Equilibrium Effects of Higher Education Subsidies to Public Schools

Vinícius Botelho (Insper)

April 19, 2024

Abstract

This paper sets up a structural supply and demand model estimated using Brazilian data to determine how much capacity constraints in public schools affect the equilibrium price responses from counterfactual changes in subsidies to public schools (conditional) and subsidies to all schools (unconditional). We show that public school capacity constraints substantially affect their private counterparts' price response: before public degrees become capacity-constrained, subsidizing them increases private prices because students with high price sensitivity switch from private to public degrees; however, the effect is reversed when public degrees are capacity-constrained because of the impact of subsidies on admission requirements. Using the model, we estimate that a budget-neutral reform increasing the monthly tuition in public degrees by BRL 300 and providing a BRL 70 monthly unconditional scholarship increases enrollment by 1.0% and consumer surplus by BRL 0.6, even though market prices rise by 22.5%.

1 Introduction

Although there is substantial evidence favoring subsidies to higher education (Chetty et al., 2020; Dynarski et al., 2022), there is still an ongoing debate about how they should be designed. One of the open questions in subsidy design is whether subsidies should be restricted to public schools (conditional) or not (unconditional). Because private schools maximize profits, their equilibrium response to subsidy changes may differ from that of public schools. So, it may be optimal to subsidize private and public colleges differently, depending on the social utility function.

One key dimension to understanding private schools' response to subsidy changes is how they compete with public and other private schools. Hence, capacity constraints in public schools should matter to how private schools respond to subsidies. Understanding the interaction between private and public players has practical relevance because more than 95% of the countries have private and public college supply (Levy, 2018). However, this topic has not received much attention in the literature on higher education. In this paper, we hope to contribute to understanding it.

Specifically, we focus on the role of capacity constraints in public schools in the effect of conditional and unconditional subsidies on private prices. To do so, we set up and estimate a structural supply and demand model for the higher education market in Brazil. We follow standard discrete choice literature in which degrees have heterogeneous levels of attractiveness, costs, and capacity. Private degrees choose prices, and public degrees choose admission scores. Students have heterogenous price sensitivities (according to their socioeconomic status and test scores) and choose degrees in a random utility model with heterogeneous choice sets (according to their test scores and eligibility for affirmative action). Using the model, we simulate counterfactual conditional and unconditional subsidies. Subsidy changes can be positive (scholarship) or negative (tuition). The scholarship is a cash transfer to the student and may exceed the price paid by her. We decompose the equilibrium impacts of subsidy changes into two components: the direct effect, which measures the impact of subsidy changes when public degrees do not face capacity constraints, and the indirect effect, which isolates the impact of subsidy changes caused by changes in the extent to which public degrees are capacity-constrained. We ask what is the effect of conditional and unconditional subsidies on private prices; then, we estimate the direct and indirect effects.

We find that increasing unconditional subsidies reduces private prices. According to our model, a flat unconditional monthly scholarship of BRL 400[1] would reduce private prices by 6.1%. The indirect effect, alone, causes prices to fall 2.1%. Therefore, capacity constraints account for a sizable share of the overall effect. On the other hand, a flat conditional scholarship of BRL 400 would reduce private prices by less than 0.1%. In this case, the direct and indirect effects go in opposite directions. The direct effect makes prices rise 3.1%, but the indirect effect causes prices to fall by 3.2%. Therefore, the indirect effect is determining the sign of the net effect. In other words, capacity constraints in public schools are changing the direction of the effect of changes in conditional subsidies on private prices.

Section 2 describes the data and the higher education market in Brazil. We argue that Brazil offers an interesting study case for our problem because there is data on capacity utilization and admission requirements for higher education degrees. Usually, in problems similar to ours, capacity is not observed, and the process through which goods are allocated under excess demand is unknown. Fortunately, Brazil has a long tradition of admitting students based on test scores, the *vestibulares*, and most public schools admit students through a centralized admissions system that ranks students using the same high school test (*ENEM*).

Our core data come from merging the census of higher education institutions, in which we observe degree capacity, location, and whether degrees are public or private, administrative

[1] Approximately USD 80.

registries from students who took the national high school examination (*ENEM*), in which we observe students' test scores and socioeconomic status, and administrative registries of student loan contracts (*FIES*), in which we observe prices of private institutions. The data was merged in a secure room at the official agency for education statistics in Brazil (*INEP*). The dataset we use to estimate our demand function uses information from more than 3 million students. We enrich the main dataset with other administrative registries to present a broader picture of the higher education market.

The data show that capacity utilization near 100% is common in public degrees, and oversubscription is large, justifying why public schools choose admission scores in the model. The law prohibits public schools from charging tuition, justifying why public schools do not choose prices. On the other hand, most private schools display capacity utilization levels not suggestive of capacity constraints, justifying why they only choose prices in the model. We also show that private and public providers hold sizable market shares in Brazil, suggesting there is room for a strong interaction between these two types of players.

