Can Affirmative Action Affect Major Choice?*

Fernanda Estevan[†] Thomas Gall[‡] Louis-Philippe Morin[§]

May 21, 2019

Abstract

Around the world, students from a disadvantaged background are underrepresented in prestigious and lucrative fields of study, such as medicine and STEM. We know little about whether universities can affect individuals' major choice and promote increased social mobility. In this paper, we provide evidence that universities can change individuals' choice of major. We use a natural experiment that expanded the set of majors to which lower SES applicants could be admitted. We find that this change in policy, which was implemented at a very selective university, increased the likelihood of lower SES students to apply for, and get accepted to more prestigious majors.

Keywords: post-secondary education, affirmative action, major choice, social mobility.

*This research was supported by the Social Sciences and Humanities Research Council of Canada (430-2013-001033), the British Academy and Newton Fund (AF140079), and the Sao Paulo Research Foundation-FAPESP (17/50134-4) and CALDO (2017/148485). We gratefully acknowledge COMVEST, UNICAMP's admission office, for providing the data and assistance during the project. We thank Pierre Brochu and Adam Lavecchia for their helpful comments. We also thank Gustavo Katague, Derek Rice, Duangsuda Sopchokchai, Bogdan Urban, Pedro Feitosa, and Tiago Ferraz for their excellent research assistance. All remaining errors are ours.

†Sao Paulo School of Economics - FGV, Rua Itapeva, 474, Sao Paulo, SP, Brazil; email: fernanda.estevan@fgv.br.

‡Department of Economics, School of Social Sciences, University of Southampton, Southampton, SO17 1BJ, UK; email: t.gall@soton.ac.uk.

§Department of Economics, University of Ottawa, 120 University, Ottawa ON K1N 6N5, Canada; email: lmorin@uottawa.ca.

1 Introduction

There is considerable heterogeneity in the labour market returns to the post-secondary field of study.[1] Recent studies suggest that the field of study is more important than school quality in determining future earnings (e.g., Altonji et al., 2016; Kirkeboen et al., 2016). Moreover, low-income students are underrepresented in lucrative and prestigious majors, such as medicine and STEM.[2] The combination of significant heterogeneity in the returns to majors and the link between parental socioeconomic status (SES) and major choice may explain some of the observed (lack of) intergenerational social mobility.

An important question is whether universities can affect individuals' major choice (Altonji et al., 2016) and promote increased social mobility. The answer certainly depends on whether students' (or perhaps parents') personal preferences or circumstances govern the individual choice of higher education (see, e.g. Arcidiacono, 2004; Beffy et al., 2012; Corak, 2013; Gemici and Wiswall, 2014; Wiswall and Zafar, 2015).

We contribute to the literature by providing direct evidence that universities can change individuals' choice of major. To do so, we exploit the introduction of an affirmative action policy at the university admission stage at Universidade Estadual de Campinas (UNICAMP), a large, highly ranked, and research-intensive public university in the state of Sao Paulo, Brazil. Given the structure of the affirmative action put in place at UNICAMP and its admission rules, the university essentially enlarged the major choice set of applicants targeted by the affirmative action, without affecting applicants' perception of their own ability.

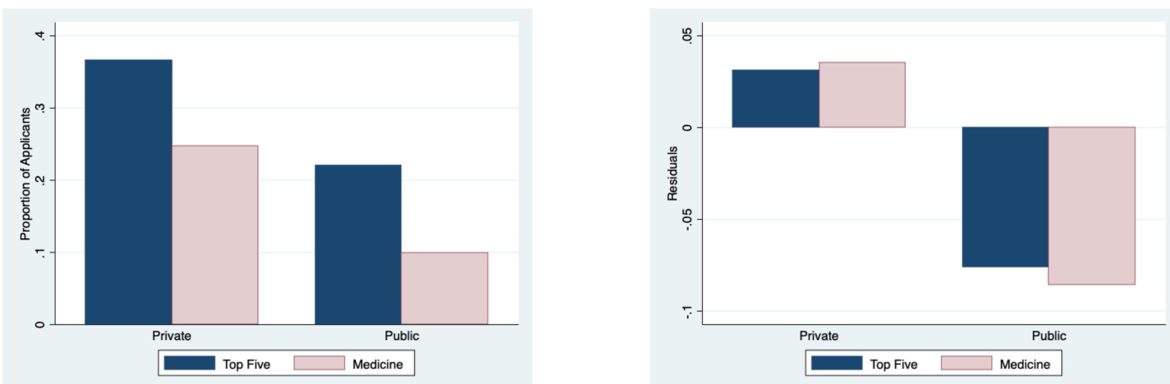
We benefit from an unusually favourable setting to investigate the causal impact of affirmative action on major choice. As in the majority of OCDE countries (Kirkeboen et al., 2016), Brazilian universities have "college-major-specific admission rules" (Bordon and Fu, 2015). Therefore, applicants must choose both the university and major they want to attend before being admitted, instead of applying to a university and then selecting a major during their undergraduate studies, as in the US. The affirmative action policy, first implemented for the 2005 UNICAMP admission, consisted in giving university applicants from public

[1] For example, in the US it is common to find that the returns to engineering are significantly larger than to education (Altonji et al., 2012, 2016). See also Hastings et al. (2013) for Chile and Kirkeboen et al. (2016) for Norway.

[2] This fact has caught the attention of the popular press and academic community around the world. In the UK, 80% of medical students have parents in higher managerial or professional occupations (Carrell, 2016), and the percentage of applicants from lower SES occupational groups ranged between 2.3% and 8.4% depending on the medical school (Steven et al., 2016). In the US, there is a similar discussion on the underrepresentation of low-income students and minorities in STEM fields (Camera, 2017; NSF, 2019). In Brazil, the fraction of public school graduates accepted in competitive majors such as Law and STEM in selective universities, i.e., flagship universities from a Brazilian state, was close to zero at the beginning of the 2000s (Cavalcanti et al., 2010).

high schools bonus points on the admission exam. In Brazil, public high school graduates typically have a lower socioeconomic background and are underrepresented in public universities, arguably the better universities of the country. While public high school students represented 84 per cent of high school graduates in the state of Sao Paulo in 2003 (INEP, 2005), they only made up about 28 per cent of UNICAMP's intake in 2004.[3]

Public high school students are also underrepresented in selective majors when compared to their private high school counterparts. Panel (a) in Figure 1 presents the proportion of UNICAMP's applicants who chose a top-five major (in terms of admission cutoff[4]) or medicine in 2003-2004 based on the type of high school they have attended (i.e., private or public). Private school students are more likely to apply for medicine, and a top-five major by 15 percentage points and these differences only shrink to approximately 11 percentage points when one controls for an end-of-high-school standardised exam, i.e., ENEM scores, in Panel (b). Thus, a significant gap in application behaviour persists even when accounting for applicants' academic credentials.



(a) Unconditional

(b) Conditional on ENEM

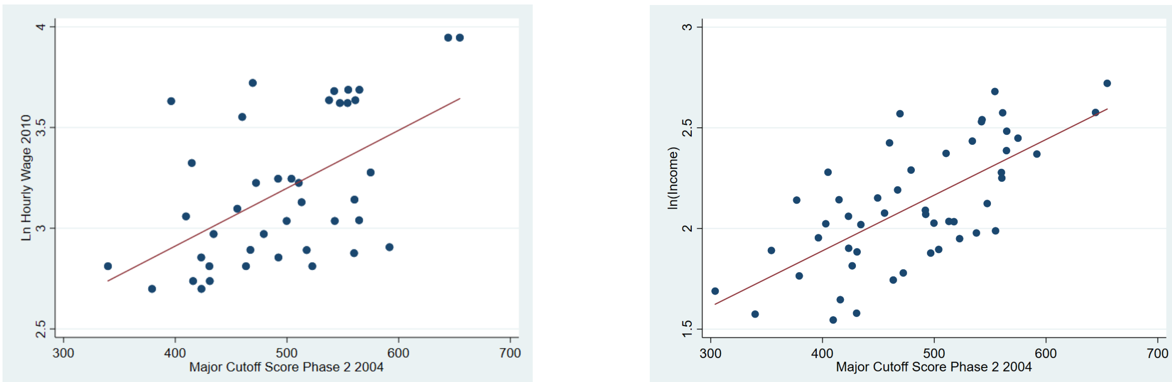
Figure 1: Proportion of Applicants Choosing a Top-Five Field of Study and Medicine

Our identification strategy exploits the quasi-natural experiment generated by the introduction of the affirmative action policy, a comprehensive individual-level dataset, and several features of the Brazilian and UNICAMP university systems. In our dataset, we observe all applicants major choices (up to three options). Applicants choose a field of study when they register to take UNICAMP's entrance exam and receive information regarding the previous year's cutoff scores per field of study (the lowest exam score admitted), which are useful

[3]These figures are similar at the University of Sao Paulo (USP), another selective university in the state of Sao Paulo.

[4]A major admission cutoff is essentially the lowest admission exam score admitted to the major. Selective majors have higher admission cutoffs. The top-five majors are the most selective STEM majors: computer engineering, control, and automation engineering, electrical engineering (day), electrical engineering (evening), and medicine (UNICAMP).

predictors of current year's cutoffs. These cutoff scores vary significantly by major and are substantially higher for lucrative occupations, as shown in Panel (a) of Figure 2 where we plot the average log hourly wage in the state of Sao Paulo (IBGE, 2010) as a function of the UNICAMP cutoff scores associated with the major the individual graduated from.[5] There is a clear positive relationship between the cutoff score of the major an individual graduated from and the average salary: a one standard deviation increase in the cutoff is associated with a 22 per cent increase in wages.



(a) Wages (b) Parental Income

Figure 2: Correlation between Admission Exam Cutoffs, Wages and Parental Income

There is also a strong correlation between socioeconomic background measured by parental income and the cutoff score of a major a student applies to, as shown in Panel (b) of Figure 2. Individuals selecting majors with relatively larger cutoff scores have, on average, higher total family income. At first sight, this association could be simply due to an intergenerational correlation of occupational preferences, as documented in van de Werfhorst and Luijkx (2010). However, the correlation in real application choices might be due to factual constraints. As most Brazilian universities at that time, UNICAMP uses a version of the well-known Boston mechanism to allocate applicants to university places (Carvalho et al., 2014). This allocation mechanism gives strong incentives for applicants to behave strategically by naming as their first choice the major that they prefer *among the majors they are likely to be accepted in*. Hence, by adding a bonus to an applicant's entrance exam score, the set of attainable majors expands, which could affect application choice, unless the preferred major in the new choice set has already been in the pre-policy set.

We illustrate this mechanism with a simple educational choice model in Section 3. In our model, students are characterised by an innate ability level, choose a high school when young and the amount of effort they provide, leading to an end-of-high-school grade, i.e., ENEM

[5]We benefit from the fact that the Brazilian Population Census inquires individuals on the major they completed.

score. We show that the ENEM score and high school attended are sufficient statistics for the expected entrance score. The model also predicts that the bonus policy may lead beneficiaries to apply for more competitive majors if their preferred major was not in their choice set initially. In contrast, private high school students may react to the (general equilibrium) change in cutoff scores by selecting majors that require a lower admission score.

We find that the expansion of individuals' choice sets caused by the affirmative action policy affected students' choice of major. Applicants who benefited from the affirmative action ended up choosing more competitive, more prestigious, and more lucrative fields of study following the introduction of the policy. Moreover, we also show that the policy increased the likelihood of being admitted to more selective majors. Thus, our results strongly suggest that individual circumstances are a relevant factor in major choice, potentially affecting the occupational choice and social mobility. Additionally, we find that the affirmative action policy has reduced the parental income gradient associated with some selective major decisions.

We contribute to a growing literature on how socioeconomic background affects major choice and whether public policy could alter its impact.[6] Montmarquette et al. (2002) and Boudarbat and Montmarquette (2009) study the impact of expected earnings on major choice, and suggest that it may vary for distinct socioeconomic groups. In particular, Boudarbat and Montmarquette (2009) claim that altering the choice through the expected earnings channel would require substantial increments in future earnings. A now common policy targeted at increasing the presence of students from disadvantaged socioeconomic backgrounds in post-secondary institutions is affirmative action, which could, in principle, alter major choice (Arcidiacono et al., 2015). Arcidiacono (2005) estimates a structural model of the college decision-making process, which includes the choice of field of study. He shows that removing race-based affirmative action has little impact on future black earnings as it impacts only the quality of college individuals attend and returns to college quality are estimated to be quite low. In his setting, changes in admission policy mainly affect the choice of institution, but not the choice of major.

