cohorts (2004–2011). See notes to Table A4 for details.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A7. Heterogeneity in RD estimates by field of study — Affirmative action tracks

	(A)	(B)	(C)	(D)	(E)
	Field of study				
Dependent variable	Business	Health	Human -ities	Natural sciences	Social sciences

Panel A. Returns to UERJ enrollment 6–9 years after applying (2SLS)

Graduated from UERJ program	0.697***	0.801***	0.528***	0.411***	0.722***
	(0.059)	(0.029)	(0.045)	(0.038)	(0.033)
Formal employment	0.071	−0.055	0.075	0.037	−0.053
	(0.077)	(0.053)	(0.058)	(0.055)	(0.051)
Log hourly wage	0.224*	0.204**	0.079	0.090	0.095
	(0.131)	(0.087)	(0.103)	(0.094)	(0.091)
Monthly earnings (2019 USD)	270.006**	217.408*	97.492	50.829	3.713
	(128.029)	(116.354)	(87.611)	(113.661)	(96.297)
N (formal employment)	784	1,895	1,295	1,681	2,492
N (log hourly wage)	698	1,382	1,009	1,269	1,719

Panel B. Returns to UERJ enrollment 10–13 years after applying (2SLS)

Graduated from UERJ program	0.724***	0.803***	0.556***	0.482***	0.715***
	(0.072)	(0.036)	(0.053)	(0.048)	(0.041)
Formal employment	0.080	0.018	−0.024	0.105	0.026
	(0.146)	(0.075)	(0.093)	(0.083)	(0.077)
Log hourly wage	0.240	0.172	0.025	−0.055	−0.141
	(0.220)	(0.127)	(0.145)	(0.133)	(0.122)
Monthly earnings (2019 USD)	327.743	243.571	17.347	−153.081	−45.814
	(239.378)	(159.063)	(157.375)	(163.135)	(151.932)
N (formal employment)	423	1,022	709	830	1,336
N (log hourly wage)	403	903	624	727	1,089

Notes: This table displays RD coefficients estimated on the sample of Black/public school applicants. Each column shows the result for applicants to different fields of study. Column (A) shows the results for applicants to business programs; column (B) for health programs; column (C) for humanities programs; column (D) for natural sciences programs, and column (E) for social sciences programs. See Table 1, Panel A, for the programs included in each field of study and Appendix Tables B2–B4 for the number of applicants by program/cohort.

Panels A–B display 2SLS RD coefficients, β, from equation (2). The dependent variables are program completion, formal employment, and earnings, each measured 6–9 years after applying (Panel A) and 10–13 years after applying (Panel B). See Appendix B.1 for variable definitions.

Parentheses contain standard errors clustered at the individual level. We use the Calonico et al. (2014) bandwidth for each outcome; sample sizes refer to the dependent variables indicated in the parentheses after N.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A8. Summary statistics for Rio de Janeiro universities in 2010

(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)
						in 2010 US dollars		
University name	Abbr.	Ownership	*Folha* national ranking	Undergrad enrollment	Graduate enrollment	Annual revenue (millions)	Annual expenses (millions)	Expenses per student
Univ. Federal do Rio de Janeiro	UFRJ	Federal	3	50,342	12,453	1,254.4	1,254.4	19,976
Univ. do Estado do Rio de Janeiro	**UERJ**	**State**	**11**	**30,144**	**5,767**	**463.9**	**465.5**	**12,962**
Pont. Univ. Católica do Rio de Janeiro	PUC-Rio	Private	13	17,061	3,352	291.3	265.8	13,022
Univ. Federal Fluminense	UFF	Federal	15	48,809	5,720	767.9	1,268.3	23,259
Univ. Federal Rural do Rio de Janeiro	UFRRJ	Federal	48	14,826	2,116	308.8	250.0	14,759
Univ. Federal do Estado do Rio de Janeiro	UNIRIO	Federal	67	14,418	934	140.5	109.2	7,112
Univ. Estácio de Sá	UNESA	Private	89	181,832	492	341.0	237.6	1,303
Univ. Gama Filho	UGF	Private	110	21,020	243	94.8	97.1	4,568
Univ. Veiga de Almeida	UVA	Private	147	21,983	184	62.4	55.7	2,512
Univ. Salgado de Oliveira	Universo	Private	155	48,130	173	246.5	114.8	2,377
Univ. Castelo Branco	UCB	Private	160	71,524	0	33.8	33.1	463
Univ. Cândido Mendes	UCAM	Private	168	21,454	458	69.8	83.9	3,827

55

Notes: This table displays summary statistics for universities in Rio de Janeiro. The sample includes private universities in the municipality of Rio, federal universities in the state of Rio, and UERJ. These are the universities we use to define enrollment outcomes in Table 3.

Columns (A)–(C) show the university's name, abbreviation, and ownership type. Column (D) reports the university's rank in the 2012 national ranking by the newspaper *Folha*. Column (E) shows the number of undergraduate students enrolled in each institution in 2010, which we compute from the individual-level dataset of Brazil's higher education census (*Censo da Educação Superior*). Column (F) shows the number of graduate students at each institution in 2010, which we compute from the CAPES census of graduate programs (*Discentes dos Programas de Pós-Graduação stricto sensu no Brasil*). Columns (G)–(H) report annual revenue and expenses in 2010 (converted to US dollars) from the school-level dataset of the *Censo da Educação Superior*. Column (I) shows annual expenses (column H) divided by total enrollment (columns E + F).

TABLE A9. Top employers of UERJ alumni

#	Firm	(A) No. UERJ graduates hired	(B) No. alumni hired per 1000 workers	(C) Firm size (mean)	(D) Located in Rio	(E) Public firm	(F) Prop. of employees w/ college	(G) Firm mean hourly wage (2019 USD)
Panel A. Top 10 firms by total number of UERJ alumni employees								
1	City Hall of Rio de Janeiro	1,161	13.30	87,274	Yes	Yes	0.461	6.891
2	State Secretary of Education	1,093	11.97	91,309	Yes	Yes	0.398	3.959
3	State University of Rio de Janeiro (UERJ)	409	56.29	7,266	Yes	Yes	0.690	13.062
4	Brazilian Petroleum (Petrobras - HQ)	384	62.91	6,104	Yes	No	0.780	27.690
5	State Secretary of Health	377	15.35	24,563	Yes	Yes	0.330	3.132
6	State Court of Law	321	21.27	15,093	Yes	Yes	0.718	16.466
7	Center for Payment of the Army	307	1.91	161,115	No	Yes	0.172	6.048
8	Federal University of Rio de Janeiro (UFRJ)	238	22.97	10,362	Yes	Yes	0.715	13.051
9	State Public Ministry	227	71.22	3,187	Yes	Yes	0.740	25.150
10	City Hall of Duque de Caxias	221	17.83	12,395	Yes	Yes	0.874	6.914
–	All other firms	–	–	510	0.777	0.084	0.384	6.971
Panel B. Top 10 firms by number of UERJ alumni hired per 1000 workers								
1	National Bank of Econ. & Social Dev.	217	109.23	1,987	Yes	No	0.875	38.010
2	Accenture	184	107.13	1,718	Yes	No	0.810	13.268
3	Petrobras - EDIHB	176	102.91	1,710	Yes	No	0.841	24.554
4	General Public Defender of the State	142	80.84	1,757	Yes	Yes	0.536	24.971
5	Petrobras - Research Center	137	72.20	1,898	Yes	No	0.693	22.740
6	State Public Ministry	227	71.22	3,187	Yes	Yes	0.740	25.150
7	Petrobras - Vibra Energy	86	68.34	1,258	Yes	No	0.781	20.840
8	TIM Cellular	112	67.67	1,655	Yes	No	0.813	13.542
9	Pedro II Federal Public School	139	63.35	2,194	Yes	Yes	0.828	10.560
10	Petrobras - EDISE	384	62.91	6,104	Yes	No	0.780	27.690
–	All other firms	–	–	537	0.777	0.085	0.384	6.964

Notes: This table displays summary statistics for top employers of UERJ alumni from the programs in our RD sample (Panel A of Table 1). Panel A lists the top ten firms ranked according to column (A), which is the number of UERJ graduates hired across all cohorts in our data. Panel B lists the top ten firms ranked according to column (B), which is the number of UERJ graduates (column A) divided by the firm size (column C) and multiplied by 1000. Column (C) shows the average firm size (number of employees). Column (D) indicates whether the firm is located in the state of Rio. Column (E) indicates whether the firm is public. Column (F) shows the proportion of the firm's employees with a college degree (from any school). Column (G) shows the firm mean hourly wage, measured in 2019 USD. The last row of each Panel shows the average of all other firms that hired at least one UERJ graduate in our sample.