Section 3 describes the structural supply and demand model, and Section 4 describes our empirical strategy for estimating it. To estimate demand parameters, we follow Berry et al. (2004). To account for the endogeneity between prices and unobserved degree attractiveness we use a cost shifter: the average wages of higher education employees and college professors in the region where the degree is located. Section 5 shows model estimation results. We show that price elasticities are higher for poor students with low test scores, and the distance between students and degrees has a sizable effect on utility. One of the possible explanations for the fact that students with lower test scores have higher price-elasticity even after controlling for their socioeconomic status is that they face lower returns to education, reducing their willingness to pay for college.

Section 6 shows three sets of counterfactual results. In the first set of counterfactuals, we evaluate the impact of capacity constraints on the baseline equilibrium. Without capacity constraints in the public sector, private profits would fall by 9.0%, compared to the baseline levels. So, in the baseline, capacity constraints in public schools are a source of sizable market power for their private counterparts.

In the second set of counterfactuals, we evaluate conditional and unconditional subsidy changes. Because we are concerned about the enrollment of poor students, we simulate the equilibrium outcomes of counterfactual conditional and unconditional subsidies targeted to them. When subsidies are conditional, only eligible students who choose public schools receive the subsidy; when subsidies are unconditional, all eligible students in higher education receive the subsidy. Counterfactual subsidies are flat, restricting simulations to policies with smaller pass-through (Sahai, 2023). Subsidies are a cash transfer, so they are not limited by tuition. Subsidy changes can be positive (increasing the subsidy from the baseline level) or negative (reducing the subsidy from the baseline level).

When unconditional subsidies rise, private prices fall monotonically. A monthly scholarship of BRL 400 would reduce private prices by 6.1%. Two mechanisms explain this effect. The first is related to the direct effect: subsidies to private schools make price-elastic students in the outside option switch to private degrees. The second is related to the indirect effect: higher subsidies to public schools increase their admission scores, and the rise in admission scores in public schools forces low-score students to switch from public to private degrees or the outside option. Because low-score students are also more price-elastic than their peers, the rise in admission scores in public degrees raises the proportion of price-elastic students attending private, reinforcing the fall in private prices. The indirect effect, alone, reduces prices by 2.1%. Consequently, capacity constraints account for a sizable share of the overall effect.

When conditional subsidies change, the effect is non-monotonic and very close to zero. However, the small net effect hides the sizable direct and indirect effects, which move in opposite directions.

As we reduce conditional subsidies, the proportion of capacity-constrained degrees falls, and the indirect effect becomes closer to zero. Then, the direct effect dominates: lower subsidies for public degrees make price-elastic students switch from public to private degrees, reducing private prices. On the other hand, as we increase conditional subsidies, the proportion of capacity-constrained degrees rises, and the indirect effect grows. As more public degrees become capacity-constrained, more low-score students are forced to switch from public to private degrees, increasing the price elasticity of students attending the private choice, and reducing private prices. The direct effect works in the opposite direction: more subsidies to public schools would induce price-elastic students in private degrees to switch to public, increasing private prices. As the subsidy grows, the indirect effect dominates.

In these exercises, the subsidy budget changes in every simulation. Therefore, welfare considerations would require hypotheses on the effects of raising or reducing taxation to match changing subsidy expenditures. We build our third set of counterfactuals by joining this concern with the fact that private prices fall when unconditional subsidies rise. We look for a budget-neutral equilibrium in which we charge tuition in public schools and use tuition revenue to fund an unconditional subsidy. We find that charging small tuition (BRL 300 monthly)[2] and providing unconditional subsidies can be a welfare-improving reform: consumer surplus rises (even though market prices rise by 22.5%), private prices fall (by 3.1%), profits rise (by 15.2%), enrollment rises (by 1.0%).

This paper connects to several strands of the literature. The first set of papers looks at the pass-through of subsidies to private prices (Kargar and Mann, 2022; Dobbin et al., 2022; Lucca et al., 2018; Turner, 2017; Cellini and Goldin, 2014; Turner, 2012; Singell and Stone, 2007). We learn from this literature that subsidies to private schools may increase private

[2]Approximately USD 60.

prices, and we add to this literature by showing that capacity constraints in public schools account for a sizable portion of the effect of unconditional subsidies on private prices. We also show that the composition effect matters empirically, providing a case in which increases in unconditional subsidies reduce prices.