Our paper is close to Papay et al. (2016), Avery et al. (2018) and Bond et al. (2018). In these papers, the authors investigate the effect of positive exam performance/signal shocks (Massachusetts Comprehensive Assessment System (MCAS), SAT and Advanced Placement

[6]There is growing literature analysing other factors determining the choice of field of study. Individuals sort themselves based on comparative advantage (Kirkeboen et al., 2016), expected earnings (Arcidiacono et al., 2012; Wiswall and Zafar, 2015), nonpecuniary factors (Beffy et al., 2012), and their (mis)perceptions about their ability to do well in a given program (Zafar, 2011; Stinebrickner and Stinebrickner, 2014). There are also some recent studies investigating whether peers (Anelli and Peri, 2015), role models (DellaVigna, 2010; Ferrando and Gille, 2016) or exposure to major (Fricke et al., 2018) can alter the choice of major.

test scores, respectively) on the college application and major completion. Papay et al. (2016) find that conditional on the same MCAS mathematics score, students attending urban schools and from low family income are more likely to attend college if they receive a more positive test score 'label'. They find no effects for the English score label and suburban or higher-income students. Bond et al. (2018) find that positive SAT shocks affect college-application portfolios (e.g., students apply to more selective colleges), but the magnitude of the effect is small. Avery et al. (2018) find that a higher Advanced Placement score (for similar raw scores) increases the likelihood to major in that exam subject (but no significant effect on attended college quality).

There is an essential difference in the channels through which score/signal shocks and our bonus-point policy can operate. In particular, Bond et al. (2018) argue that the score shocks can affect college choice by affecting: 1) the perceived admission probability and 2) the beliefs about the likelihood to do well at more selective colleges. Papay et al. (2016) emphasise the role of point 2) in explaining their findings given their labels "do not carry official consequences." Avery et al. (2018) also make a similar argument (in particular point 2), above) for major choice. While this may be true for score shocks, UNICAMP's bonus points do not bring new information regarding the applicant's ability, and we, therefore, do not expect the affirmative action to affect applicant's beliefs regarding her likelihood to do well in a given program. Still, we document significant effects from the policy on 1). An interesting finding in Bond et al. (2018) is that higher-ability students react (update) more to SAT shocks than lower-ability students. Similarly, we show that individuals in the top quartiles of ENEM respond more to the affirmative action policy.

We organize the paper as follows. In Section 2, we describe UNICAMP's admission system and its affirmative action policy. We provide some theoretical pointers to explain the potential impact of the affirmative action policy on major choice in Section 3. In Sections 4 and 5, we present our data and our identification strategy. We show the main results, explore some heterogeneous effects, and present robustness checks in Sections 6, 7, and 8, respectively. We explore the effect of the affirmative action on the parental-income gradient in Section 9. We conclude in Section 10.

2 UNICAMP's Admission and Affirmative Action Policy

To select students, UNICAMP organises an annual admission exam. Every year, around 50,000 applicants register in September to write the admission exam. UNICAMP's admission exam consists of two parts, Phase 1 (henceforth P_1), which eliminates about 70 per cent of applicants, and then Phase 2 (henceforth P_2), which is conditional on passing P_1. Both P_1 and P_2 are based on high school subjects (e.g., chemistry, mathematics and Portuguese).

Upon registration, applicants must choose up to three majors, which they rank first, second, and third. The choice of major is crucial for many reasons. First, students are admitted in a specific major, not just the university. Once accepted, it is difficult for a student to change major, and it typically involves retaking the admission exam. Second, the applicant's major choice determines the pool of competitors and, hence, the minimum P_1 score necessary for advancing to P_2 and the P_2 cutoff scores for being admitted. Since UNICAMP is a prestigious free-of-charge university, most majors are quite competitive. Each year, about 10 per cent of applicants are accepted. Still, there are stark differences in acceptance rates (and cutoff scores) across majors (see Figure 2). Finally, UNICAMP's admission process uses a version of the well-known Boston mechanism (Abdulkadiroglu and Sonmez, 2003) to assign students to majors. Under this assignment mechanism, students have an incentive to apply to majors they are likely to get in. For example, an individual who prefers to study medicine may opt for nursing, which requires a lower score on the admission exam, to ensure that she can pursue a university education.

Applicants who choose a given major and fulfil minimal requirements in each subject are ranked based on their overall admission exam score (NPO, for *nota padronizada de opção* in Portuguese), which is a weighted average of P_1 (possibly including ENEM score) and P_2 scores.[7] Importantly, the NPO ranking initially considers only applicants who choose a given major as their first choice.[8] Over the 2001-2008 period, two-thirds of majors were always filled after this first round of offers, meaning that only applicants who opted for these majors as the first choice were ever considered for admission in these majors. Only two out of 60 majors always had some slots available after the first round in every admission process over this period. In majors/years where seats are available after the first round, applicants who choose the major as second or third choices and fulfil the minimum grade criteria receive admission offers based on their NPO ranking. After that, if there are still seats available, then grade requirements are reduced, but applicants who rank the major as their first choice continue to have priority for admission. Indeed, we find that between

[7]There are eight high school subjects covered in P_2. An applicant is automatically eliminated if she receives 0 on any given subject. There are also minimum grade requirements for priority subjects, which vary depending on the major chosen. *Exame Nacional do Ensino Médio* (ENEM) is an end-of-high-school exam that can count for up to 20 per cent of P_1 score.

[8]There are four groups of majors for which first and second (but not third) options are considered simultaneously: Electric Engineering (daytime and evening), Chemical Engineering (daytime and evening), Medicine Unicamp and Medicine Faculdade de Medicina de Sao José do Rio Preto (FAMERP), and Nursing Unicamp and Nursing FAMERP.

2001 and 2008 nearly 90 per cent of admitted applicants obtained their first choice.[9] Thus, selecting a very competitive major as the first choice, such as medicine, may prevent an applicant from being admitted in nursing even if her NPO score was way above the cutoff score for nursing.

In 2004, UNICAMP implemented an affirmative action program granting a 30-point bonus (on their NPO) to applicants who did their entire high school in a public school (and passed P_1). The bonus was sizeable, corresponding to 30% of a standard deviation. Applicants who completed their entire high school in a public school and declared themselves visible minority were granted an additional 10 points (for a 40-point bonus). Note that visible minorities from private high schools were not eligible for any bonus point. Although we would expect visible minorities from public high schools to react the most from UNICAMP's affirmative action, we do not focus on the potential race dimension of the policy for two reasons. First, from a practical point of view, visible minority applicants represent a small proportion of applicants (about 14 per cent of our 2004 sample and 9 per cent of our 2004 private high-school applicants), which makes a precise estimation of the race-related parameters of interest a challenge. Second, and importantly, the race is self-reported, and we are worried that some students who would not have previously self-reported as a visible minority could have done so following the introduction of the policy, given the incentives. If so, the race would be an endogenous and therefore a lousy control variable.[10]

While the affirmative action program was announced a few months before UNICAMP registration for the 2005 admission exam, most applicants learned about the program while registering for 2005 UNICAMP exam (in September 2004). Since they had to choose their major at the same moment, it is unlikely that the affirmative action program affected their choice of major already in 2005. Therefore, we focus mainly on the impact of the affirmative action policy on major choice for the 2004 and 2006 admission exams and use additional data from 2001 to 2003 and 2007 for robustness checks.

3 An Education Choice Model

Before specifying our empirical exercise, we now illustrate the potential impact that a policy such as the one just described could have on applicants' major choice.

[9]Estevan et al. (2019) explain the UNICAMP admission process in more details.

[10]Nevertheless, we have estimated regressions where we interact race with our regressors of interest (with our public high school dummy, with our affirmative action dummy, and with the interaction of these two dummy variables) as robustness checks. Including these variables does not affect our estimates of interest and the race-related parameters are, for the vast majority, small and statistically insignificant. These results are available upon request.

The model follows the spirit of the literature on modelling major choice based on updating expectations (e.g., Arcidiacono et al., 2012; Stinebrickner and Stinebrickner, 2014; Wiswall and Zafar, 2015; Bond et al., 2018), but taking into account that the affirmative action policy generates a deterministic point bonus on the exam score, instead of a noisy signal on ability.

The economy is populated by a continuum I of students. Individuals are characterised by their innate academic ability θ_i and have disutility from exerting effort e_i at a strictly convex utility cost $c_i(e_i)$. When young, they choose a high school s_i from a discrete set S (we abstract from capacity constraints at high school) and an effort level e_i. Ability, effort and school determine individual academic human capital at the end of high school, given by function $h(\theta_i, e_i, s_i)$, which is strictly increasing in θ_i and e_i. Academic human capital captures an individual's skill at sitting exams. Suppose that an individual's final high school grade g_i is the realisation of a random variable $\tilde{g}_i$ whose mean is determined by the academic human capital:

$$\tilde{g}_i = h(\theta_i, e_i, s_i) + \epsilon, \tag{1}$$

where ϵ is a random variable with mean 0, distributed according to a distribution function F_ϵ. This assumption also implies that realised high school final grades g_i are unbiased estimators of individual academic human capital $h(\theta_i, e_i, s_i)$.

Suppose there is a discrete set M of university majors that a student can choose from. Denote the major choice of student i by $m_i \in M$.[11] A student i derives utility $u_i(m)$ from enrolling in a major $m \in M$. Different students may have different preferences of majors, so that $u_i(m) > u_i(m')$ does not imply that $u_j(m) > u_j(m')$ for $i \neq j$ and $m \neq m'$, but suppose that each individual's ranking is strict, i.e. $u_i(m) = u_i(m')$ if and only if $m = m'$. Given θ_i, s_i, e_i and the realised grades g_i and g_j for all other $j \neq i \in I$, an individual i chooses a major $m_i \in M$ to apply to. Note that in this formulation individual preferences over majors do not depend on effort e_i or high school s_i.

An exam governs university admission. Again the exam result $\tilde{t}_i$ is a random variable whose mean depends on academic human capital $h(.)$, as well as the realised high school grade g_i (both because our studied admission exam uses ENEM as a potential input, and because g_i may carry some information on a student's ability to sit exams beyond their academic human capital):

$$\tilde{t}_i = \alpha h(\theta_i, e_i, s_i) + (1 - \alpha) g_i + \nu, \tag{2}$$

[11] The model assumes only one major is chosen. This is motivated by the allocation mechanism used by UNICAMP, in which the first choice matters disproportionately (see Section 2). For models allowing for choices of a portfolio of educational options, see e.g. Epple et al. (2006), Chade et al. (2014) or Fu (2014).

where $\alpha \in (0,1)$ and ν is a random variable with mean 0, distributed according to a distribution function F_ν.[12] Note that, because of (1), high school final grades g_i are also an unbiased predictor of entrance exam scores $\tilde{t}_i$.

Each major m has a capacity $k(m)$, yielding an admission cutoff $\bar{t}(m)$ such that the measure of individuals who choose major m equals $k(m)$. We assume that all majors are oversubscribed (which will be the case if the mass of applicants is large enough and every major is preferred to the outside option for a sufficient measure of students).

To summarise, the timeline is:

- individuals born with innate ability θ_i,
- individuals choose school s_i, and then effort level e_i at school s_i, which yields academic human capital $h(\theta_i, e_i, s_i)$,
- the high-school final exam (ENEM) yields grade $\tilde{g}_i$, a random variable with mean $h(\theta_i, e_i, s_i)$,
- individuals choose a major m_i and then take the university admission exam, yielding test score $\tilde{t}_i$, a random variable with mean $\alpha h(\theta_i, e_i, s_i) + (1-\alpha) g_i$.
- Given test scores $(t_i)_{i \in I}$ that yield major specific admission cutoffs $\bar{t}(m)$, an applicant i is admitted to the major applied for m_i if $t_i + AA_i \geq \bar{t}(m_i)$, where AA_i denotes the score bonus through affirmative action available to individual i.