T𝐀BLE A10. RD estimates for number of UERJ alumni in an applicant's firm and labor market

Dependent variable	(A)	(B)	(C)	(D)	(E)	(F)
	1995–2001 general track		2004–2011 general track		2004–2011 AA tracks	
	Mean below	RD coef	Mean below	RD coef	Mean below	RD coef
Panel A. Employment 6–9 years after application (2SLS)						
# UERJ alumni per 1000 workers at firm	6.561	4.652*** (1.624)	7.120	9.738*** (2.583)	3.334	8.523** (3.580)
# UERJ alumni per 1000 workers in labor market	5.021	1.646** (0.774)	5.333	0.718 (1.000)	3.817	4.639*** (1.353)
# UERJ alumni per 1000 workers at firm (net of labor market)	1.540	3.505** (1.390)	1.786	7.279** (2.989)	−0.483	5.186* (2.867)
N (# UERJ alumni at firm)	2,029	25,324	2,698	34,127	396	5,060
Panel B. Employment 10–13 years after application (2SLS)						
# UERJ alumni per 1000 workers at firm	5.873	3.648*** (1.078)	5.620	5.954** (3.025)	3.224	−0.092 (2.767)
# UERJ alumni per 1000 workers in labor market	4.421	0.433 (0.693)	4.515	0.316 (0.940)	2.960	2.382*** (0.894)
# UERJ alumni per 1000 workers at firm (net of labor market)	1.452	3.179*** (1.159)	1.105	4.552 (2.819)	0.264	−0.821 (1.601)
N (# UERJ alumni at firm)	2,242	25,535	2,039	28,217	277	4,179

Notes: This table presents RD estimates for the effects of UERJ enrollment on the number of UERJ alumni in an applicant's firm and labor market. Columns (A), (C), and (E) show means of each dependent variable for applicants in each group who scored $(-0.1, 0)$ SDs below the cutoff. Columns (B), (D), and (F) show 2SLS RD coefficients, β, from equation (2), which measure the effects of UERJ enrollment on employment outcomes. Panel A measures outcomes 6–9 years after application, and Panel B measures outcomes 10–13 years after application. In the first row of each panel, the dependent variable is the number of alumni per 1000 workers at the applicant's firms (averaged over the time period); this replicates results from Table 4. In the second row, the dependent variable is the number of alumni per 1000 workers in the applicant's labor markets (averaged over the time period). We define labor markets as a municipality × 5-digit industry code, and we compute the number of workers in each firm and labor market using its mean size over all years of our data. In the last row of each panel, the dependent variable is the *difference* between the dependent variables in the first and second rows. Parentheses contain standard errors clustered at the individual level. We use the Calonico et al. (2014) bandwidth for each outcome; sample sizes refer to the dependent variables indicated in the parentheses after N.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A11. Effects of UERJ enrollment on firm, occupation, industry, and municipality mean wages

Dependent variable	(A) 1995–2001 general track	(B)	(C) 2004–2011 general track	(D)	(E) 2004–2011 AA tracks	(F)
	Mean below	RD coef	Mean below	RD coef	Mean below	RD coef

Panel A. Outcomes 6–9 years after application (2SLS)

Dependent variable	Mean below	RD coef	Mean below	RD coef	Mean below	RD coef
Firm mean wage (log)	3.303	0.018 (0.043)	3.475	-0.095^* (0.053)	3.073	0.106^* (0.062)
Occupation mean wage (log)	3.271	0.028 (0.034)	3.388	-0.062 (0.039)	3.017	0.053 (0.041)
Industry mean wage (log)	3.219	0.009 (0.031)	3.201	-0.024 (0.037)	3.000	0.044 (0.036)
Municipality mean wage (log)	3.186	0.009 (0.017)	3.175	-0.013 (0.018)	3.137	-0.005 (0.020)
N (firm mean wage)	2,024	30,345	2,681	31,087	394	4,306

Panel B. Outcomes 10–13 years after application (2SLS)

Dependent variable	Mean below	RD coef	Mean below	RD coef	Mean below	RD coef
Firm mean wage (log)	3.572	0.093^* (0.053)	3.581	-0.053 (0.062)	3.223	0.049 (0.071)
Occupation mean wage (log)	3.428	-0.009 (0.033)	3.509	0.025 (0.041)	3.157	-0.071 (0.053)
Industry mean wage (log)	3.314	0.039 (0.034)	3.271	0.029 (0.040)	3.094	0.018 (0.044)
Municipality mean wage (log)	3.209	-0.025 (0.021)	3.206	-0.027 (0.020)	3.131	0.017 (0.023)
N (firm mean wage)	2,236	24,701	2,010	21,071	275	3,133

Notes: This table presents RD estimates for the effects of UERJ enrollment on mean wages at the firm, occupation, industry, and municipality levels. Columns (A), (C), and (E) show means of each dependent variable for applicants in each group who scored $(-0.1, 0)$ SDs below the cutoff. Columns (B), (D), and (F) show 2SLS RD coefficients, β, from equation (2), which measure the effects of UERJ enrollment on mean hourly wages associated with four different characteristics of individuals' jobs: 1) firm; 2) occupation (using 4-digit CBO codes); 3) industry (using 4-digit CNAE codes); and 4) municipality. Panel A measures outcomes 6–9 years after application, and Panel B measures outcomes 10–13 years after application. Parentheses contain standard errors clustered at the individual level. We use the Calonico et al. (2014) bandwidth for each outcome; sample sizes refer to the dependent variables indicated in the parentheses after N.

TABLE A12. OLS regressions on alumni firm variables

Covariate	(A)	(B)	(C)	(D)
	Dependent variable: Log firm mean hourly wage			
Any alumni firm	0.444 (0.004)			
Firm w/ 50+ alumni per 1000 workers		0.398 (0.011)		
Firm w/ 25−50 alumni per 1000 workers		0.589 (0.010)		
Firm w/ 10−25 alumni per 1000 workers		0.639 (0.006)		
Firm w/ 5−10 alumni per 1000 workers		0.488 (0.007)		
Firm w/ 1−5 alumni per 1000 workers		0.439 (0.005)		
Firm w/ 0−1 alumni per 1000 workers		0.303 (0.005)		
Pre-AA alumni			0.481 (0.004)	
Only post-AA alumni			0.274 (0.005)	
General track alumni from same cohort				0.533 (0.005)
General track alumni from diff. cohort				0.455 (0.004)
Only AA alumni from same cohort				0.293 (0.008)
Only AA alumni from diff. cohort				0.070 (0.010)
Admission score	0.170 (0.002)	0.164 (0.002)	0.173 (0.002)	0.164 (0.002)
N	549,675	549,675	549,675	549,675

Notes: This table shows OLS estimates of the wage premia associated with employment at UERJ alumni firms. The sample includes all UERJ applicants. Regressions are at the applicant × year level and include observations 6–9 years after UERJ application. The dependent variable is log firm mean hourly wage. We use four types of our alumni firm variables as covariates:

- Column (A): Any alumni firm;
- Column (B): Firms categorized by the number of UERJ alumni they hired relative to their mean size;
- Column (C): Firms that hired any alumni from the pre-AA cohorts (1995–2001) vs. firms that hired alumni *only* from the post-AA cohorts (2004–2011);
- Column (D): Firms defined by the alum's cohort (same or different than the applicant's cohort) and application track (general or AA). We define these outcomes to be non-overlapping, i.e., the variables in the lower rows equal one only if the firm did not hire alumni who meet the criteria for the higher rows.

All regressions control for the applicant's standardized admission score and application pool × calendar year dummies. Parentheses contain standard errors clustered at the individual level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Dependent variable	(A) Pre-AA mean	(B) DD coefficients	(C)	(D)
	All top enrollees	All top enrollees	URM top enrollees	Non-top enrollees

Panel A. SES indices based on first and last names

Dependent variable	(A)	(B)	(C)	(D)
First-name-based log wage	2.424	0.004 (0.003)	−0.004 (0.016)	−0.008** (0.004)
First-name-based wage	4.464	0.037 (0.022)	0.006 (0.140)	−0.080*** (0.029)
Last-name-based log wage	2.400	−0.006 (0.007)	−0.007 (0.019)	−0.026*** (0.007)
Last-name-based wage	4.406	−0.071 (0.052)	−0.052 (0.142)	−0.191*** (0.063)

Panel B. Predicted log wage based on characteristics, scores, and SES name indices

Dependent variable	(A)	(B)	(C)	(D)
Predicted log wage	3.303	−0.023 (0.028)	−0.012 (0.048)	−0.163*** (0.043)
Predicted log wage (if in RAIS)	3.255	−0.034 (0.028)	−0.011 (0.052)	−0.156*** (0.043)
N (enrollees)	16,466	35,866	1,631	30,854

Notes: This table displays DD estimates of the effect of affirmative action exposure on student SES as measured by first and last names. Column (A) shows the mean of each dependent variable for all top enrollees in the 1995–2001 cohorts. Columns (B)–(D) display estimates of π from equation (3) for all top enrollees, URM top enrollees, and non-top enrollees.

In Panel A, the dependent variables are SES indices defined based on the applicant's first or last name. To define these indices, we take the nationwide RAIS data for the year 2019 and compute the leave-individual-out mean hourly wage associated with each first name and each last name. We exclude first and last names that appear fewer than ten times. The dependent variables in Panel A are these first and last name mean wages in both logs and levels.

In Panel B, the dependent variables are the predicted values from a regression of log hourly wage (6–9 years after application) on each of the variables in Panels B–C of Table 5 plus the SES name indices from Panel A of this table. These are similar to the dependent variables in Panel D of Table 5, except these predicted wages are also based on the SES name indices.