The second set of papers looks at the side effects of public school subsidies. Peltzman (1973), assuming that higher education is indivisible and consumed only once, showed that increasing subsidies to public schools can reduce the investment in higher education when public degrees are low-quality. Specifically, the aggregate effects of the subsidy would depend on the difference between the number of students who increase their investment level, switching from the outside option to low-quality public degrees, and the number of students who reduce their investment level, switching from high-quality private to low-quality public degrees. Many papers (Cohodes and Goodman, 2014; Cellini, 2009; Long, 2004) have shown that the concern is empirically relevant. We add to this literature by showing another concern related to conditional subsidies: when public degrees are capacity-constrained, increasing subsidies reduces the choice set of low-score students, concentrating subsidies on students with lower price elasticity, and reducing enrollment and consumer surplus.

The third set of papers looks at the importance of subsidies to higher education (Chetty et al., 2020; Dynarski et al., 2022). We learn from this literature that subsidies are important to bring higher education consumption closer to the socially optimum level and that market failures leading to underinvestment in higher education affect more students from poorer backgrounds. We add to this literature by showing the limitations of price subsidies to increase the enrollment of students from poor socioeconomic backgrounds, even when subsidies are targeted at public schools, which do not pass the subsidies they receive through prices.

The fourth set of papers looks at the assignment of students to capacity-constrained schools (Gale and Shapley, 1962; Balinkski and Sönmez, 1999; Agarwal and Somaini, 2018; Bó and Hakimov, 2019; Fack et al., 2019). We learn from this literature how the matching process of students and schools depends on capacity constraints and school preferences; we also learn the main properties of student-school assignment mechanisms. We use its results to simplify how we model the admission process in public schools, preserving the main properties from the actual mechanism that assigns students and schools. Because it simplifies the application-admission game, our model offers a tractable way to make capacity-constrained public schools and profit-maximizing private schools interact.

The fifth set of papers looks at the public provision of goods. After all, this problem is relevant for other markets as well, such as primary and secondary schools (Dinerstein and Smith, 2021), pharmacies (Atal et al., 2022), food distribution (Jiménez-Hernández and Seira, 2022; Handbury and Moshary, 2021; Banerjee et al., 2019), retail stores (Busso and Galiani, 2019), healthcare (Curto et al., 2019), and financial services (Fonseca and Matray, 2022).

Furthermore, Epple et al. (2017) provide an equilibrium model to analyze the interaction of private and public providers in higher education when conditional and unconditional subsidies change. However, their model does not incorporate for-profit schools (private schools are non-profit organizations maximizing quality) and they do not discuss the equilibrium implications of capacity constraints in public institutions. Even though they acknowledge schools face capacity constraints, they limit the discussion to how capacity constraints would affect the players' first-order condition. Moreover, differences in admission policies are driven by peer effects and preferences, not capacity constraints. So, we believe our paper adds to the debate by highlighting the equilibrium impacts of capacity constraints in public degrees.

2 Higher Education in Brazil

In this section, we discuss the institutional framework of the higher education market in Brazil, present our data, and discuss how the data informs our model choices regarding the problem of private and public schools. We also discuss descriptive evidence regarding the distribution of subsidies and enrollment between students from high and low socioeconomic statuses (SES).

2.1 Institutional Framework

Schools supply degrees, to which students apply. When choosing degrees, students simultaneously choose a major, a school campus, and the time of day classes will be held, such as morning, afternoon, or evening.

The market has private and public schools, which can be federal, state, or municipal. By law, public schools do not charge tuition. Unlike many other countries, public education in Brazil has high quality: according to the U.S. News Ranking for 2022-2023, there are 16 Brazilian public universities among the best 30 universities in Latin America. The combination of high quality and no tuition makes the public option attractive to all students, raising concerns about whether subsidies effectively increase student enrollment in higher education.

Yearly, Brazil holds a national high school examination called ENEM (*Exame Nacional do Ensino Médio*). It is a non-mandatory test students can take in high school or after graduation, commonly used for admission in public schools. The test is divided into four subjects (Mathematics, Language, Natural Sciences, and Social Sciences) and an essay.

Most public schools admit students using a centralized system, SISU (*Sistema de Seleção Unificada*), which ranks students according to ENEM scores, reported preferences, and affirmative action eligibility. SISU is similar to the Iterative Deferred Acceptance Mechanism, extensively described in Bó and Hakimov (2019).[3] It opens twice yearly, and admissions follow a cutoff score mechanism. Only test scores and affirmative action eligibility are taken into account in the admission process.

2.2 Data and Descriptive Statistics

Our core student and degree data come from three sources.

The first dataset is the 2015 ENEM (INEP, 2015), which is the administrative registry with the data from all test-takers in the 2015 edition of ENEM. This data provide us test scores, socieconomic status, affirmative action eligibility, and location of all test-takers. We divide students into groups according to the percentile of their test scores, the quintile of their socioeconomic status, their affirmative action eligibility, and the combined statistical area in which they took ENEM.[4] Socioeconomic status is defined as family *per capita* income.[5] We only kept in the sample ENEM test takers in 2015 who already had a high school diploma or were graduating that year.