Turn now to individual decisions. When choosing which major to apply for, an individual solves

$$\max_{m_i \in M} Prob(t_i \geq \bar{t}(m_i) - AA_i) u_i(m_i),$$

where $Prob(.)$ denotes the probability of being accepted into major m_i, and AA_i denotes the policy (i.e. $AA_i > 0$ if individual i is entitled to a bonus and $AA_i = 0$ otherwise). Given that there is a continuum of applicants and a discrete set of majors, and invoking a law of large numbers, for each $m \in M$ it must hold that $\bar{t}(m) = \overline{h}(m)$, where

$$\overline{h}(m) : \int I(\alpha h(\theta_j, e_j, s_j) + (1-\alpha) g_j \geq \overline{h}(m)) \cdot I(m_j = m) dj = k(m),$$

[12] After applying for a major, individuals have some time to prepare for the entrance exam. We choose to ignore the problem of choosing the optimal effort in exam preparation since students do not seem to have changed their effort provision meaningfully in response to the policy (Estevan et al., 2019). We conjecture, however, that allowing for the second round of effort choice would not change our analysis qualitatively, but could affect the quantitative outcome, in particular, the measure of applicants who effectively gain access to a major they would not have had access to without the policy.

where $I(.)$ denotes an indicator function, $k(m)$ was the capacity of major m and the integral is with respect to the Lebesgue measure. The probability of being accepted into the major m is then just

$$Prob(t_i \geq \bar{t}(m) - AA_i) = Prob(\nu \geq \overline{h}(m) - AA_i - \alpha(h(\theta_i, e_i, s_i) + (1-\alpha)g_i).$$

Therefore an individual will prefer major m over another major m' if

$$\begin{aligned} &Prob(\nu \geq \overline{h}(m) - AA_i - \alpha h(\theta_i, e_i, s_i) - (1-\alpha)g_i)u_i(m) \\ &> Prob(\nu \geq \overline{h}(m') - AA_i - \alpha h(\theta_i, e_i, s_i) - (1-\alpha)g_i)u_i(m'). \end{aligned}$$

This immediately implies two useful revealed preference properties.

Fact 1. *Given θ_i, e_i, s_i and $(t_j)_{j\in I}$, if an individual i chooses m then $u_i(m) > u_i(m')$ for all m' with $\bar{t}(m') < \bar{t}(m)$.*

That is, any major with a lower equilibrium cutoff than the one that is actually chosen will yield lower utility than the one chosen. In a similar vein, applicants who are eligible for the AA policy will choose the same major or a major with higher equilibrium cutoff, which they prefer, if the AA policy is active (i.e., the year is 2006 in our case).

Fact 2. *Given θ_i, e_i, s_i and $(t_j)_{j\in I}$, if an individual i chooses m when $AA_i = 0$ and m' when $AA_i > 0$, then $u_i(m') > u_i(m)$ and $\bar{t}(m) < \bar{t}(m')$.*

Moreover, since high school grades are sufficient statistics for admission exam scores, the expected difference $E[g_i - t_i] = E[\alpha\epsilon\nu + (1-\alpha)]0 = 0$, which implies that the high school attended should not have an independent effect on major choice.

Fact 3. *Under (1) and (2), $E[g_i - t_i] = 0$. If $E[g_i - t_i] = 0$ then $m_i = m$ and $m_j = m'$ for $h(\theta_i, e_i, s_i) = h(\theta_j, e_j, s_j)$ imply that $u_i(m) > u_i(m')$ and $u_j(m) < u_j(m')$.*

That is, under our assumption any school effects on the admission exam performance should be present in the high school exit grade, and different major choices for individuals with the same academic human capital (and thus similar high school exit grades) must be due to different preferences.

Discussion

Effect Sizes

To gain some idea about which applicants would be affected the most by a bonus policy, recall that an individual receiving a bonus AA_i will switch from major m to another major

m' (with $u_i(m) > u_i(m')$ and $\overline{h}(m) < \overline{h}(m')$) if both

$$\frac{Prob(\nu \geq \overline{h}(m) - \alpha h(\theta_i, e_i, s_i) - (1-\alpha)g_i)}{Prob(\nu \geq \overline{h}(m') - \alpha h(\theta_i, e_i, s_i) - (1-\alpha)g_i)} > \frac{u_i(m')}{u_i(m)},$$

and

$$\frac{u_i(m')}{u_i(m)} > \frac{Prob(\nu \geq \overline{h}(m) - AA_i - \alpha h(\theta_i, e_i, s_i) - (1-\alpha)g_i)}{Prob(\nu \geq \overline{h}(m') - AA_i - \alpha h(\theta_i, e_i, s_i) - (1-\alpha)g_i)}.$$

Denoting by $x_i = E[\tilde{t}_i]$ the predicted admission exam grade, the condition implies that

$$\frac{Prob(\nu \geq \overline{h}(m) - x_i)}{Prob(\nu \geq \overline{h}(m') - x_i)} - \frac{Prob(\nu \geq \overline{h}(m) - x_i - \Delta)}{Prob(\nu \geq \overline{h}(m') - x_i - \Delta)} > 0. \tag{3}$$

Hence, the larger the difference on the LHS of condition (3) the more likely it is that an applicant i will switch major choice for a bonus of Δ, because the set of preferences that would induce switching increases). Since $Prob(\nu \geq \overline{h}(m) - x_i) = 1 - F_\nu(\overline{h}(m) - x_i)$, the probability of switching major choice after receiving a bonus Δ is (weakly) monotone in

$$\frac{1 - F_\nu(\bar{t}(m) - x_i)}{1 - F_\nu(\bar{t}(m') - x_i)} - \frac{1 - F_\nu(\bar{t}(m) - x_i - \Delta)}{1 - F(\bar{t}(m') - x_i - \Delta)}.$$

For small Δ this difference is well approximated by the differential:

$$\begin{aligned}
\frac{\partial \frac{1-F_\nu(\overline{h}(m)-x_i)}{1-F_\nu(\overline{h}(m')-x_i)}}{\partial x_i} &= \frac{(1 - F_\nu(\overline{h}(m) - x_i))f_\nu(\overline{h}(m') - x_i) - (1 - F_\nu(\overline{h}(m') - x_i))f_\nu(\overline{h}(m) - x_i)}{(1 - F_\nu(\overline{h}(m') - x_i))^2} \\
&= \left(\frac{f_\nu(\overline{h}(m') - x_i)}{1 - F_\nu(\overline{h}(m') - x_i)} - \frac{f_\nu(\overline{h}(m) - x_i)}{1 - F_\nu(\overline{h}(m) - x_i)}\right)\frac{1 - F_\nu(\overline{h}(m) - x_i)}{1 - F_\nu(\overline{h}(m') - x_i)}.
\end{aligned}$$

This expression is guaranteed to be greater (less) than zero if F_ν has an increasing (decreasing) hazard rate on $[\overline{h}(m) - x_i, \overline{h}(m') - x_i]$. The normal distribution has the property that its hazard rate is strictly increasing and convex, which implies the following statement.

Fact 4. *If the error term ν follows a normal distribution, the LHS of Condition (3) increases in x_i, and thus the necessary condition for switching majors under the policy becomes slacker.*

General Equilibrium Effects: Cutoffs

There is a possible general equilibrium effect as the policy will tend to increase entrance exam scores by awarding the bonus to some applicants, thus increasing the cutoffs. This may also affect private school students' behaviour in equilibrium. Since under affirmative action, students' choices will change toward majors with higher cutoffs, if at all, cutoffs cannot decrease, and some must increase if some individuals change their major choice. Facing higher cutoffs, some students who do not receive a bonus may now find a major with lower cutoff more attractive since the probability of being accepted in their original choice decreases. This points to a differential effect of the affirmative action policy on public and private school students: when cutoffs generally increase, private school students react to the higher cutoffs by aiming at lower cutoff majors, whereas public school applicants choose higher cutoff majors, because the bonus overcompensates the increase in cutoffs and increases the acceptance probability in the higher cutoff major. That is, private school applicants may move down the ladder, but this effect is entirely due to the change in equilibrium cutoff values, whereas public school applicants move up the ladder and choose higher cutoff majors, even at the new, higher levels.

Theoretical Identification: Summary

To summarise, the reasoning derived above yields allows us to state some predictions on the effects of introducing the policy on individual major choice. Our arguments relied in essence on the following set of assumptions:

Assumption 1. *(1) $\overline{h}(m) = \overline{t}(m)$ is observable,*

(2) the relation of $u_i(m)$ to observable individual characteristics (such SES, origin, school, academic capital captured by ENEM) is stationary and

(3) the high school final grade g_i is a sufficient statistic for $h(\theta_i, e_i, s_i)$.

These assumptions imply the following set of predictions:

Prediction 1. *Conditional on ENEM grades and individual characteristics, the difference in individuals' major choices over time*

(1) will be zero for applicants in the absence of an AA bonus policy,

(2) will on average increase the cutoff of major choices for applicants who are eligible for a bonus when the policy is active, and

(3) will be zero for applicants who are not eligible for a bonus when the policy is active, conditional on the equilibrium cutoffs for major admission.

That is, recipients of the bonus are expected to apply to majors with higher entry thresholds, i.e. more prestigious and rewarding majors when controlling for individual covariates.

Applicants who are not eligible for the bonus may apply for less prestigious majors, because all thresholds will increase under the policy, but will follow the same application pattern conditional on new equilibrium entry thresholds. Prediction (1) corresponds to common trends, (2) to a positive effect of the intention to treat, and (3) to no effect on the control group beyond the general equilibrium externality through entry thresholds.

Moreover, the policy effects can be expected to be heterogeneous in individual ability. Fact 4 implies that if the error term is distributed normally the set of payoffs functions such that an individual will find it profitable to upgrade their major choice expands with academic ability. Hence, we would expect that any policy effect would be more pronounced for applicants with higher ENEM scores.

Thus equipped with theoretical guidance, we will now turn our attention towards the data and devote the remainder of this paper to an empirical examination of the affirmative action policy at UNICAMP.

4 Data

We use administrative individual-level data from the 2004 and 2006 UNICAMP admission exams (*vestibular*), obtained from COMVEST.[13] 2004 is the last pre-affirmative action year while 2006 is the first year for which we expect a reaction to the affirmative action policy in terms of major choice. Indeed, a fair number of applicants who registered for the 2005 exam learnt about the affirmative action policy while registering, giving them little time to think about changing their planned major.[14] Hence, we might not expect applicants to react to the affirmative action policy (in terms of major selection) in 2005. Although our main results are for 2004 and 2006, we have exam information from 2001 to 2007. We use this information to investigate the parallel-trend assumption (necessary for our identification strategy to be valid) and to check the robustness of our results in Section 8.[15]

Our dataset contains socioeconomic characteristics of all applicants who registered to take the admission exam, including age, gender, race, previous university attendance (yes or no), type of high school (public or private), and both parents' education levels and occupations, providing a rich set of control variables for our regressions. Parental income will also be used

[13] Comissão Permanente para os Vestibulares (COMVEST) is UNICAMP's admission office (https://www.comvest.unicamp.br/).

[14] Note that it is not clear whether, on the 2005 exam registration form, the question regarding major selection came before applicants were informed about the affirmative action or after.

[15] The main reason we do not use 2007 for our main results is that UNICAMP's main competitors, Universidade de Sao Paulo (USP) and the Universidade Estadual Paulista "Julio de Mesquita Filho" (UNESP), also implemented affirmative action policies in 2007, which could contaminate our results for years after 2006. We show that including 2007, however, does not affect our results in Section 8.

as the main regressor of interest (interacted with the affirmative action binary variable) when we investigate the effect of affirmative action on the parental socioeconomic gradient. We recovered the municipality of residence (99.6%) and the name of the institution where they completed high school (96.3%) for nearly our entire sample. Importantly, we also observe applicants' three major choices and all the grades obtained during the admission exam.

We complete our UNICAMP dataset with publicly available information on major-specific P_2 cutoff scores for the previous year. In our regression analysis, the P_2 cutoff score associated with the applicant's major choice will be one of our dependent variables of interest. Arguably, a greater cutoff is associated with a more competitive major. This information is not only available online, but it is also provided to all applicants when they register for the exam. Therefore, applicants should have a good idea of the competitiveness of each considered major before selecting them. Figure 3 illustrates some of the information provided to the 2004 prospective candidates, regarding the levels of competition/selectivity of different majors (COMVEST, 2004). For each major, applicants can see the number of seats available in the previous year (*vagas*), the number of applicants (*inscritos*), the ratio of applicants to seats in P_1 (*rel. C/V 1ª fase*), the number and percentage of applicants who pass P_1 (*aprovados 1ª fase*), the P_1 cutoff to move on to P_2 (*pontuação do último convocado para a 2ª fase*), the ratio of applicants to seats in P_2 (*rel. C/V 2ª fase*), and the cutoff on P_2 (*nota padronizada do último matriculado*).