Parentheses contain standard errors clustered at the program level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A14. DD estimates for graduation, employment, and earnings 10–13 years after application

	(A)	(B)	(C)	(D)
	Pre-AA mean	DD coefficients		
Dependent variable	All top enrollees	All top enrollees	URM top enrollees	Non-top enrollees
Panel A. Graduation and formal employment				
Graduated from UERJ program	0.568	0.011 (0.018)	−0.003 (0.071)	0.010 (0.028)
Formal employment	0.768	−0.013 (0.011)	−0.053 (0.044)	−0.010 (0.013)
Panel B. Earnings				
Log hourly wage	3.600	−0.115** (0.053)	−0.129 (0.138)	−0.252*** (0.072)
Monthly earnings (2019 USD)	2,005.914	−224.443** (90.068)	−83.051 (189.155)	−469.037*** (133.630)
Firm mean hourly wage (log)	3.565	−0.114** (0.044)	−0.185 (0.142)	−0.191*** (0.055)
Panel C. Employment at firms with pre- and post-AA alumni				
Pre-AA alumni	0.595	−0.043 (0.027)	0.011 (0.084)	−0.023 (0.034)
Only post-AA alumni	0.089	0.034** (0.016)	−0.070 (0.059)	0.034 (0.021)
Panel D. Alumni firm employment by application track and cohort				
General track alumni from same cohort	0.463	−0.093*** (0.026)	−0.121 (0.076)	−0.046 (0.030)
General track alumni from diff. cohort	0.248	0.046** (0.018)	0.031 (0.078)	0.023 (0.017)
Only AA alumni from same cohort	0.000	0.031*** (0.006)	0.024 (0.018)	0.036*** (0.009)
Only AA alumni from diff. cohort	0.012	0.005 (0.004)	−0.016 (0.015)	0.003 (0.004)
N (enrollees)	16,466	31,016	1,157	26,484
N (wage observations)	12,614	23,381	928	20,091

Notes: This table displays DD estimates of the effect of affirmative action exposure on graduation, earnings, and employment at alumni firms measured 10–13 years after application. Column (A) shows the mean of each dependent variable for all top enrollees in the 1995–2001 cohorts. Columns (B)–(D) display estimates of π from equation (3) for all top enrollees, URM top enrollees, and non-top enrollees. The dependent variables are defined similarly to those in Tables 2 and 4. In Panel C, we define alumni firms using graduates from the pre- and post-AA cohorts. In Panel D, we define alumni firms using the alum's cohort (same or different than the applicant's cohort) and application track (general or AA). We define the outcomes in Panels C–D to be non-overlapping, i.e., the variables in the lower rows of each panel equal one only if the firm did not hire alumni who meet the criteria for the higher rows. Parentheses contain standard errors clustered at the program level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A15. DD estimates by field of study — Top enrollees

Dependent variable	(A) Business	(B) Human -ities	(C) Natural sciences	(D) Social sciences
Panel A. Graduation and formal employment				
Graduated from UERJ program	−0.011	0.046	−0.020	0.052
	(0.043)	(0.034)	(0.039)	(0.025)
Formal employment	0.022	−0.021	−0.029	−0.030
	(0.032)	(0.029)	(0.018)	(0.026)
Panel B. Earnings				
Log hourly wage	−0.035	−0.212**	−0.082	−0.181*
	(0.040)	(0.094)	(0.071)	(0.070)
Monthly earnings (2019 USD)	−118.732	−211.522***	−86.695	−301.923
	(91.622)	(60.937)	(77.045)	(185.385)
Firm mean hourly wage (log)	−0.100**	−0.137**	−0.066	−0.079
	(0.018)	(0.055)	(0.072)	(0.058)
Panel C. Employment at firms with pre- and post-AA alumni				
Pre-AA alumni	−0.072	−0.026	−0.062	−0.158***
	(0.034)	(0.032)	(0.037)	(0.023)
Only post-AA alumni	0.035	0.047	0.059	0.088**
	(0.054)	(0.038)	(0.034)	(0.026)
Panel D. Alumni firm employment by application track and cohort				
General track alumni from same cohort	−0.057**	−0.095**	−0.103**	−0.151**
	(0.008)	(0.036)	(0.036)	(0.034)
General track alumni from diff. cohort	−0.049**	0.077***	0.036	0.071**
	(0.007)	(0.021)	(0.024)	(0.016)
Only AA alumni from same cohort	0.043*	0.023	0.046***	0.044**
	(0.011)	(0.020)	(0.008)	(0.010)
Only AA alumni from diff. cohort	0.019	0.011	0.016***	−0.017**
	(0.012)	(0.011)	(0.004)	(0.006)
N (enrollees)	2,895	13,410	11,266	5,445
N (wage observations)	2,343	10,312	8,409	3,515

Notes: This table examines heterogeneity by field of study in our DD estimates for top enrollees. The sample and dependent variables are the same as in column (B) of Table 6, but we estimate regressions for programs in four field of study groups: (A) business, (B) humanities, (C) natural sciences, and (D) social sciences. See Table 1 for the programs included in each field of study group. We cannot estimate our DD specification for health programs because there is no variation in our binary measure of exposure to affirmative action. All outcomes are measured 6–9 years after UERJ application.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A16. Robustness of DD estimates — Top enrollees

Dependent variable	(A) Bench-mark	(B) Pre-crisis years	(C) Linear trends	(D) Demo-graphics	(E) Selectivity controls	(F) No field of study	(G) Actual AA share
Panel A. Graduation and formal employment							
Graduated from UERJ program	0.013 (0.021)	0.004 (0.019)	0.037* (0.021)	0.013 (0.021)	0.011 (0.020)	0.039* (0.021)	0.006 (0.020)
Formal employment	−0.027* (0.015)	−0.022 (0.014)	−0.039** (0.019)	0.005 (0.008)	−0.015 (0.021)	−0.044** (0.020)	−0.023 (0.015)
Panel B. Earnings							
Log hourly wage	−0.132*** (0.045)	−0.118** (0.047)	−0.133** (0.061)	−0.110** (0.043)	−0.067** (0.033)	−0.152** (0.061)	−0.125** (0.052)
Firm mean hourly wage (log)	−0.095** (0.035)	−0.094*** (0.032)	−0.120*** (0.043)	−0.075** (0.033)	−0.059** (0.028)	−0.125*** (0.044)	−0.082* (0.043)
Panel C. Employment at firms with pre- and post-AA alumni							
Pre-AA alumni	−0.055** (0.023)	−0.061*** (0.020)	−0.072*** (0.026)	−0.050* (0.025)	−0.035* (0.021)	−0.064** (0.024)	−0.064*** (0.021)
Only post-AA alumni	0.049** (0.023)	0.044** (0.020)	0.080*** (0.025)	0.048** (0.023)	0.017 (0.018)	0.073*** (0.024)	0.055** (0.024)
Panel D. Alumni firm employment by application track and cohort							
General track alumni from same cohort	−0.098*** (0.021)	−0.094*** (0.016)	−0.093*** (0.024)	−0.096*** (0.022)	−0.090*** (0.022)	−0.087*** (0.025)	−0.091*** (0.023)
General track alumni from diff. cohort	0.042** (0.016)	0.037** (0.017)	0.046*** (0.016)	0.045** (0.017)	0.034 (0.021)	0.044** (0.017)	0.029 (0.018)
Only AA alumni from same cohort	0.036*** (0.009)	0.029*** (0.009)	0.041*** (0.007)	0.036*** (0.008)	0.029*** (0.008)	0.041*** (0.007)	0.035*** (0.008)
Only AA alumni from diff. cohort	0.010** (0.005)	0.007 (0.005)	0.007 (0.004)	0.010** (0.005)	0.006** (0.003)	0.008* (0.004)	0.011** (0.005)
N (enrollees)	35,866	28,591	35,866	35,866	35,866	35,866	35,866
N (wage observations)	26,445	20,889	26,445	26,445	26,445	26,445	26,445

Notes: Column (A) reproduces our benchmark DD results for top enrollees (column B in Table 6). Column (B) includes only outcomes measured in 2003–2014. Column (C) includes program-specific linear trends estimated in the 1995–2001 cohorts. Column (D) includes controls for age, gender, race, qualifying exam score, and writing field exam score. Column (E) includes cohort dummies interacted with dummies for quartiles of program selectivity (x-axis of Figure 1). Column (F) excludes the field of study group interactions, $f(m)$. Column (G) defines ExposureToAA$_m$ as each major's affirmative action share in the 2004–2011 cohorts (y-axis of Figure 1), scaled to represent a 20 percentage point increase. Parentheses contain standard errors clustered at the program level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A17. Number of students in Enade sample by exam year, field, and university

| | Exam year | | | | | | | | | | | | |
Exam field	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Total
Dentistry	270			156			415			475			1,316
Medicine	342			228			636			738			1,944
Nursing	385			258			479			426			1,548
Nutrition	158			117			255			317			847
Physical education	251			162			141	176		224	288		1,242
Social work	113			116			138			283			650
Biology		647			474			1,014			1,263		3,398
Chemistry		213			262			364			416		1,255
Computation		400			316			373			633		1,722
Engineering I (Civil)		342			454			441			1,036		2,273
Engineering II (Electrical)		333			340			522			730		1,925
Engineering III (Mechanical)		185			205			381			465		1,236
Engineering IV (Chemical)		173			120			180			291		764
Engineering VI (Production)		108			164			270			365		907
Geography		506			531			708			1,098		2,843
History		526			509			600			1,163		2,798
Language		861			905			1,247			1,336		4,349
Math		388			451			420			602		1,861
Philosophy		96			108			140			119		463
Physics		192			185			139			325		841
Social science		286			193			316			399		1,194
Teaching		835			1,396			4,584			3,807		10,622
Accounting			194			614			428			610	1,846
Business			710			1,029			1,018			1,128	3,885
Design			198			288			211			320	1,017
Economics			392			410			305			574	1,681
Law			537			1,226			1,229			1,574	4,566
Psychology			210			266			159			311	946
Social communication			288			316			184			385	1,173
UERJ	141	676	242	124	718	294	192	1,086	253	304	1,643	512	6,185
Other federal & state universities	1,378	5,415	2,287	913	5,895	3,855	1,872	10,789	3,281	2,159	12,693	4,390	54,927
Full sample	1,519	6,091	2,529	1,037	6,613	4,149	2,064	11,875	3,534	2,463	14,336	4,902	61,112

Notes: This table shows the number of students in our Enade sample for Table 7. The sample includes 2004–2015 Enade exam takers at UERJ and other federal and state universities that did not implement affirmative action until 2012 or later. We define each university's year of affirmative action adoption as the first year in which more than 10 percent of new enrollees entered through a reserved quota in the *Censo da Educação Superior* data, or, if it is earlier, the year of affirmative action adoption cited in Júnior and Daflon (2014) or Vieira and Arends-Kuenning (2019). The federal universities in our sample are: UFAC, UFAM, UFC, UFCG, UFCSPA, UFERSA, UFLA, UFMS, UFMT, UFPEL, UFRR, UFV, UFVJM, UNIFAL-MG, UNIFAP, UNIFEI, UNIR, and UNIRIO. The state universities in our sample are: UECE, UERJ, UERR, UNESP, UNITINS, and URCA. We exclude fields with no UERJ exam takers, and we drop any institution × major pair that does not have exam takers in every year in which the exam was offered.

TABLE A18. DDD estimates for UERJ's Enade exam taker characteristics and scores

	(A)	(B)	(C)	(D)
	UERJ pre-AA mean	DDD coefficients		
			White	
	All	All	private HS	Other
Dependent variable	students	students	students	students

Panel A. Characteristics of Enade exam takers

# exam takers	38.250	−5.024	−8.316***	−4.539
		(3.766)	(1.346)	(3.958)
White	0.750	−0.180***		−0.138***
		(0.017)		(0.014)
Private HS student	0.588	−0.133***		−0.061***
		(0.019)		(0.011)
Female	0.511	0.075***	0.026	0.122***
		(0.009)	(0.019)	(0.016)
Age	26.028	0.397	1.018**	0.101
		(0.280)	(0.389)	(0.374)
Mother has a HS degree	0.726	−0.069***	−0.008	−0.051**
		(0.016)	(0.030)	(0.020)
HH income / minimum wage	8.619	−1.706***	−0.082	−2.102***
		(0.152)	(0.205)	(0.121)
Predicted general score	0.567	−0.012***	−0.002	−0.012***
		(0.001)	(0.002)	(0.001)

Panel B. Enade scores (proportion correct answers)

Overall score	0.566	−0.024***	−0.030**	−0.027***
		(0.008)	(0.011)	(0.009)
Field-specific component	0.535	−0.020**	−0.035***	−0.020**
		(0.009)	(0.011)	(0.009)
General component	0.661	−0.035***	−0.014	−0.048***
		(0.008)	(0.014)	(0.007)

N (programs × years)	36	1,664	1,664	1,664
N (exam takers)	747	61,112	16,851	37,992

Notes: This table displays triple-difference (DDD) estimates of the effect of affirmative action on the characteristics (Panel A) and scores (Panel B) of UERJ's Enade exam takers. The sample includes 2004–2015 Enade exam takers from UERJ and other federal and state universities that did not implement affirmative action until 2012 or later. See Appendix Table A17 for details on our Enade sample and the exam fields. Column (A) shows dependent variable means for UERJ exam takers in 2004–2006. The other columns show θ coefficients from the DDD regression:

$$Y_{mjt} = \gamma_{mj} + \gamma_{mt} + \pi_{f(m)}[\text{UERJ}_j \times \text{Post}_t] + \theta[\text{UERJ}_j \times \text{Post}_t \times \text{ExposureToAA}_m] + \varepsilon_{mjt}.$$

Regressions are at the exam field (m) by institution (j) by year (t) level, with observations weighted by the number of exam takers. (In the first two rows of Panel A, we weight by the number of 2004–2006 exam takers in each mj cell.) We include field × institution dummies, γ_{mj}, field × year dummies, γ_{mt}, and an indicator for UERJ interacted with an indicator for the 2007–2015 cohorts, $\text{UERJ}_j \times \text{Post}_t$. We interact $\text{UERJ}_j \times \text{Post}_t$ with dummies for the field of study groups listed in Table 1, $f(m)$ (i.e., business, health, humanities, natural sciences, and social sciences). The variable of interest is $\text{UERJ}_j \times \text{Post}_t \times \text{ExposureToAA}$, where ExposureToAA is an indicator for UERJ programs in which the share of affirmative action enrollees in the 2004–2011 cohorts was 30 percent or higher. Columns (B)–(D) include all students, white private high school students, and non-white and/or public school students. "Predicted general score" is the predicted value from a regression of general component scores on age and dummies for gender, race, private high school, mother's education, father's education, and family income. Parentheses contain standard errors clustered at the institution level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE A19. DD estimates for Enade exam taker characteristics of UERJ's competitors

	(A)	(B)	(C)	(D)
	2004–06 mean	DD coefficients		
Dependent variable	UFRJ and UFF	UFRJ and UFF	UFRJ	UFF
White	0.750	0.009	0.017**	−0.004
		(0.011)	(0.007)	(0.009)
Private HS student	0.662	−0.019	−0.030*	−0.005
		(0.018)	(0.016)	(0.015)
White private HS student	0.516	−0.008	−0.011	−0.004
		(0.014)	(0.014)	(0.013)
Female	0.532	0.008	0.007	0.010
		(0.011)	(0.011)	(0.011)
Age	25.746	0.190	0.288	0.000
		(0.250)	(0.200)	(0.262)
Mother has a HS degree	0.758	−0.013	−0.026**	0.007
		(0.017)	(0.012)	(0.012)
Father has a HS degree	0.776	−0.026**	−0.021*	−0.036***
		(0.011)	(0.011)	(0.011)
HH income / minimum wage	8.510	−0.446***	−0.413***	−0.506***
		(0.093)	(0.124)	(0.077)
Predicted general score	0.568	0.002*	0.002*	0.002*
		(0.001)	(0.001)	(0.001)
N (programs × years)	61	1,729	1,601	1,611
N (exam takers)	2,372	68,267	62,398	58,270

Notes: This table displays DD estimates of the effect of UERJ's adoption of affirmative action on the characteristics of Enade exam takers at two of its competitor schools: UFRJ and UFF. The sample includes 2004–2015 Enade exam takers from UFRJ, UFF, and other federal and state universities nationwide that did not implement affirmative action until 2012 or later. See Appendix Table A17 for details on our Enade sample and the exam fields. Column (A) shows dependent variable means for UFRJ and UFF exam takers in 2004–2006. The other columns show π coefficients from the DD regression:

$$Y_{mjt} = \gamma_{mj} + \gamma_{mt} + \pi[\text{Competitor}_j \times \text{Post}_t] + \varepsilon_{mjt}.$$

Regressions are at the exam field (m) by institution (j) by year (t) level, with observations weighted by the number of exam takers (except in the first row of Panel A, where we weight by the number of 2004–2006 exam takers in each mj cell.) We include dummies for field × institution, field × year, and the interaction between indicators for UFRJ/UFF and the 2007–2015 cohorts (Competitor$_j$ × Post$_t$). In column (B), the sample includes UFRJ, UFF, and all schools in the control group. Column (C) includes only UFRJ plus control schools, and column (D) includes only UFF plus control schools. "Predicted general score" is the predicted value from a regression of general component scores on age and dummies for gender, race, private high school, mother's education, father's education, and family income. Parentheses contain standard errors clustered at the institution level.

  * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B. Empirical appendix

B.1. **Variable definitions.** This section describes the main variables in our paper.