When registering for the admission exam, individuals who took the ENEM exam can authorise UNICAMP to obtain their ENEM scores from the Ministry of Education. Not surprisingly, almost all applicants in our sample do so (i.e., around 90 per cent in each year's sample). For all applicants, UNICAMP calculates P_1 results with and without ENEM scores counting for 20 per cent of the grade and chooses the largest of the two scores. Therefore, the incentives to give permission are high, as ENEM scores can increase but not decrease P_1 scores.

To concentrate on our population of interest, we restrict our sample to individuals who took the exam for immediate admission (i.e., not as a practice test) and who completed their high school education in Brazil.[16] We also exclude applicants who registered but did not write the exam (4 per cent) and those applying for majors that require an aptitude test (5 per cent) as the selection into those majors requires specific abilities.[17]

Out of our sample of interest, we first discard individuals with missing socioeconomic variables (9.9 per cent), and without ENEM scores (7.7 per cent). Finally, given that we

[16]Individuals who want to write the admission exam as a practice test must indicate it in the registration form. Each year, roughly 4 per cent of exam takers write it as a practice test. About 0.5 per cent of applicants completed their high school abroad.

[17]We keep 'Dentistry' since its aptitude test only evaluates applicants' psychomotor coordination.

Dados do Vestibular Unicamp 2003

Relação candidatos-vaga nas 1ª e 2ª fases

Cursos	Vagas	Inscritos	Rel. C/V 1ª fase	Aprovados 1ª fase Nº	Aprovados 1ª fase %	Pontuação do último convocado para a 2ª fase (0 - 120)	Rel. C/V 2ª fase	Nota Padronizada do último matriculado
Arquitetura e Urbanismo (N)	30	986	32.9	241	24.4	64.9	8.0	753.51
Artes Cênicas (I)	25	489	19.6	92	18.8	60.0	3.7	664.43
Ciência da Computação (N)	50	1524	30.5	383	25.1	60.0	7.7	677.08
Ciências Biológicas (I)	45	1939	43.1	362	18.7	71.5	8.0	584.81
Ciências Biológicas (N)	45	854	19.0	318	37.2	60.0	7.1	525.22
Ciências Econômicas (I)	70	1508	21.5	562	37.3	64.3	8.0	683.38
Ciências Econômicas (N)	35	742	21.2	282	38.0	61.0	8.1	675.85
Ciências Sociais (I)	55	747	13.6	259	34.7	60.0	4.7	608.59
Ciências Sociais (N)	55	774	14.1	174	22.5	60.0	3.2	599.39
Dança (I)	25	269	10.8	83	30.9	47.8	3.3	596.15
Educação Artística (I)	30	261	8.7	91	34.9	52.2	3.0	598.55
Educação Física (I)	50	971	19.4	156	16.1	57.1	3.1	486.16
Educação Física (N)	50	562	11.2	150	26.7	48.0	3.0	454.31
Enfermagem - Unicamp (I)	40	823	20.6	125	15.2	60.0	3.1	428.84
Enfermagem - Famerp (I)	60	647	10.8	191	29.5	50.9	3.2	407.16
Eng. Agrícola (I)	70	512	7.3	233	45.5	52.2	3.3	478.92
Eng. Alimentos (I)	80	1205	15.1	439	36.4	60.0	5.5	655.69

Source: COMVEST (2004)

Figure 3: Applicant Manual Information on Previous Year Cutoffs

include the municipality of residence and high school fixed effects in our regressions, we eliminate singletons (municipalities and high schools, for which we observe only one individual) to avoid overstating the statistical significance of our coefficients of interest (Correia, 2015, 2016). This last restriction reduces our sample by 2.2 per cent. The final sample has 34,346 and 32,062 applicants in 2004 and 2006, respectively.

Table 1 presents some descriptive statistics for our sample of interest. By far the most popular major is medicine. Almost 20 per cent of all applicants choose medicine as their first major choice. Given this popularity, the fraction of applicants admitted to medicine is small: less than three per cent of individuals who applied to medicine are admitted (or 0.5 per cent of all applicants).[18] All the top-five most competitive majors are STEM majors. They are, in decreasing order of admission cutoff: Medicine, Computer Engineering (daytime), Control and Automation Engineering (evening), Electrical Engineering (daytime), Electrical Engineering (evening). The admission rate for these STEM majors is just around five per cent (1.5 per cent of all applicants apply for and obtain admission in these top-five majors). Roughly 30 per cent of our applicants completed their entire secondary education in a public high school, and the sample is evenly split between females and males.

[18] The number of places available in medicine stayed at 110 per year between 2000 and 2008.

Some applicant characteristics changed over our period of analysis, and we will have to control for these differences in our econometric analysis. Given that ENEM is a potentially important control variable for our analysis, we also estimated our regressions using variants of our 'raw' ENEM measure (e.g., normalized to zero each year, normalized using the average at level of the state of Sao Paulo each year, or normalized using the mean of individuals who should not be affected by the affirmative action) and the results are almost identical.[19] Parental education (both the father and mother education) has also improved between 2004 and 2006 among UNICAMP applicants. In particular, there were fewer parents without a high school diploma in 2006. In terms of occupation, there was an improvement for mothers while the situation stayed more or less the same for fathers. As we discuss in Estevan et al. (2019), time trends explain these variations, as successive cohorts are better-educated. However, given that these changes in education and occupation attainments could potentially influence applicants' major choice, we will control for these variables in each of our regressions.

Judging by the proportion of students applying to medicine or a top-five major, there is suggestive evidence that private high-school applicants, who represent more than 70 per cent of our sample, shied away from competitive majors. Indeed the proportion of students applying to these majors decreased between 2004 and 2006. Although such a decrease is in line with our theoretical pointers, presented in Section 3, it may also be due to other trends (unrelated to the affirmative action). Our econometric strategy will clarify how we attempt to measure the causal effect of the affirmative action policy on applicants' major choice.

5 Empirical Strategy

Given the quasi-random assignment nature of UNICAMP's affirmative action policy, we estimate its effects on major choice with a straightforward difference-in-difference model. Formally, for all our outcomes of interest, our regression equation takes the following form:

$$Z_{i,s,m,t} = \alpha P_i + \beta(P_i \times AA_t) + \phi ENEM_i + \boldsymbol{X_i}\boldsymbol{\Gamma} + \mu_m + \eta_s + \tau_t + \varepsilon_{i,s,m,t}, \tag{4}$$

where $Z_{i,s,m,t}$ is one of our major-choice measures (described below) for applicant i observed in year t, who attended high school s, and living in municipality m. P_i is equal to 1

[19] These results are available upon request. We use the unadjusted (or 'raw') ENEM score in our main estimations, as ENEM is meant to capture the ability signal received by the applicants. It is not clear to us whether applicants, when collecting and evaluating their ENEM performance, would take into account factors that could affect the average score (such as the test difficulty), defeating the purpose of normalizing the scores.

if the applicant went to a public secondary school, 0 otherwise; AA_t is equal to 1 if the applicant applied during UNICAMP's affirmative-action years, 0 otherwise; $ENEM_i$ is the grade obtained on the ENEM exam (an ability signal received by the applicant), and $\boldsymbol{X_i}$ is a vector of controls for the applicant's personal characteristics (i.e., gender, a quartic function of age, a dummy for previous university experience, parental educational attainment, and parental occupation). μ_m, η_s and τ_t are municipality, (last attended) high school and year fixed effects, respectively.[20] In all our regressions, our standard errors will be two-way cluster-robust (at the high-school and municipality levels).

In each of our regressions, the parameter of interest is β, the difference-in-difference parameter, which will measure how much the affirmative action policy changed the gap in the outcome variable between private and public high-school applicants. Since our regression framework controls for individual ENEM and a host of individual characteristics, according to our theoretical framework, the coefficient will identify the policy effects on the recipients of the bonus under Assumption 1, in Section 3.

We investigate the impact of UNICAMP's affirmative action on major choice by considering different margins. In particular, we look at 1) the likelihood to choose medicine, by far the most popular and arguably the most competitive major at UNICAMP, 2) the probability of selecting one of the top five most competitive majors based on the 2003 P_2 cutoffs, and 3) the 2003 P_2 cutoff of the program applied to. The first two outcomes concentrate on very competitive programs and are associated with prestigious occupations. If there is a positive correlation between major choice and the ability signal (ENEM), then the estimated β might be hiding some of the effects of the affirmative action as we could expect the 'action' only to occur for high ability students. For this reason, we also look at more comprehensive measures of major choice. By looking at the major cutoffs, we can more easily detect a change in major choice, for example, if candidates choose a slightly more competitive/prestigious major following the introduction of the admission policy. Comparing our results based on our different measures of major competitiveness/prestige, we will also be able to detect potential heterogeneity in the impact of the affirmative action.

Since a change of application behaviour is most relevant from a socioeconomic mobility point of view if students are also admitted in their chosen major, we will use the same empirical strategy to examine the (joint) probability of applying and being accepted to medicine and a top-five major. Finally, we will use a slightly modified version of equation (4) to investigate the effect of UNICAMP's affirmative action on the major-choice parental

[20]Note that, in specifications in which we include school fixed effects, the identification of the α parameter will operate through applicants who switched between public and private schools during their secondary education. Hence, one should be cautious when comparing the point estimates when controlling for school fixed effects or not. This is particularly the case if switchers are different from typical high school students.

income gradient—the main difference is to replace P_i by the parental income (in natural logs) in equation (4), see Section 9 for details.

6 Main Results

We begin by looking at whether UNICAMP's affirmative action policy lead public high-school students to apply for more selective majors. Given that we find an increase in application for more selective majors, we then turn to investigate whether the rise in application translated into more public high-school applicants being admitted in selective majors.

6.1 Applications to Prestigious Majors

Tables 2 and 3 present the estimated effect of the affirmative action on the probability to apply for medicine and a top-five major, respectively. The first three columns of these tables present results when we do not control for ENEM. Columns (2) and (3) sequentially add fixed effects for the applicant's municipality of residence and her (last attended) high school. The last three columns repeat the same exercise, but also control for ENEM scores. Comparing the parameter estimates for 'Public HS $\times$ AA' with and without ENEM controls (e.g. column (3) vs column (6)) will inform us on whether weaker public high school in terms of ENEM scores applied for UNICAMP, following the affirmative action, i.e., whether the partial correlation between 'Public HS $\times$ AA' and ENEM is positive or negative.[21]

Both Tables 2 and 3 suggest that the affirmative action significantly increased the likelihood of public high school students to apply for a selective major, relative to their private high school counterparts, and the patterns are similar across specifications and tables. In both cases, the estimates are slightly larger once we control for ENEM. Focussing on the specification (6) (our preferred specification), the point estimates for medicine and a top-five major are 1.0, and 2.0 percentage points, respectively. The estimates are statistically significant at conventional levels, with p-values of 0.053 for medicine and 0.005 for top-five majors, and the magnitude of these estimated effects is also considerable. Compared to the proportion of public high school applicants who chose these majors in 2004, the estimates suggest a 10.8, and 9.8 per cent increase following the implementation of the affirmative action for medicine, and top-five majors, respectively.[22]

[21] Given that columns (3) and (6) control for school fixed effects, the estimates for the 'Public High School' parameter from these specifications come from applicants who switched secondary school and therefore may not be comparable the ones in specifications where we do not control for such fixed effects.

[22] The proportions of public high school candidates who applied for medicine and a top-five major in 2004 are 9.4, and 20.8 per cent, respectively.

While Tables 2 and 3 consider whether public high-school students applied more for the most selective programs offered at UNICAMP, Table 4 investigates whether our findings hold more generally by considering application for all majors and using the previous year P_2 cutoff (the smallest P_2 score admitted to the major) as a selectivity measure. These cutoff scores are available to applicants when they register to the admission exam, as illustrated in Figure 3.

Results from Table 4 are in line with our previous findings.[23] Again the estimates are larger when controlling for ENEM. When focussing on the specification (6), the results suggest that the affirmative action policy led students from public high-school students to select more competitive majors. The estimated effect is 4.3 points, which might look small when compared to the average major cutoff for public high-school applicants in 2004, but it still represents 20 per cent of the observed 2004 private-public difference in the specification (5), 21.827 points.

Overall, the results suggest that the affirmative action policy did increase the selectivity of the majors chosen by public high-school applicants, relative to their private high-school counterparts. The effect is sizeable, especially for the most selective majors.