- **Admission score.** Raw admission scores are based on applicants' subject scores on different subjects of the field exam (*Exame discursivo*), plus bonus points from their qualifying exam performance (*exame de qualificação*). We standardize raw admission scores to represent an individual's distance from the admission cutoff in their application pool in SD units. For this, we subtract the score of the last admitted student in the application pool, and divide by the SD of scores for all applicants to the same program/cohort. We adjust these SDs to be comparable across cohorts because the number of field exam takers varies significantly over time due to changes in UERJ's standards for the qualification exam.
- **Alumni firm employment.** We define UERJ applicant i to major m as obtaining a job at an alumni firm if their firm ever employed another individual j who graduated from major m (the "alum"). We define different versions of this variable based on the alum's cohort, application track, and year of employment. We also define versions that classify firms based on the number of alumni they hired as a proportion of their mean size over all years of our data.
- **Demographic characteristics.** From the UERJ admission data, we observe age at application, gender, race, mother's education, and household income. Age is available in all cohorts; other characteristics are available only in certain cohorts (see Appendix Table A1). These variables come from a survey that applicants completed as part of the application process. We also use gender and race from the RAIS data, which we observe for any applicant who appears in this dataset. We use indicators for three racial groups: *branco* (white), *pardo* (brown), and *preto* (Black).
- **Field exam subject scores.** An applicant's scores on subjects of the field exam (*exame discursivo*). We use an applicant's writing exam score (which is common to all applicants), and their mean score across 2–4 other subjects (which vary depending on the cohort and major they are applying to). We observe field exam subject scores in the 1995–2001 and 2007–2011 cohorts.
- **Firm.** We define firms at the establishment level. Establishments are identified by their 14-digit CNPJ (short for *Cadastro Nacional da Pessoa Jurídica*, or National Registry of Legal Entities). The CNPJ is a tax identifier for legally incorporated identities. The first eight digits identify the company. The rest of the digits identify the branch or subsidiary of the company.
- **Firm mean wage.** The leave-individual-out mean hourly wage at a given firm.
- **Firm size.** The total number of workers employed by the firm at the end of each year.

- **Formal employment.** An indicator that takes the value one if an applicant appears in the employee-employer matched dataset (RAIS).
- **Hourly wage.** We calculate the hourly rate of each worker as the ratio between a worker's inflation-adjusted monthly earnings and the hours worked per month. Hours worked reflects the number of hours per week at which the firm hired the worker according to the worker's contract, which may differ from the hours actually worked in any given week.
- **Industry mean wage.** The leave-individual-out mean hourly wage of all UERJ applicants working in a given industry. We define industries at the 4- or 5-digit level of the Brazilian National Classification of Economic Activities (*Classificação Nacional de Atividades Economicas*) level.
- **Monthly earnings (2019 USD).** This variable represents a worker's average monthly salary in a given year. To report this variable, establishments have to calculate the worker's total earnings for the year and divide them by the number of months the firm employed the worker. We adjust earnings for inflation using the consumer price index. We express earnings in 2019 US dollars using the Brazilian Real/US Dollar exchange rate.
- **Municipality mean wage.** The leave-individual-out mean hourly wage of all UERJ applicants working at a given municipality. Municipalities are defined by the location of the worker's establishment.
- **Occupation mean wage.** The leave-individual-out mean hourly wage of all UERJ applicants with a given occupation. We define occupations at the 4-digit of the Brazilian Occupational Code Classification (*Classificação Brasileira de Ocupações*) level.
- **Qualifying exam score.** An applicant's standardized score from the qualifying exam (*exame de qualificação*). This exam includes eight subject tests common to all applicants: Biology, Chemistry, Geography, Foreign language (French, Spanish, or English), History, Literature/Portuguese, Mathematics, and Physics. Students that fail to achieve a minimum score on the qualifying exam cannot take the second round discursivo exam. We observe qualifying exam scores in the 1995–2001 cohorts.

We measure graduation and labor market outcomes in two time periods: 6–9 years after application, and 10–13 years after application. For earnings and wage indices, we use the mean value over each four-year period. For binary variables, we use the maximum value over the four year period.

B.2. **Data and merging.** Our base dataset includes a list of all individuals who passed UERJ's first-round admission exam and applied to an undergraduate program in the years 1995–2001 and 2004–2011 (UERJ, 2020a). This dataset includes the program(s)/cohort that each individual applied to, their admission score in the second exam of the admission process, and their admission decisions. The 2004–2011 records include the track each applicant

applied through. In addition, we have access to socioeconomic variables for the 2007–2011 application cohorts.

We combine the UERJ admission records with two individual-level administrative datasets. The first dataset is from UERJ, and it includes the graduation outcomes of all the students who enrolled in UERJ in 1995–2011 (UERJ, 2020b). These records contain the student's program, enrollment date, status as of December 2020 (i.e., graduated, dropped out, or still enrolled), and final year in the program.

The second administrative dataset is called the *Relação Anual de Informações Sociais*, or RAIS (RAIS, 2021). It includes employment outcomes collected by the Ministry of Labor. We have access to the RAIS for the 2003–2019 period. This dataset has information on all workers with a formal-sector job. The RAIS contains information about both the worker and the firm. Worker information includes demographic variables (e.g., age, gender, and race), educational attainment, occupation, hours worked, and earnings. Firm-level variables include the number of employees, industry, and geographic location.

We merge the admission data with the graduation records using the university ID of each individual. Most individuals match uniquely on the ID, but in cases with duplicated IDs, we corroborate the quality of the matches using individuals' names and programs. We fix a few cases in which different individuals have the same university ID. We match 94.8 percent of individuals in the graduation records to the admission records using the university IDs. We use the names and application years of the remaining unmatched individuals to match them to the graduation records. Overall, we match 97.8 percent of the individuals in the graduation records to the admission records.

Lastly, we link the combined dataset from the above merge to the RAIS dataset using individuals' national ID numbers (*Cadastro de Pessoas Físicas*, or CPF for short), birth dates, and names. For this, we follow a two-step procedure. First, we match individuals for whom we have the CPF available in the UERJ records.[34] Second, for individuals who remain unmatched, we merge them using their names and dates of birth. We define a match from this process as observations that have either: 1) the same CPF number; or 2) the same birth date and an exact name match. We match 77.4 of the individuals in merged UERJ records to at least one year of the RAIS through this process. Out of the matched individuals, 66.1 percent were matched using the CPF, and the remaining 33.9 percent were matched using names and dates of birth.

One way to benchmark the merge rate with the RAIS is to compare it with the share of individuals with similar demographic characteristics who have a formal-sector job in Brazil. To do this, we use data from the 2015 Brazilian household survey (*Pesquisa Nacional por*

---

[34] The UERJ records contain the CPF nearly all individuals who applied in 2000–2001 and 2004–2011. Before 2000, the CPF is rarely available. Virtually all workers in the 2003–2019 RAIS datasets have a CPF.

TABLE B1. Timeline of events during the 2010 admission process

Event	Date
First date for applicants to take the qualifying exam	06/21/2009
Results of the qualifier exam are published	07/01/2009
Second date for applicants to take the qualifying exam	09/13/2009
Results of the qualifier exam are published	09/23/2009
Applicants who passed the qualifier exam take the field exam	12/13/2009
Results of the field exam are published	01/16/2010
Results of the field exam are published	01/30/2010
First round of admission offers is sent	01/30/2010
Second round of admission offers is sent	02/12/2010
Admitted students can enroll in first-semester programs	03/02/2010 – 03/03/2010
First day of classes - 1st semester	**03/10/2010**
Third round of admission offers is sent	03/16/2010
Fourth round of admission offers is sent	07/02/2010
Fifth round of admission offers is sent	07/16/2010
Newly admitted applicants can enroll in second-semester programs	07/28/2010 – 07/29/2010
First day of classes - 2nd semester	**08/10/2010**

Notes: This calendar is summarized from information in these two UERJ documents:
- http://sistema.vestibular.uerj.br/portal_vestibular_uerj/arquivos/arquivos2010/ed/03_anexo1_WEB.pdf
- http://sistema.vestibular.uerj.br/portal_vestibular_uerj/arquivos/arquivos2010/calendario/calendario_eq.pdf

Amostra de Domicílios, abbreviated PNAD), which includes information on the informal economy. Our proxy of working in the formal sector is having the right to a pension when retired.[35] The share of economically active individuals aged 25–37 with at least a high-school degree who have a job in the formal sector is 62.4 percent. This suggests that our merge identified most individuals with formal sector jobs.

B.3. **UERJ's admission process.** Applicants can gain admission to UERJ at one of several stages. The admission process begins with applicants taking a common qualifying exam. Applicants who pass this exam then take a field exam. UERJ ranks applicants based on their field exam scores and sends admissions offers to accepted applicants up to the capacity of each program. The remaining applicants are either rejected (if their score in the field exam is below a minimum threshold) or waitlisted. The first admission offers are typically sent in January, and admitted students have several weeks to accept or reject their offer. UERJ sends a second round of admission offers to waitlisted applicants based on the number of offers that were declined. This process is repeated up to five times per application year

---

[35] International organizations define informality in two different ways. Under the *legal* definition, a worker is considered informal if she does not have the right to a pension when retired. An alternative to the legal definition is the *productive* definition, where a worker is considered informal if she is a salaried worker in a small firm (i.e., it employs less than five workers), a non-professional self-employed, or a zero-income worker. We use the legal definition in the main text. The share of workers with a formal job under the productive definition is slightly lower than the one based on the legal definition.

if there are remaining open seats, and the last admission offers may occur as late as July. Appendix Table B1 provides an example of this process for the 2010 cohort.

The admission thresholds in our RD analysis are given by the admission score of the final student who gained admission in each application pool (after all waitlist offers). Any applicant who scored above this threshold could have been admitted to UERJ, although some of these students chose to enroll in other universities by the time they would have gotten in off the waitlist. Potential for non-random sorting around the admission cutoff arises because applicants have control over whether they accept or reject their admission offer. Students just above the final cutoff may therefore be those who particularly want to attend UERJ. We present tests for non-random sorting around the admission cutoff in Section 2.2.

B.4. **Sample.** Our initial dataset includes all applicants to UERJ undergraduate majors who passed the first-round qualifying exam and who have a valid second-round admission score (i.e., non-missing, non-zero). UERJ has several campuses; its main campus is in the municipality of Rio de Janeiro, and it has five smaller campuses in other municipalities in the state: Baixada Duque de Caixas (DDC), Nova Friburgo (NF), Resende (RES), São Gonçalo (SGO), and Teresópolis (TER). The number of undergraduate programs changes across cohorts of our data because UERJ split some large programs into smaller "sub-programs" and added some new majors.

Our raw data includes 71 different sub-programs across all cohorts and campuses. We group these 71 sub-programs into 43 programs to create a consistent set over time. We create these groups using documentation from UERJ detailing how large programs were divided into sub-programs. We exclude six new majors that UERJ created after the introduction of affirmative action: computing engineering (NF), geography (DDC), math (DDC), pedagogy (SGO), tourism (TER), and actuarial sciences (RIO). Appendix Tables B2-B4 show the 43 programs in our data and the sub-programs that they are derived from.

We use data from these 43 programs to create two different samples to analyze the impacts of UERJ's affirmative action policy. For our RD sample, we exclude programs where fewer than 30 percent of the 2004–2011 students entered through an affirmative action track. The second column in Appendix Tables B2-B4 shows the percent of students that entered through an affirmative action track in each program group during 2004–2011. Bolded figures denote programs where this figure is above 30 percent. 24 programs meet this criteria. Within these programs, we also exclude program-cohort-admission track triplets with fewer than five applicants below the admission threshold. We also exclude all applicants to the disabled/indigenous track since these quotas rarely filled up. In Appendix Tables B2-B4, we highlight in bold the program-cohort pairs in each admission track that satisfy our sample restrictions and appear in our RD sample.

For our DD sample, we focus on applicants who *enrolled* in UERJ. Our DD sample includes the 24 programs in our RD sample plus 19 other programs with lower take-up rates in the affirmative action tracks. These programs are unbolded in Appendix Tables B2-B4.

TABLE B2. Number of applicants by cohort — General track

#	Program	Prop. AA	Program name(s)	1995	1996	1997	1998	1999	2000	2001	2004	2005	2006	2007	2008	2009	2010	2011
1	Accounting	**0.364**	Accounting	**351**	**463**	**450**	**471**	**476**	**469**	**1160**	**350**	**442**	**502**	**374**	**484**	**492**	**551**	**492**
2	Art	0.287	Artistic education	198	233	210	230	235	234	547								
			Art								384	413	402					
			Art history											114		85	160	127
			Visual arts (bach.)											127				
			Visual arts (license)											125				
			Visual arts												326	334	292	328
3	Biology	**0.494**	Biology	**194**	**295**	**225**	**292**	**297**	**351**	**1899**	**659**	**1059**	**1156**	**973**	**873**	**1160**	**1028**	**1148**
4	Biology (SGO)	0.260	Biology	94	229	151	236	222	235	643	271	246	380	274	325	252	209	227
5	Business	**0.428**	Business	**466**	**583**	**459**	**590**	**590**	**593**	**2200**	**537**	**964**	**983**	**824**	**864**	**1071**	**943**	**1108**
6	Cartographic eng.	0.126	Cartographic eng.	43	112	86	119	117	117	185	69	79	129	104	115	156	218	148
7	Chemical eng.	**0.465**	Chemical eng.						317	897	420	662	811	838	817	1149	1128	1290
8	Chemistry	**0.352**	Chemistry	**352**	**474**	340	**352**	**353**	160	408	212	206	336	319	317	363	349	321
9	Computer science	0.325	Information science	**633**	**705**	**567**	**587**	**590**	**593**	**2029**	**592**	**699**	**775**	**637**	**603**	**665**	**548**	
			Computer science															**742**
10	Dentistry	**0.404**	Dentistry	**357**	**357**	**350**	**356**	**356**	**359**	**1235**	**450**	**447**	**632**	**446**	**458**	**441**	**503**	**605**
11	Economics	0.286	Economics	442	541	408	539	548	547	1355	529	640	664	538	532	752	709	754
12	General eng.	**0.307**	Engineering	1065	**1252**	**1182**	**1407**	**1410**	**1417**									
			Civil eng.							736	291	512	574	511	691	908	905	1310
			Electrical eng.							2409	614	922	1070	695	765	1048	1066	1109
			Textile eng.	18	**153**	42												
13	Geography	**0.468**	Geography	**113**	**119**	**115**	**157**	**198**	**196**	**943**	**337**	**578**	**524**	**523**	**587**	**532**	**544**	**523**
14	Geog. Ed. (SGO)	0.275	Geography	137	214	131	232	237	237	591	291	366	369	259	259	334	225	202
15	Geology	**0.321**	Geology	28	**84**	**65**	**88**	**89**	**86**	**203**	**94**	**144**	**216**	**185**	**264**	**409**	**317**	**329**
16	History	**0.457**	History	**383**	**298**	**288**	**392**	**395**	**393**	**1991**	**723**	**981**	**1007**	**863**	**719**	**828**	**830**	**772**
17	Hist. Ed. (SGO)	**0.333**	History	**202**	**234**	**185**	**238**	**236**	**236**	**592**	**306**	**311**	**412**	**318**	**264**	**292**	**203**	**226**
18	Industrial design	**0.456**	Industrial design	**157**	**174**	**171**	**208**	**175**	**207**	**836**	**357**	**480**	**655**	**517**	**539**	**627**	**627**	**699**
19	Journalism	**0.481**	Social communication	**350**	**355**	**350**	**358**											
			Journalism					237	239	1679	528	717	803	737	672	940	777	1130
			Public relations					159	239	948	311	469	552	500	479	645	619	717
20	Language (SGO)	0.205	Language	298	351	326	469	470	475	910	332	334	395	208	246	251	180	233
21	Language I	0.259	Language I	282	267	263	296	292	294	978								
			Literature/English								403	328	403	325	285	318	281	295
			Port./German								48	81	60	49	36	52	46	42
			Port./Japanese								48	65	66	58	62	65	36	35

Table B2. Number of applicants by cohort — General track *(continued)*

#	Program	Prop. AA	Program name(s)	1995	1996	1997	1998	1999	2000	2001	2004	2005	2006	2007	2008	2009	2010	2011
22	Language II	0.293	Language II	317	350	363	388	388	513	1117								
			Port./France								77	166	74	137	78	99	104	78
			Port./Italian								93	165	87	95	80	77	86	74
			Port./Spanish								245	200	312	220	206	232	184	216
23	Language III	0.379	Language III	255	322	305	295	296	294	1212								
			Port./Greek								30	21	63	21	29	10	28	15
			Port./Latin								57	71	50	65	36	47	31	44
			Port./Literature								454	348	460	338	295	330	294	305
24	Law	0.460	Law	1455	1477	1479	1483	1496	1486	5940	2079	2271	3182	2487	2468	2909	2884	3734
25	Math	0.158	Math	243	354	537	586	592	589	1181	490	506	515	337	367	367	322	302
26	Math Ed. (SGO)	0.143	Math	88	198	160	235	232	235	283	148	138	142	111	121	118	86	98
27	Mech. eng.	0.353	Mech. eng.							516	206	371	508	585	505	728	680	822
28	Mech. eng. (NF)	0.186	Mech. eng.					170	228	143	74	144	238	196	285	334	315	380
29	Medicine	0.454	Medicine	541	546	546	550	552	551	4122	1473	1749	2754	2838	2025	2639	2669	3971
30	Nursing	0.431	Nursing	187	312	274	313	311	389	1478	563	565	762	475	585	536	507	499
31	Nutrition	0.411	Nutrition	236	316	368	311	390	444	2203	566	706	818	561	599	689	516	661
32	Oceanography	0.229	Oceanography	70	90	78	110	112	116	279	121	483	367	262	268	271	270	228
33	Teaching	0.234	Teaching	557	614	577	623	616	621	1852	724	880	809	511	464	555	451	509
34	Teaching (DDC)	0.208	Teaching	267	251	267	354	347	343	637				143	155	151	138	183
			Teaching I								201	160	225					
			Teaching II								75	51	121					
35	Philosophy	0.222	Philosophy	246	292	264	288	288	287	593	362	272	381	276	228	251	199	186
36	Physical ed.	0.206	Physical education	177	236	238	290	295	352	1611	447	506	600	413	405	470	384	413
37	Physics	0.135	Physics	196	321	299	402	432	434	664	295	410	472	397	289	378	325	320
38	Prod. eng.	0.384	Prod. eng.							694	294	466	539	572	526	696	578	792
39	Prod. eng. (RES)	0.152	Prod. eng.	144	284	264	288	290	293	338	185	284	375	301	356	400	372	394
40	Psychology	0.480	Psychology	268	390	355	394	391	383	1527	719	887	984	789	800	920	854	1138
41	Social science	0.408	Social sciences	283	297	285	294	294	293	1311	472	594	631	628	448	547	468	440
42	Social work	0.432	Social work	261	276	266	276	275	274	1240	315	374	511	325	324	348	348	427
43	Statistics	0.087	Statistics	148	286	343	465	468	465	475	203	107	303	134	179	139	190	149

Notes: This table displays the number of applicants in the general track for each program/cohort in our data. See Table B4 for details on the table structure and statistics.