6.2 Admission to Prestigious Majors

Table 5 presents the estimated effect of the affirmative action on the (joint) probability to apply for and be admitted to medicine. The outcome variable is hence a binary variable equal to 1 if the student applied for and was accepted to medicine, and zero otherwise. The sample is composed of all applicants (not just medicine applicants). We focus on the joint probability as opposed to the admission probability conditional on applying because consistently estimating the effect of the affirmative action on the latter becomes challenging given the policy affected the likelihood of applying for medicine.

Table 5 suggests that UNICAMP's affirmative action increased the likelihood of public high school applicants to apply for and be admitted in medicine, relative to their private high school counterparts. As expected, columns (1), (2), (4) and (5) suggest that, before the affirmative action, public high school applicants were less likely to apply and be admitted to medicine. Columns (1) to (3) (when we do not control for ENEM) do not suggest that the affirmative action increased the likelihood of public high-school applicants to be admitted to medicine, but the magnitude of the estimates are still substantial. Though not statistically significant, the point estimates for these specifications are large. They fluctuate between 0.08 and 0.15 percentage points, which represent increases of 80 and 150 per cent,

[23]Note that the number of observations is slightly smaller in Table 4 as some majors were introduced in 2004 and therefore did not have a 2003 cutoff. We exclude these majors from our sample.

respectively. The estimates for the effect of affirmative action become statistically significant at conventional significance levels when we control for ENEM. The estimated impact is also substantial: compared to the proportion of public high school applicants who were admitted to medicine in 2004 (0.1 percent), the estimate found in column (6) suggests a 153 per cent increase in the proportion of public high school applicants choosing and being admitted to medicine, following the implementation of the affirmative action.

Given that we are looking at the joint probability of applying and being accepted, there are few explanations for the large size of the estimate: 1) the probability of being admitted, conditional on applying, could have 'mechanically' increased significantly due to the 30-point bonus; and/or 2) public high school applicants, knowing that it would be easier to be admitted, may have increased their likelihood to apply to medicine. The observed increased in medicine might also not be representative of what we observe for a typical student, as medicine attracts applicants that are significantly better in terms of ENEM scores and, importantly, more homogeneous in terms of P_2 scores. This homogeneity is essential here since it is associated with a larger relative magnitude of the 30-point bonus (in terms of public high school applicant shift within the P_2 score distribution) relative to other majors. To investigate whether we observe similar patterns for a more substantial part of the applicant distribution, we now turn to the likelihood of applying and being admitted to a top-five major.

Table 6 presents results from estimating equation (4) on the joint probability to apply and be admitted to a top-five major. Again, the regressions are estimated on all applicants, whether they applied for a top-five major, or not. The results are in line with the ones found in Table 5. Table 6 suggests that public high school applicants increased their likelihood to apply and be admitted to a top-ten major by between 0.56 and 0.83 percentage points following the affirmative action. The magnitude of the effect is also large: specification (6) suggests a 93 per cent increase when compared to the proportion of public high school applicants who applied and were admitted to a top-five major in 2004 (0.87 per cent). Note that controlling for ENEM (columns (3) through (6)) does not change our estimated affirmative action effect dramatically. If anything, seeing that its magnitude slightly increases when controlling for ENEM suggests that slightly weaker public high school students (based on ENEM scores) applied to UNICAMP following the affirmative action. Finding that the magnitude of the affirmative action effect is slightly smaller for top-five major relative to medicine is not surprising as we expect the salience of the 30-point bonus to be larger for majors in which applicants are more homogeneous (and with a smaller initial proportion of public high-school applicants.

To summarize, the results presented in Tables 2 through 6 all suggest a large positive impact of the affirmative action on public high school students' major choice and their representation among selective programs.

7 Heterogeneity

We now investigate potential heterogeneity in the effect of the affirmative action on major choice. First, given that many studies find that females and males react differently to educational reforms, we investigate whether a specific gender drives our results. Second, motivated by Fact 4 resulting from our theoretical framework and by the results of Bond et al. (2018) who find that higher-ability students react (update) more to SAT shocks than lower-ability students, we investigate whether applicants with higher ENEM scores responded differently than the rest of the applicants.

7.1 Gender Differences

In order to detect gender differences in the reaction to the affirmative action, we augment our regression equation (4) with the gender interaction terms 'Female × Public HS,' 'Female × AA,' and 'Female × Public HS × AA.' The coefficient estimate for 'Female × AA' will inform us if females from private high schools reacted differently from their male counterparts while the sum of the coefficient estimates for 'Female × AA' and 'Female × Public HS × AA' will inform us whether females from public high schools reacted differently than their male counterparts. Note that now the β estimate will capture the effect of the affirmative action on the application (admission) public-private gap among male applicants.

Table 7 presents the results when we include gender interaction terms. When we concentrate on Medicine, we do not observe any gender differences in terms of the effect of affirmative action on application behaviour. The main difference between the results presented in this table and those from Table 2 is that the although the estimates for β are almost identical, the standard errors are larger in Table 7, making β no longer statistically significant at standard levels (the p-value is .164). When we look at the probability to apply and be admitted, our results suggest that female applicants from private high schools saw their admission rates drop the most following the affirmative action. The admission probability of public high-school female applicants does not seem to have changed differently than their male counterparts' following the introduction of the affirmative action as the sum of the coefficients for 'Female × AA' and 'Female × Public HS × AA' is not statistically different from 0 (p-value=0.239).

The results regarding admission to a top-five program (i.e. column (3)) do not suggest any gender differences in the effect of affirmative action. However, when we look at the probability to apply for a top-five major, we see interesting results. While the comparison of the parameter estimates for 'Public HS × AA' in Tables 3 and 7 suggests similar average effects for male applicants to the overall effect, the results for the female interaction terms suggest that females reacted differently. In particular, female applicants from private high schools increased their likelihood to apply for a top-five major relative to their male counterparts, so much so that the increase is similar to the one observed among females and males from public high schools.[24] Given that we observe only relative differences, it is not possible to say whether female applicants from private high schools increased their application likelihood more, or decreased less than their male counterparts. If we rule out the possibility that they increased their application likelihood following a decrease in admission probability, our results suggest that our results for top-five majors are driven by private high-school male applicants who became significantly less likely to apply.

Overall, our results do not suggest substantial gender differences in applicants' reaction to the affirmative action, at least within public high-school applicants. One caveat regarding allowing for the affirmative action to affect females and males differently is that our estimates seem to become slightly imprecise.

7.2 ENEM Quartile Differences

Bond et al. (2018) found that higher-ability students reacted more to an SAT shock than lower-ability ones. Although their SAT shocks are likely to affect both students' likelihood to be accepted in more competitive colleges and their perceived ability whereas our treatment is expected to affect the former only, it is conceivable that we observe the same pattern with our affirmative action policy.

To investigate the potential heterogeneity in the affirmative-action effect across student ability, we split our sample into for quartiles based on their ENEM score (where Q1 is the bottom quartile and Q4, the top). We estimate equation (4) separately for each quartile, for both 1) the likelihood to apply for medicine and 2) the likelihood to apply for a top-five major. We present the results in Tables 8 and 9.[25] The results are similar whether we look at medicine or top-five majors. In both cases, the affirmative action seems to have had the most significant positive impact on students with ENEM scores in the top quartile. A

[24] The sum of the coefficients for 'Female × AA' and 'Female × Public HS × AA' is not statistically different from 0 (p-value=0.224), suggesting that female and male applicants from public high schools reacted similarly to the introduction of the affirmative action.

[25] The number of observations per regression does not represent exactly 25% of (and add up to less than) the original sample since, as we split our sample in four, we end up with an additional number of singletons, which are automatically dropped from the regressions (about 6% of our observations).

bit surprisingly, the affirmative action seems to have decreased the likelihood to apply for medicine or a top-five major for public high school students in Q1.[26] When we concentrate on top-five programs, we see that the effect steadily increases as we move from the bottom to the top quartile.[27] These findings are in line with Bond et al. (2018) as well as with our theoretical framework (see Fact 4), suggesting that higher ability students react more to the incentive from the affirmative action (and this is more so for the most competitive program).

Appendix Tables A.1 and A.2 present the results for a similar exercise done on the probability to be admitted to medicine and top-five majors, respectively. Somewhat not surprisingly given the competitiveness of these programs, the effect of the affirmative action is concentrated in the top quartile. The effects are very large for students in Q4, for both medicine and top-five majors.[28]

8 Robustness Checks

8.1 Pre-Trends and Additional Years

The central assumption underlying the use of a difference-in-difference estimator is that, in the absence of the affirmative action policy, the application behaviour of private and public high school applicants would have followed the same trend. To check the validity of this assumption, we use data from 2001 to 2004 (a pre-affirmative action period) and estimate the change in the public-private gap in the probability to apply for medicine and a top-five major. We do so by running regressions of our variables of interest on the same regressors as in equation (4), but where we replace $P_i \times AA_t$ with a series of interaction terms between our year fixed effects and P_i. The coefficient estimates will capture how the public-private gap changed from year to year, before the affirmative action.

Table 10 presents the results from the above exercise for the joint probability of applying and being admitted and the likelihood of applying to our two types of majors. When we look at the results for medicine, we can see that the trends in the public-private gaps were stable before the affirmative action for both outcomes of interest. Column (1) suggests that the gap in admission to medicine stayed unchanged over the 2001-2004 period. An F-test

[26] We also see that the proportion of 2004 public high school students with ENEM scores in Q1 applying to medicine is larger than for similar students in Q2 (8.3% vs 6.9%), which suggests that Q1 applicants could be more 'overly optimistic' when it comes to their likelihood of being admitted. In our sample, no one with an ENEM score in Q1 was admitted to medicine or a top-five major, and a handful of students in Q2 were.

[27] Something similar occurs for medicine except that the parameter for applicants in Q3 is not statistically significant.

[28] Since no applicant in Q1 was accepted in medicine or a top-five major, we cannot estimate our regressions on these individuals.

on the joint statistical significance of the interaction-term parameters does not reject the hypothesis that these parameters are all equal to zero at conventional significance levels.

When we look at the likelihood to apply for medicine, while two out of the three estimated coefficients are significant at conventional significance levels, we can see that these estimates are essentially identical suggesting that, although different from 2001, the private-public gap stayed stable after 2002.[29] In fact, despite having two statistically significant interaction terms, an F-test on the joint significance of the three interaction terms suggest they are not different from 0.

The results are very similar for top-five majors. Only one of the six parameters is significant at 10 per cent (with a p-value of .097). None of the two F-tests rejects the null hypothesis that the interaction terms are all equal to 0. Both the magnitude of interaction-term parameter estimates and their lack of statistical significance suggest that the gap in application and admission to top-five majors stayed very stable during the years preceding the affirmative action policy. These findings suggest that our results in Tables 2 and 3 are not capturing a differential in trends where the public high school applicants apply more and more to selective majors over time.

Overall, the results from Table 10 suggest that our main results are not driven by the pre-affirmative action trends in the public-private gaps. In particular, the gap in applying for our majors of interest was stable between 2002 and 2004. Tables A.3 and A.4 complement the analysis of the private-public pre-trends by showing the results of a placebo exercise using only 2003-2004, the two years preceding the affirmative action policy. Doing so informs us on whether the 2004 (our base year in the main analysis) private-public application gap was abnormally large. Concentrating on a shorter period gets around the issue of having parameters of essential control variables (e.g., ENEM) potentially changing over the long run. The results support the absence of diverging pre-existing trends, as the coefficient estimate for the placebo interaction term is very close to zero and statistically insignificant.

Given that the pre-affirmative trends for private and public high school applicants are similar, we can investigate whether the impact of the affirmative action estimated over the 2004-2006 period stays similar if we expand the pre and post periods. Table 11 presents the results from estimating the effect of UNICAMP's affirmative action when we expand the covered period to 2002-2007 (still leaving 2005 out of the analysis, as in Tables 2 through 9). The results for the joint probability of applying and being admitted presented in Table 11 are surprisingly similar to the ones focussing on 2004 and 2006, despite adding two pre-policy and one post-policy years of observations. If anything, the results for applying to medicine or a top-five major suggest slightly larger estimates when we include the additional years of

[29] The p-values for Public HS $\times$ 2003 and Public HS $\times$ 2004 are .048 and .090, respectively.

observations, but the differences are all within two standard errors of the original estimates. In all cases, the estimated effects are highly statistically and economically significant. These findings suggest that the choice of using 2004 as 'control year' as opposed to using multiple years is not crucial to the analysis.[30]

8.2 Age Restriction

In our main regressions, we do not make significant age restrictions. In our sample, some students are younger than 17, and some are older than 23 (the ages at which we would expect applicants to apply for university).[31] Here we investigate whether our estimates are robust to restricting our sample to 17-23 years of age, corresponding to the 5th and 95th age percentiles in our original sample.