#	Program	Prop. AA	Program name(s)	2004	2005	2006	2007	2008	2009	2010	2011
1	Accounting	**0.364**	Accounting	74	50	**71**	**41**	**31**	**37**	**50**	**50**
2	Art	0.287	Artistic education Art Art history Visual arts (bach.) Visual arts (license) Visual arts	84	26	19	9 7	10	5 17	2 16	4 13
3	Biology	**0.494**	Biology	**126**	**70**	**88**	**64**	**38**	**64**	**51**	**48**
4	Biology (SGO)	0.260	Biology	66	19	37	21	19	13	19	25
5	Business	**0.428**	Business	**111**	**53**	**85**	**59**	**54**	**54**	**67**	**75**
6	Cartographic eng.	0.126	Cartographic eng.	12	2	4	3	2	8	5	14
7	Chemical eng.	**0.465**	Chemical eng.	**43**	45	**42**	**34**	27	**56**	**42**	**55**
8	Chemistry	**0.352**	Chemistry	29	11	**24**	**16**	**24**	**13**	17	9
9	Computer science	**0.325**	Information science Computer science	**111**	40	**57**	**34**	**43**	**36**	38 	 **35**
10	Dentistry	**0.404**	Dentistry	**57**	15	**25**	17	17	**23**	**28**	**30**
11	Economics	0.286	Economics	77	34	40	24	32	30	26	31
12	General eng.	**0.307**	Civil eng. Engineering Electrical eng. Textile eng.	58 113	26 54	**35** **87**	**30** **56**	23 38	**58** **64**	**47** **72**	**74** **72**
13	Geography	**0.468**	Geography	**80**	**56**	**49**	**49**	**25**	**40**	**32**	20
14	Geog. Ed. (SGO)	0.275	Geography	73	55	35	23	17	20	21	26
15	Geology	**0.321**	Geology	**14**	4	12	4	7	**20**	10	**16**
16	History	**0.457**	History	**174**	**81**	**114**	**84**	**54**	**55**	**53**	**55**
17	Hist. Ed. (SGO)	**0.333**	History	**71**	29	**50**	**27**	24	20	9	24
18	Industrial design	**0.456**	Industrial design	**33**	11	**33**	17	**18**	**33**	**33**	**36**
19	Journalism	**0.481**	Social communication Journalism Public relations	 **94** **37**	 **30** **13**	 **54** **28**	 **33** 21	 **33** **23**	 **36** **25**	 **38** **34**	 **44** **35**
20	Language (SGO)	0.205	Language	108	26	57	24	21	15	14	16
21	Language I	0.259	Language I Literature/English Port./German Port./Japanese	 53 8	 17 2 3	 27 1 6	 9 1 2	 8 1	 15 5 4	 19 1 1	 10 2 5

#	Program	Prop. AA	Program name(s)	2004	2005	2006	2007	2008	2009	2010	2011
22	Language II	0.293	Language II								
			Port./France	15	18	7	9	1	5	8	3
			Port./Italian	22	13	13	9	6	4	6	7
			Port./Spanish	79	19	30	24	11	11	20	15
23	Language III	**0.379**	Language III								
			Port./Greek	**7**	2	4	3		1	1	1
			Port./Latin	**11**	4	8	4	1	3	4	2
			Port./Literature	**123**	**37**	**67**	33	27	16	20	**25**
24	Law	**0.460**	Law	**284**	**113**	**217**	**132**	**147**	**173**	**198**	**203**
25	Math	0.158	Math	133	41	47	28	23	18	16	9
26	Math Ed. (SGO)	0.143	Math	35	13	18	11	10	3	8	7
27	Mech. eng.	**0.353**	Mech. eng.	**31**	9	16	**17**	**24**	**27**	**38**	**31**
28	Mech. eng. (NF)	0.186	Mech. eng.	16	6	16	9	11	21	18	43
29	Medicine	**0.454**	Medicine	**135**	**61**	**73**	**83**	**65**	**106**	**145**	**189**
30	Nursing	**0.431**	Nursing	**119**	**44**	**74**	**25**	**46**	**52**	**37**	**37**
31	Nutrition	**0.411**	Nutrition	**93**	33	**80**	**50**	**35**	**41**	**29**	**44**
32	Oceanography	0.229	Oceanography	13	19	17	8	6	8	9	10
33	Teaching	0.234	Teaching	284	91	137	74	32	37	40	41
34	Teaching (DDC)	0.208	Teaching				16	9	7	12	9
			Teaching I	59	19	32					
			Teaching II	27	5	11					
35	Philosophy	0.222	Philosophy	70	14	30	18	13	11	15	11
36	Physical ed.	0.206	Physical education	105	27	51	24	26	19	19	25
37	Physics	0.135	Physics	70	25	32	17	13	17	13	15
38	Prod. eng.	**0.384**	Prod. eng.	21	18	13	14	**22**	**39**	21	**39**
39	Prod. eng. (RES)	0.152	Prod. eng.	31	8	17	12	8	17	20	21
40	Psychology	**0.480**	Psychology	**181**	**76**	**85**	**56**	**47**	**62**	**63**	**77**
41	Social science	**0.408**	Social sciences	**75**	36	**49**	**43**	17	20	28	24
42	Social work	**0.432**	Social work	**127**	**56**	**81**	**54**	27	**34**	28	**38**
43	Statistics	0.087	Statistics	31	5	16	1	9	4	3	5

Notes: This table displays the number of applicants in the public high school track for each program/cohort in our data. See Table B4 for details on the table structure and statistics.

TABLE B4. Number of applicants by cohort — Black track

#	Program	Prop. AA	Program name(s)	2004	2005	2006	2007	2008	2009	2010	2011
1	Accounting	**0.364**	Accounting	**47**	17	25	17	13	23	**33**	**42**
2	Art	0.287	Artistic education								
			Art	53	12	10					
			Art history				3			6	7
			Visual arts (bach.)				3				
			Visual arts (license)				2				
			Visual arts					6	8	14	6
3	Biology	**0.494**	Biology	**64**	**28**	**28**	**32**	14	**39**	**33**	**44**
4	Biology (SGO)	0.260	Biology	28	9	13	8	4	7	7	6
5	Business	**0.428**	Business	**59**	38	**42**	17	25	**54**	**42**	**47**
6	Cartographic eng.	0.126	Cartographic eng.	4	1	2	2		2	9	5
7	Chemical eng.	**0.465**	Chemical eng.	26	13	18	16	16	**34**	**43**	**47**
8	Chemistry	**0.352**	Chemistry	**19**	6	10	8	2	14	8	15
9	Computer science	**0.325**	Information science	63	19	**30**	20	17	14	21	
			Computer science								17
10	Dentistry	**0.404**	Dentistry	**42**	5	**18**	12	16	14	**24**	**24**
11	Economics	0.286	Economics	36	14	20	5	6	25	30	22
12	General eng.	**0.307**	Civil eng.	21	7	12	11	9	25	**31**	**70**
			Engineering								
			Electrical eng.	**61**	22	30	25	13	32	40	**62**
			Textile eng.								
13	Geography	**0.468**	Geography	**49**	23	18	14	19	21	**33**	22
14	Geog. Ed. (SGO)	0.275	Geography	52	21	18	12	11	7	15	11
15	Geology	**0.321**	Geology	7	4	4	2	6	**13**	12	**13**
16	History	**0.457**	History	**124**	**37**	**39**	**37**	24	**50**	**36**	**31**
17	Hist. Ed. (SGO)	**0.333**	History	**44**	10	19	11	11	16	11	9
18	Industrial design	**0.456**	Industrial design	**25**	6	**14**	7	8	**21**	8	**18**
19	Journalism	**0.481**	Social communication								
			Journalism	**67**	**13**	**25**	19	**23**	**29**	**31**	**31**
			Public relations	**36**	**22**	**22**	11	8	**25**	**30**	**32**
20	Language (SGO)	0.205	Language	47	15	16	6	3	10	4	9
21	Language I	0.259	Language I								
			Literature/English	31	8	6	4	5	11	7	12
			Port./German	4	1		2	1	1		
			Port./Japanese	6		1	1		2	2	

TABLE B4. Number of applicants by cohort — Black track *(continued)*

#	Program	Prop. AA	Program name(s)	2004	2005	2006	2007	2008	2009	2010	2011
22	Language II	0.293	Language II								
			Port./France	10	3	2	4	4	2	5	3
			Port./Italian	7	12	3	1	2	2	2	4
			Port./Spanish	36	9	18	17	9	6	7	8
23	Language III	**0.379**	Language III								
			Port./Greek	7		3	1		1	2	1
			Port./Latin	6	4	2	5		3	1	
			Port./Literature	**71**	18	**24**	12	11	13	13	19
24	Law	**0.460**	Law	**271**	89	**138**	**122**	**125**	**174**	**192**	**247**
25	Math	0.158	Math	54	16	9	15	10	5	7	6
26	Math Ed. (SGO)	0.143	Math	16	2	1	1	1	1	3	3
27	Mech. eng.	**0.353**	Mech. eng.	**28**	7	10	9	11	**22**	**18**	**27**
28	Mech. eng. (NF)	0.186	Mech. eng.	3	1	2	2		6	7	8
29	Medicine	**0.454**	Medicine	**123**	**38**	**53**	**58**	**48**	**114**	**120**	**187**
30	Nursing	**0.431**	Nursing	**109**	17	**35**	19	20	**32**	**23**	**26**
31	Nutrition	**0.411**	Nutrition	**60**	**24**	**25**	10	16	22	19	28
32	Oceanography	0.229	Oceanography	11	6	9	2	4	2	6	2
33	Teaching	0.234	Teaching	157	40	57	25	16	27	23	34
34	Teaching (DDC)	0.208	Teaching				12	12	15	8	11
			Teaching I	55	13	17					
			Teaching II	17	8	5					
35	Philosophy	0.222	Philosophy	52	11	9	4	2	14	6	4
36	Physical ed.	0.206	Physical education	56	6	24	9	5	6	11	12
37	Physics	0.135	Physics	37	7	10	8	4	9	8	4
38	Prod. eng.	**0.384**	Prod. eng.	9	6	5	3	12	**23**	10	**36**
39	Prod. eng. (RES)	0.152	Prod. eng.	14	5	8	2	1	5	7	7
40	Psychology	**0.480**	Psychology	**123**	**32**	**37**	**28**	28	**43**	**42**	**53**
41	Social science	**0.408**	Social sciences	**56**	19	23	22	22	15	**34**	29
42	Social work	**0.432**	Social work	**119**	**36**	**48**	**27**	30	**41**	22	**52**
43	Statistics	0.087	Statistics	21	4	8		1	2	3	6

Notes: This table displays the number of applicants in the Black track for each program/cohort in our data. The first column shows the 43 programs in our RD and DD samples. The second column shows the proportion of 2004–2011 enrollees in each who were from any affirmative action track (*y*-axis of Figure 1); **bold** numbers in this column show programs with ExposureToAA$_m$ = 1 in our benchmark DD specification (3). The third column shows the subgroups that comprise each program. Remaining columns show the number of applicants to each program/cohort; **bold** numbers denote program/cohorts that we include in our RD sample.

B.5. **Fuzzy merge of UERJ and higher education census data.** In Section 3, we examine the effects of UERJ enrollment on college selectivity using data from a census of all Brazilian college enrollees, called the *Censo da Educação Superior* (INEP, 2019). This subsection describes the merge between UERJ applicants and the higher education census.

We focus on universities in the state of Rio de Janeiro since most Brazilian college students enroll in a university in their home state. We include only 2009–2011 UERJ applicants in this analysis because the higher education census does not exist at the individual level prior to 2009.

We do not observe individuals' ID numbers in the higher education census, so we link the census to the UERJ records using a fuzzy merge based on exact day of birth, gender, and year of enrollment.[36] In the census data, we compute the *total* number of students at a particular university with a given birthdate, gender, and enrollment year. We merge these variables into our UERJ sample using birthdate, gender, and year of *application.* We then use these totals as dependent variables in our RD specification.

The resulting dependent variables reflect the total number of enrollees in a particular university in Rio de Janeiro who have the same birthdate/gender/enrollment-year triplet as a UERJ applicant. The ideal dependent variable—if we could uniquely identify individuals in the census—would be an indicator variable that takes the value one if a given UERJ applicant enrolled in a given university and zero otherwise. If no college student at a university has the same birthdate/gender/enrollment-year triplet as the applicant, we know that the applicant did not enroll in that university in that year (barring errors in the merge variables). However, if one or more enrollees at the university share the same combination of those three variables, we cannot tell with certainty whether the applicant ended up enrolling in the university.[37] Thus our dependent variables contain additional measurement error.

B.6. **Decomposition of DD estimate for top enrollees' log hourly wages.** Our main result in Section 5 is that top enrollees' hourly wages declined by 0.132 log points in UERJ majors with high exposure to affirmative action relative to less-exposed majors (Table 6, Panel B, column B). This section describes back-of-the-envelope calculations on the proportion of this estimate that can be explained by compositional, networking, and learning mechanisms.

First, we find that 25 percent of our main DD estimate can be explained by changes in the observable characteristics of top enrollees. In Panel D of Table 5 we combine applicants'

---

[36] Each of these variables is available in the public version of the *Censo da Educação Superior* that we downloaded from the website of a Brazilian Ministry of Education agency called INEP (*Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira*). Some of these variables are no longer available in the current version of this dataset on INEP's website.

[37] For the average UERJ applicant in our data, there are 29 students with the same birthdate, gender, and enrollment year across all Rio de Janeiro universities.

observable characteristics—both demographics and admission scores—into a log wage index, and use this as an outcome variable in our DD specification (3). For individuals who appear in the RAIS data, the observable characteristics of top enrollees declined by 0.033 log wage points in majors with greater exposure to affirmative action relative to less-exposed majors (last row in Panel D of Table 5, column B). Although this point estimate is not statistically significant, it is 25 percent of the magnitude of our main DD estimate for top enrollees' log hourly wages, i.e., $-0.033/(-0.132) = 25\%$.

Next, we find that 10–17 percent of our DD coefficient for top enrollees' hourly wages can be explained by networking mechanisms. We estimate the contribution of networking mechanisms by multiplying the DD estimates for employment in alumni firms (Panels C–D of Table 6) by the OLS wage premia associated with employment in these firms (Appendix Table A12). In Panel C of Table 6, we find that greater exposure to affirmative action reduced top enrollees' likelihood of employment at firms that hired pre-AA alumni by 5.5 percentage points, and it increased their likelihood of employment at firms that hired *only* post-AA alumni by 4.9 percentage points. Appendix Table A12 shows that firms with pre-AA alumni had an hourly wage premium of 0.481 log points relative to other firms in our sample, whereas firms with only post-AA alumni had an hourly wage premium of 0.274 log points. Under the assumption that these OLS wage premia reflect causal effects, the change in access to pre-AA and post-AA alumni firms can explain 10 percent of our main DD estimate for hourly wages, i.e., $(-0.055*0.481+0.049*0.274)/(-0.132) \approx 10\%$. If we do a similar calculation using the estimates from Panel D of Table 6—which measure employment with alumni from different application tracks and cohorts—we find that these estimates can explain 17 percent of our main DD estimate for hourly wages.[38]

Lastly, we estimate that 32 percent of the negative wage effect for top enrollees can be explained by learning mechanisms. We find that affirmative action reduced UERJ's white private high school students' proportion of correct answers on the Enade exam by 2.2 percentage points (first row in Panel B of Table 7, column C).[39] All else equal, white students

---

[38] Specifically, the DD estimates in Panel D of Table 6 and their associated OLS firm wage premia are:

- General track alumni from the same cohort: DD coef = -0.098; OLS wage premium = 0.533;
- General track alumni from different cohorts: DD coef = 0.042; OLS wage premium = 0.455;
- Only AA alumni from the same cohort: DD coef = 0.036; OLS wage premium = 0.293;
- Only AA alumni from different cohorts: DD coef = 0.010; OLS wage premium = 0.070.

Thus we can explain $(-0.098*0.533+0.042*0.455+0.036*0.293+0.010*0.070)/(-0.132) \approx 17\%$ of our main DD estimate for hourly wages.

[39] We find a slightly larger point estimate (-3.0pp) in a triple-difference specification that also compares UERJ majors with more- and less-exposure to affirmative action (Appendix Table A18).

from private high schools would have had to obtain *higher* entrance exam scores to be admitted to UERJ in the cohorts with affirmative action.[40] Thus we think that the 2.2 percentage point decline in Enade scores is, if anything, likely to underestimate the decline in scores that we would find in our top enrollee sample. We cannot estimate the relationship between hourly wages and the proportion of correct answers on the Enade exam because we do not have these two datasets linked at the individual level. As an alternative benchmark, we use Reyes (2023)'s estimate that a one percentage point increase in the proportion of correct answers on Brazil's national college entrance exam (ENEM) is associated with a 0.0192 log point increase in early-career wages. Under the assumption that the relationship between correct answers and wages is the same for the ENEM and Enade exams, the decline Enade scores for white private high school students can explain 32 percent of our DD estimate for top enrollees' wages, i.e., $(-2.2 * 0.0192)/(-0.132) = 32\%$.

---

[40] However, we find limited evidence that affirmative action changed the demographics characteristics of UERJ's white private high school students who took the Enade exam (Panel A of Table 7).