Results from Table 12 are very similar to those presented in Tables 2 through 6, although the statistical significance for admission to medicine is just outside the standard significance levels with a p-value of 0.121. Overall, these results suggest that our results are not driven older (or younger) applicants than the typical ones.

9 Affirmative Action and the Parental-Income Gradient

By modifying applicants' major choice and their probability of being admitted to more selective majors, the affirmative action policy could have impacted the intergenerational mobility of low SES individuals. We now investigate whether the affirmative action policy changed the link between parental income and major selection/admission. To investigate this question, we slightly modify our regression model presented in equation (4). Essentially, we replace our public-school indicator variable by the log of the parental income.[32] For these regressions, we do not control for high-school fixed effects or previous university attendance as they are (probably) determined by parental income. Tables 13 through 16 present the results for medicine and top-five majors. We do the estimations for both the 2004-2006 and 2002-2007 periods (Panels A and B, respectively).

[30]One potential confounding factor in 2007 is that USP introduced its affirmative action policy that year and could have had consequences on UNICAMP applicants' major decision. We also obtain similar results using 2002-2006, which are available upon request. Finding that the results for 2002-2006 are very similar to 2002-2007 suggest that USP's affirmative action policy did not have strong 'general equilibrium' effects.

[31]For competitive programs, some applicants take a few years to prepare for the admission test, which explains why we have many applicants in their early 20s).

[32]Appendix A for a description of how we construct the income variable.

Panel A of Tables 13 and 14 suggest that the link between parental income and major selection changed immediately following the affirmative action. Concentrating on the specification (4), we see that before the affirmative action, a 10-percent increase in parental income increased the likelihood of applying to medicine (a top-five major) by 0.7 (0.6) percentage points. For medicine, this gradient decreased by 0.12 percentage points (or 17 per cent) following the affirmative action. The decrease is even more substantial for top-five programs, representing a drop of 27 per cent. As with our previous results, the effect of the affirmative action policy is slightly larger when we control for ENEM. The results for the 2002-2007 period (Panel B) present similar results. If anything, Panel B suggests a more pronounced decrease in the parental income gradient following the introduction of the affirmative action.

Tables 15 and 16 focus on the parental income gradient with respect to admission to medicine and a top-five major. Panel A of Tables 15 and 16 do not suggest that the parental income gradient in admission likelihood changed significantly between 2004 and 2006. The results are different when we consider the 2002-2007 (Panel B of Tables 15 and 16). Once we control for ENEM, the affirmative action seems to have decreased the parental-income gradient significantly. In fact, after the introduction of the affirmative action, the parental income gradient is no longer statistically different from zero.[33]

10 Conclusion

In this paper, we investigate whether an affirmative action policy implemented by a Brazilian university in 2005 altered the major choices of beneficiaries versus non-beneficiaries and their admission probability to more selective majors, typically associated with higher earnings.

The affirmative action policy effectively increased the choice set of targeted applicants, and we show that they reacted by choosing more selective majors. Applicants from public high schools were approximately 10 per cent more likely to select medicine (the most competitive major) or a top-five major following the implementation of the policy. This behavioural change was associated with a substantial increase in the joint probability of applying and being admitted to medicine and a top-five major, close to 100 per cent. The policy managed to close the gap in major applications between students with different parental income backgrounds, and there also is some evidence that the affirmative action decreased the parental-income gradient for the admission probability to selective majors. Finally, as Bond et al. (2018), we find that higher ability applicants reacted more to the change in admission probability triggered by the affirmative action.

[33]We obtain similar results if we consider the parental education gradient (using categories) instead of the parental income gradient. The results are available upon request.

Our results have significant policy implications. They propose an essential role for factual constraints in shaping application and admission behaviour. They also suggest that well-designed public policies may address unequal access to high-paying majors and lead to higher social mobility.

Finding that a university managed to increase the proportion of underprivileged (i.e., public high-school) applicants choosing and being admitted to more competitive and prestigious majors, suggests future research questions. In particular, it would be interesting to study whether the affirmative action had longer-term effects on outcomes such as university performance, major completion and eventually occupations and wages.

References

Abdulkadiroglu, A. and Sonmez, T. (2003). School choice: A mechanism design approach. *American Economic Review*, 93(3):729–747.

Altonji, J., Arcidiacono, P., and Maurel, A. (2016). The analysis of field choice in college and graduate school: Determinants and wage effects. In Eric A. Hanushek, S. M. and Woessmann, L., editors, *Handbook of the Economics of Education*, volume 5, pages 305–396. Elsevier.

Altonji, J. G., Blom, E., and Meghir, C. (2012). Heterogeneity in human capital investments: High school curriculum, college major, and careers. *Annual Review of Economics*, 4(1):185–223.

Anelli, M. and Peri, G. (2015). Gender of siblings and choice of college major. *CESifo Economic Studies*, 61(1):53–71.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, 121(1-2):343–375.

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

Arcidiacono, P., Hotz, V. J., and Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, 166(1):3–16.

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

Avery, C., Gurantz, O., Hurwitz, M., and Smith, J. (2018). Shifting college majors in response to advanced placement exam scores. *Journal of Human Resources*, 53(4):918–956.

Beffy, M., Fougère, D., and Maurel, A. (2012). Choosing the field of study in postsecondary education: Do expected earnings matter? *Review of Economics and Statistics*, 94(1):334–347.

Bond, T. N., Bulman, G., Li, X., and Smith, J. (2018). Updating human capital decisions: Evidence from SAT score shocks and college applications. *Journal of Labor Economics*, 36(3):807–839.

Bordon, P. and Fu, C. (2015). College-major choice to college-then-major choice. *Review of Economic Studies*, 82(4):1247–1288.

Boudarbat, B. and Montmarquette, C. (2009). Choice of fields of study of university Canadian graduates: The role of gender and their parents' education. *Education Economics*, 17(2):185–213.

Camera, L. (2017). Low-income students nowhere to be found in STEM. *U.S. News*, May 25. https://www.usnews.com/news/stem-solutions/articles/2017-05-25/low-income-students-nowhere-to-be-found-in-stem.

Carrell, S. (2016). Students from wealthy backgrounds dominate medical schools. *The Guardian*, January 22. https://www.theguardian.com/society/2016/jan/22/medical-school-students-wealthy-backgrounds.

Carvalho, J.-R., Magnac, T., and Xiong, Q. (2014). College choice allocation mechanisms: Structural estimates and counterfactuals. IZA Discussion Papers 8550, Institute for the Study of Labor (IZA).

Cavalcanti, T., Guimaraes, J., and Sampaio, B. (2010). Barriers to skill acquisition in Brazil: Public and private school students performance in a public university entrance exam. *Quarterly Review of Economics and Finance*, 50(4):395–407.

Chade, H., Lewis, G., and Smith, L. (2014). Student portfolios and the college admissions problem. *Review of Economic Studies*, 81(3):971–1002.

COMVEST (2004). Manual do vestibular UNICAMP. Comissao Permanente para os Vestibulares. http://www.comvest.unicamp.br/vest2004/manual.pdf.

Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives*, 27(3):79–102.

Correia, S. (2015). Singletons, cluster-robust standard errors and fixed effects: A bad mix. Technical report, Duke University.

Correia, S. (2016). Linear models with high-dimensional fixed effects: An efficient and feasible estimator. Technical report, Duke University.

DellaVigna, S. (2010). The Obama effect on economic outcomes: Evidence from event studies. *mimeo*, pages 1–36.

Epple, D., Romano, R., and Sieg, H. (2006). Admission, tuition, and financial aid policies in the market for higher education. *Econometrica*, 74(4):885–928.

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

Ferrando, M. and Gille, V. (2016). Does the identity of leaders matter for education? Evidence from the first black governor in the US. *mimeo*, pages 1–32.

Fricke, H., Grogger, J., and Steinmayr, A. (2018). Exposure to academic fields and college major choice. *Economics of Education Review*, 64(1):199–213.

Fu, C. (2014). Equilibrium tuition, applications, admissions, and enrollment in the college market. *Journal of Political Economy*, 122(2):225–281.

Gemici, A. and Wiswall, M. (2014). Evolution of gender differences in post-secondary human capital investments: College majors. *International Economic Review*, 55(1):23–56.

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

IBGE (2010). [2010 Population Census]. Censo Demográfico 2010 (in Portuguese). SIDRA version, http://www.sidra.ibge.gov.br/cd/defaultcd2000.asp?o=11&i=P.

INEP (2005). [Basic Education Summary Statistics]. Sinopse Estatística da Educação Básica (in Portuguese). Available on http://portal.inep.gov.br/sinopses-estatisticas-da-educacao-basica. Accessed: 2018-08-07.

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

Montmarquette, C., Cannings, K., and Mahseredjian, S. (2002). How do young people choose college majors? *Economics of Education Review*, 21(6):543–556.

NSF (2019). Women, minorities, and persons with disabilities in science and engineering: 2019. Technical report, National Center for Science and Engineering Statistics, Special Report NSF 19-304. Alexandria, VA.

Papay, J. P., Murnane, R. J., and Willett, J. B. (2016). The impact of test score labels on human-capital investment decisions. *Journal of Human Resources*, 51(2):357–388.

Steven, K., Dowell, J., Jackson, C., and Guthrie, B. (2016). Fair access to medicine? Retrospective analysis of UK medical schools application data 2009-2012 using three measures of socioeconomic status. *BMC Medical Education*, 16(11):1–10.

Stinebrickner, R. and Stinebrickner, T. R. (2014). A major in science? Initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies*, 81(1):426–472.

van de Werfhorst, H. G. and Luijkx, R. (2010). Educational field of study and social mobility: Disaggregating social origin and education. *Sociology*, 44(4):695–715.

Wiswall, M. and Zafar, B. (2015). Determinants of college major choice: Identification using an information experiment. *Review of Economic Studies*, 82(2):791–824.

Zafar, B. (2011). How do college students form expectations? *Journal of Labor Economics*, 29(2):301–348.

Appendix A – Construction of the Parental Income Variable

In order to investigate whether UNICAMP's affirmative action policy affected the link between parental income and program selection/admission, we use the family income information reported by the applicants. Total family income is reported in categories and as a function of the minimum wage (e.g., 10-15 times the minimum wage). For each of these categories except for the top one, we use the mid-point of the interval as income measure. We use 1.5 times the max value for the top-coded category (i.e. more than 40 times the minimum wage). In the regressions presented in Tables 13 through 16, we use the natural log of the mid-points as main regressor. We interact the log of the family income with the affirmative-action dummy variable (AA_t) to see whether the affirmative changed the parental-income gradient.

As a robustness check, we have also used as main regressor a family-size adjusted income measure where the mid-points are divided by the square root of the family size. The results, available upon request, are very similar to the ones where we use the unadjusted income measure.

Appendix B – Main Tables

Table 1: Descriptive Statistics

	2004	2006	Difference	
Applied to Medicine (%)	0.196	0.181	-0.014	***
Applied to Top Five Major (%)	0.308	0.282	-0.026	***
Admitted to Medicine (%)	0.005	0.005	0.001	
Admitted to Top Five Major (%)	0.015	0.015	0.000	
ENEM Score (0-10)	7.5	7.1	-0.343	***
	(1.5)	(1.5)		
Public High School (%)	0.289	0.297	0.007	**
Female (%)	0.506	0.506	-0.001	
Age	19.2	19.3	0.112	***
	(2.0)	(2.2)		
Mother without HS Degree (%)	0.241	0.221	-0.021	***
Mother with HS Degree (%)	0.322	0.320	-0.003	
Mother with Univ. Degree (%)	0.437	0.460	0.023	***
Father without HS Degree (%)	0.235	0.223	-0.012	***
Father with HS Degree (%)	0.287	0.293	0.006	*
Father with Univ. Degree (%)	0.479	0.484	0.005	
Mother with Manual Occ. (%)	0.067	0.062	-0.005	**
Mother with Mid-Top Occ. (%)	0.299	0.273	-0.026	***
Mother with Top Occ. (%)	0.276	0.309	0.034	***
Mother with Other Occ. (%)	0.359	0.356	-0.003	
Father with Manual Occ. (%)	0.120	0.128	0.008	***
Father with Mid-Top Occ. (%)	0.335	0.305	-0.030	
Father with Top Occ. (%)	0.483	0.481	-0.002	
Father with Other Occ. (%)	0.063	0.087	0.024	***
Previous University Attendance (%)	0.062	0.057	-0.006	***
Observations	34,346	32,062		
Cutoff Phase 2 (previous year)	630.6	537.2	-93.4	***
	(93.8)	(81.3)		
Observations	31,470	29,605		

Notes: Standard deviations are in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2: Affirmative Action and Medicine UNICAMP Application

	(1)	(2)	(3)	(4)	(5)	(6)
ENEM Score				0.036***	0.030***	0.027***
				(0.003)	(0.003)	(0.002)
Public High School	-0.102***	-0.100***	-0.018*	-0.087***	-0.089***	-0.018**
	(0.010)	(0.012)	(0.009)	(0.010)	(0.011)	(0.009)
Public HS × AA	0.009	0.020**	0.006	0.018**	0.026***	0.010*
	(0.007)	(0.008)	(0.005)	(0.007)	(0.008)	(0.005)
Controls:						
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	No	Yes	Yes	No	Yes	Yes
High School Fixed Effects	No	No	Yes	No	No	Yes
Mean of Dep. Var. (2004, Public)	0.094	0.094	0.094	0.094	0.094	0.094
Observations	66,408	66,408	66,408	66,408	66,408	66,408
Municipality Clusters	810	810	810	810	810	810
High School Clusters	3,556	3,556	3,556	3,556	3,556	3,556

Notes: The dependent variable is a binary variable equal to 1 if the applicant applied to Medicine at UNICAMP, 0 otherwise. The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Standard deviations are in square brackets. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Affirmative Action and Application to a Top-Five UNICAMP Major

	(1)	(2)	(3)	(4)	(5)	(6)
ENEM Score				0.044***	0.036***	0.032***
				(0.004)	(0.004)	(0.002)
Public High School	-0.078***	-0.072***	-0.018	-0.060***	-0.058***	-0.019
	(0.010)	(0.013)	(0.015)	(0.009)	(0.011)	(0.015)
Public HS × AA	0.013*	0.026***	0.016**	0.023***	0.032***	0.020***
	(0.008)	(0.009)	(0.007)	(0.008)	(0.009)	(0.007)
Controls:						
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	No	Yes	Yes	No	Yes	Yes
High School Fixed Effects	No	No	Yes	No	No	Yes
Mean of Dep. Var. (2004, Public)	0.208	0.208	0.208	0.208	0.208	0.208
Observations	66,408	66,408	66,408	66,408	66,408	66,408
Municipality Clusters	810	810	810	810	810	810
High School Clusters	3,556	3,556	3,556	3,556	3,556	3,556

Notes: The dependent variable is a binary variable equal to 1 if the applicant applied to one of the top-five most competitive UNICAMP majors (i.e., Medicine, Computer Engineering-daytime, Control and Automation Engineering-evening, Electrical Engineering-daytime, Electrical Engineering-evening) based on the 2003 Phase 2 cutoffs, 0 otherwise. The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Standard deviations are in square brackets. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Affirmative Action and Major Choice Phase 2 Cutoff

	(1)	(2)	(3)	(4)	(5)	(6)
ENEM Score				13.148***	11.558***	10.352***
				(0.423)	(0.354)	(0.298)
Public High School	-30.871***	-26.002***	-7.289**	-25.429***	-21.827***	-7.540**
	(2.861)	(2.204)	(3.138)	(2.235)	(1.787)	(3.117)
Public HS× AA	2.829	5.813***	2.587	5.894***	7.894***	4.255**
	(2.257)	(1.996)	(1.764)	(2.236)	(1.981)	(1.789)
Controls:						
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	No	Yes	Yes	No	Yes	Yes
High School Fixed Effects	No	No	Yes	No	No	Yes
Mean of Dep. Var.	595.8	595.8	595.8	595.8	595.8	595.8
(2004, Public)	[100.2]	[100.2]	[100.2]	[100.2]	[100.2]	[100.2]
Observations	61,075	61,075	61,075	61,075	61,075	61,075
Municipality Clusters	794	794	794	794	794	794
High School Clusters	3,470	3,470	3,470	3,470	3,470	3,470

Notes: The dependent variable is the previous year Phase 2 cutoff (i.e. the Phase 2 score of the last admitted applicant) of the major chosen by the applicant. The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Standard deviations are in square brackets. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Affirmative Action and Medicine UNICAMP Admission (All Applicants)

	(1)	(2)	(3)	(4)	(5)	(6)
ENEM Score				0.004***	0.004***	0.004***
				(0.001)	(0.001)	(0.001)
Public High School	-0.003***	-0.003***	1.19E-04	-0.001**	-0.001**	3.61E-05
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Public HS × AA	0.001	0.001	0.001	0.002**	0.002**	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Controls:						
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	No	Yes	Yes	No	Yes	Yes
High School Fixed Effects	No	No	Yes	No	No	Yes
Mean of Dep. Var. (2004, Public)	0.001	0.001	0.001	0.001	0.001	0.001
Observations	66,408	66,408	66,408	66,408	66,408	66,408
Municipality Clusters	810	810	810	810	810	810
High School Clusters	3,556	3,556	3,556	3,556	3,556	3,556

Notes: The dependent variable is a binary variable equal to 1 if the applicant applied to and was admitted to Medicine at UNICAMP, 0 otherwise. The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Standard deviations are in square brackets. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Affirmative Action and Admission to Top-Five UNICAMP Majors (All Applicants)

	(1)	(2)	(3)	(4)	(5)	(6)
ENEM Score				0.011***	0.011***	0.011***
				(0.001)	(0.001)	(0.001)
Public High School	-0.001	-0.001	-2.00E-04	0.004**	0.003*	-3.98E-04
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)
Public HS × AA	0.006***	0.006***	0.006***	0.008***	0.008***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Controls:						
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	No	Yes	Yes	No	Yes	Yes
High School Fixed Effects	No	No	Yes	No	No	Yes
Mean of Dep. Var. (2004, Public)	0.009	0.009	0.009	0.009	0.009	0.009
Observations	66,408	66,408	66,408	66,408	66,408	66,408
Municipality Clusters	810	810	810	810	810	810
High School Clusters	3,556	3,556	3,556	3,556	3,556	3,556

Notes: The dependent variable is a binary variable equal to 1 if the applicant applied and was admitted to one of the top-five most competitive UNICAMP majors (i.e., Medicine, Computer Engineering-daytime, Control and Automation Engineering-evening, Electrical Engineering-daytime, Electrical Engineering-evening) based on the 2003 Phase 2 cutoffs, 0 otherwise. The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Standard deviations are in square brackets. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Estimating the Impact of the Affirmative Action with Gender-Interaction Terms (2004-2006)

	(1)	(2)	(3)	(4)
	Medicine		Top-Five	
	App. & Adm.	Applied	App. & Adm.	Applied
ENEM Score	0.004***	0.026***	0.011***	0.031***
	(0.001)	(0.002)	(0.001)	(0.002)
Public High School	0.001	0.008	-0.003**	0.008
	(0.001)	(0.010)	(0.002)	(0.020)
Public HS × AA	-0.001	0.010	0.009***	0.023*
	(0.001)	(0.007)	(0.003)	(0.012)
Female × Public HS	-0.001	-0.053***	0.006*	-0.052***
	(0.001)	(0.008)	(0.003)	(0.015)
Female × AA	-0.003**	-0.004	3.05E-04	0.020***
	(0.001)	(0.006)	(0.002)	(0.008)
Female × Public HS × AA	0.005**	0.001	-0.001	-0.007
	(0.002)	(0.009)	(0.004)	(0.013)
Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes
High School Fixed Effects	Yes	Yes	Yes	Yes
Observations	66,408	66,408	66,408	66,408
Municipality Clusters	810	810	810	810
High School Clusters	3,556	3,556	3,556	3,556

Notes: The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8: Affirmative Action and Medicine UNICAMP Application: by ENEM Quartile

	Q1	Q2	Q3	Q4
	(1)	(2)	(3)	(4)
ENEM Score	-0.002	0.005	0.052***	0.163***
	(0.003)	(0.010)	(0.012)	(0.013)
Public High School	-0.005	-0.028**	-0.023	-0.065**
	(0.018)	(0.011)	(0.020)	(0.029)
Public HS $\times$ AA	-0.023**	0.023***	0.017	0.040***
	(0.010)	(0.008)	(0.013)	(0.013)
Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes
High School Fixed Effects	Yes	Yes	Yes	Yes
Mean of Dep. Var. (2004, Public)	0.084	0.070	0.092	0.154
Observations	15,581	15,496	15,520	15,666
Municipality Clusters	402	387	402	370
High School Clusters	2,078	1,789	1,602	1,443

Notes: The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 9: Affirmative Action and Application to a Top-Five UNICAMP Major: by ENEM Quartile

	Q1	Q2	Q3	Q4
	(1)	(2)	(3)	(4)
ENEM Score	0.001	1.65E-04	0.072***	0.160***
	(0.003)	(0.010)	(0.014)	(0.010)
Public High School	0.029	-0.057*	-0.043*	-0.102***
	(0.019)	(0.032)	(0.024)	(0.028)
Public HS × AA	-0.021*	0.030***	0.034**	0.046***
	(0.011)	(0.011)	(0.017)	(0.017)
Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes
High School Fixed Effects	Yes	Yes	Yes	Yes
Mean of Dep. Var. (2004, Public)	0.163	0.174	0.225	0.345
Observations	15,581	15,496	15,520	15,666
Municipality Clusters	402	387	402	370
High School Clusters	2,078	1,789	1,602	1,443

Notes: The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 10: Pre-Affirmative Action Trends (2001-2004)

	(1)	(2)	(3)	(4)
	Medicine		Top-Five	
	App. & Adm.	Applied	App. & Adm.	Applied
ENEM Score	0.005***	0.031***	0.013***	0.042***
	(0.001)	(0.002)	(0.001)	(0.002)
Public High School	0.001	-0.008	3.94E-05	0.008
	(0.001)	(0.009)	(0.002)	(0.012)
Public HS × 2002	-0.001	0.005	-0.002	-0.012*
	(0.001)	(0.006)	(0.002)	(0.007)
Public HS × 2003	-4.27E-04	0.011**	-0.003	-0.005
	(0.001)	(0.006)	(0.002)	(0.009)
Public HS × 2004	-3.68E-04	0.011*	3.26E-05	-0.002
	(0.001)	(0.007)	(0.002)	(0.011)
F-Tests†				
H_0 : Public HS × Year=0 (p-value)	0.869	0.172	0.161	0.218
Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes
High School Fixed Effects	Yes	Yes	Yes	Yes
Observations	131,436	131,436	131,436	131,436
Municipality Clusters	1,199	1,199	1,199	1,199
High School Clusters	4,119	4,119	4,119	4,119

Notes: The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. † The F-test null hypothesis is that all Public High School × Year interaction-term parameters are equal to 0. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 11: Estimating the Impact of the Affirmative Action Using Additional Years (2002-2007)

	(1)	(2)	(3)	(4)
	Medicine		Top-Five	
	App. & Adm.	Applied	App. & Adm.	Applied
ENEM Score	0.005***	0.029***	0.013***	0.039***
	(0.001)	(0.002)	(0.001)	(0.002)
Public High School	4.05E-04	-0.009*	0.001	-0.001
	(0.001)	(0.006)	(0.001)	(0.008)
Public HS times AA	0.003**	0.019***	0.010***	0.031***
	(0.001)	(0.005)	(0.002)	(0.004)
Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes
High School Fixed Effects	Yes	Yes	Yes	Yes
Observations	165,268	165,268	165,268	165,268
Municipality Clusters	1,237	1,237	1,237	1,237
High School Clusters	4,876	4,876	4,876	4,876

Notes: The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 12: Reduced Sample: 17 to 23 Years Old Only

	(1)	(2)	(3)	(4)
	Medicine		Top-Five	
	App. & Adm.	Applied	App. & Adm.	Applied
ENEM Score	0.004***	0.019***	0.012***	0.036***
	(0.001)	(0.002)	(0.001)	(0.002)
Public High School	-1.33E-05	-0.017*	-0.002	-0.019
	(0.001)	(0.010)	(0.002)	(0.017)
Public HS × AA	0.002	0.010*	0.008***	0.024***
	(0.001)	(0.006)	(0.002)	(0.007)
Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes
High School Fixed Effects	Yes	Yes	Yes	Yes
Observations	63,453	63,453	63,453	63,453
Municipality Clusters	773	773	773	773
High School Clusters	3,383	3,383	3,383	3,383

Notes: The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 13: Affirmative Action and SES Gradient – Application to Medicine

	(1)	(2)	(3)	(4)
A. 2004-2006				
ENEM Score			0.039***	0.032***
			(0.003)	(0.003)
ln(Income)	0.093***	0.084***	0.076***	0.071***
	(0.008)	(0.008)	(0.007)	(0.007)
ln(Income) × AA	-0.004	-0.009***	-0.007*	-0.012***
	(0.004)	(0.004)	(0.004)	(0.004)
Number of Observations	65,507	65,507	65,507	65,507
Number of Municipality Clusters	806	806	806	806
B. 2002-2007				
ENEM Score			0.037***	0.030***
			(0.002)	(0.002)
ln(Income)	0.100***	0.090***	0.084***	0.078***
	(0.010)	(0.010)	(0.009)	(0.009)
ln(Income) × AA	-0.009***	-0.014***	-0.013***	-0.017***
	(0.002)	(0.002)	(0.002)	(0.002)
Number of Observations	163,354	163,354	163,354	163,354
Number of Municipality Clusters	1,232	1,232	1,232	1,232
Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Age and Gender	Yes	Yes	Yes	Yes
Parental Education	No	No	No	No
Parental Occupation	No	No	No	No
Municipality Fixed Effects	No	Yes	No	Yes
High School Fixed Effects	No	No	No	No

Notes: The dependent variable is a binary variable equal to 1 if the applicant applied to Medicine at UNICAMP, 0 otherwise. AA is a binary variable equal to one if the year is after 2004, zero otherwise. Income is the parental income, measured as multiples of the minimum wage (e.g. three times the minimum wage). The ENEM score is divided by 12 so that it lies between 0 and 10. We control for age using a quartic function. Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. 'Trainees' are excluded. Cluster-robust standard errors (at the municipality level) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 14: Affirmative Action and SES Gradient – Application to Top-Five Majors

	(1)	(2)	(3)	(4)
A. 2004-2006				
ENEM Score			0.045***	0.037***
			(0.003)	(0.004)
ln(Income)	0.088***	0.077***	0.068***	0.062***
	(0.008)	(0.008)	(0.007)	(0.007)
ln(Income) × AA	-0.008**	-0.014***	-0.013***	-0.017***
	(0.004)	(0.003)	(0.004)	(0.003)
Number of Observations	65,507	65,507	65,507	65,507
Number of Municipality Clusters	806	806	806	806
B. 2002-2007				
ENEM Score			0.046***	0.037***
			(0.002)	(0.003)
ln(Income)	0.097***	0.084***	0.077***	0.069***
	(0.008)	(0.009)	(0.008)	(0.008)
ln(Income) × AA	-0.015***	-0.021***	-0.020***	-0.025***
	(0.003)	(0.002)	(0.003)	(0.002)
Number of Observations	163,354	163,354	163,354	163,354
Number of Municipality Clusters	1,232	1,232	1,232	1,232
Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Age and Gender	Yes	Yes	Yes	Yes
Parental Education	No	No	No	No
Parental Occupation	No	No	No	No
Municipality Fixed Effects	No	Yes	No	Yes
High School Fixed Effects	No	No	No	No

Notes: The dependent variable is a binary variable equal to 1 if the applicant applied to Medicine at UNICAMP, 0 otherwise. AA is a binary variable equal to one if the year is after 2004, zero otherwise. Income is the parental income, measured as multiples of the minimum wage (e.g. three times the minimum wage). The ENEM score is divided by 12 so that it lies between 0 and 10. We control for age using a quartic function. Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. 'Trainees' are excluded. Cluster-robust standard errors (at the municipality level) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 15: Affirmative Action and SES Gradient – Admission to Medicine

	(1)	(2)	(3)	(4)
A. 2004-2006				
ENEM Score			0.004***	0.004***
			(0.001)	(0.001)
ln(Income)	0.003***	0.003***	0.001***	0.001***
	(4.60E-04)	(0.001)	(3.34E-04)	(3.80E-04)
ln(Income) × AA	0.001	0.001	3.05E-04	4.17E-04
	(0.001)	(0.001)	(0.001)	(0.001)
Number of Observations	65,507	65,507	65,507	65,507
Number of Municipality Clusters	806	806	806	806
B. 2002-2007				
ENEM Score			0.005***	0.005***
			(0.001)	(0.001)
ln(Income)	0.004***	0.004***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
ln(Income) × AA	-1.45E-04	-2.72E-04	-0.001**	-0.001**
	(2.86E-04)	(2.96E-04)	(2.92E-04)	(2.99E-04)
Number of Observations	163,354	163,354	163,354	163,354
Number of Municipality Clusters	1,232	1,232	1,232	1,232
Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Age and Gender	Yes	Yes	Yes	Yes
Parental Education	No	No	No	No
Parental Occupation	No	No	No	No
Municipality Fixed Effects	No	Yes	No	Yes
High School Fixed Effects	No	No	No	No

Notes: The dependent variable is a binary variable equal to 1 if the applicant applied to and was admitted to Medicine at UNICAMP, 0 otherwise. AA is a binary variable equal to one if the year is after 2004, zero otherwise. Income is the parental income, measured as multiples of the minimum wage (e.g. three times the minimum wage). The ENEM score is divided by 12 so that it lies between 0 and 10. We control for age using a quartic function. Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. 'Trainees' are excluded. Cluster-robust standard errors (at the municipality level) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 16: Affirmative Action and SES Gradient – Admission to a Top-Five Major

	(1)	(2)	(3)	(4)
A. 2004-2006				
ENEM Score			0.011***	0.011***
			(0.001)	(0.001)
ln(Income)	0.006***	0.005***	0.001*	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
ln(Income) × AA	3.97E-04	1.58E-04	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Number of Observations	65,507	65,507	65,507	65,507
Number of Municipality Clusters	806	806	806	806
B. 2002-2007				
ENEM Score			0.012***	0.012***
			(0.001)	(0.001)
ln(Income)	0.007***	0.006***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
ln(Income) × AA	-1.40E-04	-4.03E-04	-0.001***	-0.002***
	(4.78E-04)	(0.001)	(0.001)	(0.001)
Number of Observations	163,354	163,354	163,354	163,354
Number of Municipality Clusters	1,232	1,232	1,232	1,232
Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Age and Gender	Yes	Yes	Yes	Yes
Parental Education	No	No	No	No
Parental Occupation	No	No	No	No
Municipality Fixed Effects	No	Yes	No	Yes
High School Fixed Effects	No	No	No	No

Notes: The dependent variable is a binary variable equal to 1 if the applicant applied and was admitted to one of the top-five most competitive UNICAMP majors (i.e., Medicine, Computer Engineering-daytime, Control and Automation Engineering-evening, Electrical Engineering-daytime, Electrical Engineering-evening) based on the 2003 Phase 2 cutoffs, 0 otherwise. AA is a binary variable equal to one if the year is after 2004, zero otherwise. Income is the parental income, measured as multiples of the minimum wage (e.g. three times the minimum wage). The ENEM score is divided by 12 so that it lies between 0 and 10. We control for age using a quartic function. Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. 'Trainees' are excluded. Cluster-robust standard errors (at the municipality level) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Appendix C – Additional Tables

Table A.1: Affirmative Action and Medicine UNICAMP Admission: by ENEM Quartile

	Q1	Q2	Q3	Q4
	(1)	(2)	(3)	(4)
ENEM Score	-	5.81E-05	0.002**	0.062***
	-	(6.10E-05)	(0.001)	(0.004)
Public High School	-	-7.31E-05	-0.001	-0.001
	-	(8.07E-05)	(0.001)	(0.007)
Public HS × AA	-	2.31E-04	8.25E-05	0.009*
	-	(2.39E-04)	(0.001)	(0.005)
Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes
High School Fixed Effects	Yes	Yes	Yes	Yes
Mean of Dep. Var. (2004, Public)	0.000	0.000	0.001	0.006
Observations	15,581	15,496	15,520	15,666
Municipality Clusters	402	387	402	370
High School Clusters	2,078	1,789	1,602	1,443

Notes: The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.2: Affirmative Action and Admission to Top-Five UNICAMP Majors: by ENEM Quartile

	Q1	Q2	Q3	Q4
	(1)	(2)	(3)	(4)
ENEM Score	-	0.001**	0.014***	0.101***
	-	(2.50E-04)	(0.003)	(0.007)
Public High School	-	-4.01E-04	-2.67E-04	-0.001
	-	(4.66E-04)	(0.003)	(0.010)
Public HS × AA	-	0.001*	0.002	0.033***
	-	(0.001)	(0.005)	(0.008)
Controls:				
Year Fixed Effects	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes
Municipality Fixed Effects	Yes	Yes	Yes	Yes
High School Fixed Effects	Yes	Yes	Yes	Yes
Mean of Dep. Var. (2004, Public)	0.000	0.000	.010	0.047
Observations	15,581	15,496	15,520	15,666
Municipality Clusters	402	387	402	370
High School Clusters	2,078	1,789	1,602	1,443

Notes: The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.3: Placebo Affirmative Action and Medicine UNICAMP Application (2003-2004)

	(1)	(2)	(3)	(4)	(5)	(6)
ENEM Score				0.036***	0.031***	0.027***
				(0.002)	(0.003)	(0.002)
Public High School	-0.107***	-0.103***	-0.013	-0.090***	-0.089***	-0.009
	(0.012)	(0.016)	(0.009)	(0.013)	(0.015)	(0.009)
Public HS × Placebo AA	0.009	0.010*	0.007	0.006	0.007	0.004
	(0.006)	(0.005)	(0.006)	(0.006)	(0.005)	(0.006)
Controls:						
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	No	Yes	Yes	No	Yes	Yes
High School Fixed Effects	No	No	Yes	No	No	Yes
Mean of Dep. Var. (2003, Public)	0.096	0.096	0.096	0.096	0.096	0.096
Observations	65,988	65,988	65,988	65,988	65,988	65,988
Municipality Clusters	912	912	912	912	912	912
High School Clusters	3,328	3,328	3,328	3,328	3,328	3,328

Notes: The dependent variable is a binary variable equal to 1 if the applicant applied to Medicine at UNICAMP, 0 otherwise. The Placebo AA variable takes the values of 1 if the applicant is observed in 2004, 0 otherwise. The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Standard deviations are in square brackets. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.4: Placebo Affirmative Action and Top-Five UNICAMP Application (2003-2004)

	(1)	(2)	(3)	(4)	(5)	(6)
ENEM Score				0.049***	0.043***	0.037***
				(0.002)	(0.003)	(0.002)
Public High School	-0.082***	-0.074***	-0.002	-0.058***	-0.055***	0.004
	(0.012)	(0.013)	(0.017)	(0.012)	(0.012)	(0.016)
Public HS $\times$ Placebo AA	0.012*	0.013**	0.008	0.008	0.009	0.004
	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)
Controls:						
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Personal Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Parental Education	Yes	Yes	Yes	Yes	Yes	Yes
Parental Occupation	Yes	Yes	Yes	Yes	Yes	Yes
Municipality Fixed Effects	No	Yes	Yes	No	Yes	Yes
High School Fixed Effects	No	No	Yes	No	No	Yes
Mean of Dep. Var. (2003, Public)	0.226	0.226	0.226	0.226	0.226	0.226
Observations	65,988	65,988	65,988	65,988	65,988	65,988
Municipality Clusters	912	912	912	912	912	912
High School Clusters	3,328	3,328	3,328	3,328	3,328	3,328

Notes: The dependent variable is a binary variable equal to 1 if the applicant applied to Medicine at UNICAMP, 0 otherwise. The Placebo AA variable takes the values of 1 if the applicant is observed in 2004, 0 otherwise. The ENEM score is divided by 12 so that it lies between 0 and 10. Personal characteristics consist of a quartic function of age, gender, and a previous university attendance indicator variable, while the parental education and occupation controls each consists of 16 dummy variables (8 per parent). Municipalities are the ones of the applicant's residence. Singletons (observations in municipalities or high school that we only observe once) are dropped. Standard deviations are in square brackets. Two-way cluster-robust standard errors (at the municipality and high school levels) are shown in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.