The second dataset is the 2016 Higher Education Census (INEP, 2016), which is an administrative registry containing data from an annual survey on all higher education institutions in Brazil and their students. The data provide us with students' degree choices (the degrees in which these students enrolled as first-year students in 2016), degree capacity, degree ownership status (public or private), and which degrees are free. We dropped information on distance learning degrees and other education modalities beyond high school, such as technical degrees.

The third dataset is the 2016 FIES dataset,[6] which is an administrative registry containing data from all student loan contracts. The data provide us private prices.

These datasets were accessed, merged, and enriched in a secure room at INEP (*Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira*), the official agency for education statistics in Brazil. To reduce the number of choices in the model, we assume that degrees are the combination of major, school, and municipality in which classes are held. This means aggregating degrees in the same municipality, school, and major, regardless of the specific campus or timing in which classes take place. Noteworthy, graduation diplomas usually clearly specify only the school and major in which students graduate.[7] Descriptive statistics of student and degree data are shown in Tables (1) and (2).

To measure higher education costs, we use RAIS (*Relação Anual de Informações Soci-*

[3]In contrast to the standard Iterative Deferred Acceptance Mechanism, SISU allows students to make two applications (informing their first and second choices), is designed for only four rounds, and allows students to change their choices anytime.

[4]The income levels and test scores that correspond to each group can be found in the Appendix.

[5]Details about how the family *per capita* income and affirmative action eligibility were calculated can be found in the Appendix.

[6]FIES is Brazil's largest student loan program.

[7]For further reference, see Brazilian Ministry of Education (2018).

Table 1: Descriptive statistics: students

	Mean	Std.Dev.
College enrollment	0.39	0.49
Not eligible to affirmative action	0.34	0.47
Number of observations	5,480,902	

Source: Higher Education Census (INEP, 2016) and ENEM (INEP, 2015).

Table 2: Descriptive statistics: degrees

	Mean	Std.Dev.
Prices	489.15	584.830
Proportion of public	0.27	0.445
Capacity utilization	0.51	0.326
Number of observations	21,232	

Source: Higher Education Census (INEP, 2016) and FIES (MEC, 2016).

ais), an administrative registry containing information on the labor contracts of all companies in Brazil. It contains data on the industry, wages, occupation, and hours contracted of all formal labor market contracts in 2016. We restrict the sample to employees in higher education institutions who signed 20, 40, or 44-hour contracts, which are the most common workweeks in the data and represent 65% of the contracts in higher education institutions. In Table (3) we show average wages and hours contracted of all employees and college professors, by department (as officially stated in the labor contract).[8]

2.3 Markets

As shown in Table (4), most schools are private. Among public schools, federal and state schools are the most common types.

Among students who go to college, 77% go to private degrees, and 23% go to public degrees, as shown in Table (5). Private schools have fewer students per school than their public counterparts.

We assume a market is a state, reducing the number of degree choices from 21,232 to 786 (on average). This restriction makes choice sets more realistic and reduces the computational burden of estimating and simulating the model. Moreover, only 3.3% of college students choose degrees in other states, meaning that the cost of imposing this restriction does not seem to be large. We also restrict the size of the market to 500,000 students, for computational constraints. So, we exclude the two biggest states (São Paulo and Minas Gerais), which correspond to 30.8% of the sample.

[8] Mathematics and Statistics, Architecture and Engineering, Biology and Health, Pedagogy, Language and Literature, Humanities, and Economics, Business, and Accounting.

Table 3: Descriptive statistics: labor contracts

Variable	Mean	Std.Dev	N
Wages (Mathematics and Statistics)	9,171.99	5,712.24	7,333
Wages (Architeture and Engineering)	10,718.14	6,454.42	10,798
Wages (Biology and Health)	9,497.45	6,217.29	21,587
Wages (Pedagogy)	7,977.41	6,133.14	52,728
Wages (Language and Literature)	8,639.11	5,188.05	4,989
Wages (Humanities)	8,752.22	6,325.41	16,186
Wages (Economics, Business, and Accounting)	8,138.82	6,629.28	10,810
Wages (all contracts)	5,957.00	5,615.16	552,083
Hours contracted (Mathematics and Statistics)	38.54	6.17	7,333
Hours contracted (Architeture and Engineering)	38.04	6.40	10,798
Hours contracted (Biology and Health)	36.66	8.24	21,587
Hours contracted (Pedagogy)	37.12	8.58	52,728
Hours contracted (Language and Literature)	38.08	6.64	4,989
Hours contracted (Humanities)	37.47	7.63	16,186
Hours contracted (Economics, Business, and Accounting)	35.67	9.74	10,810
Hours contracted (all contracts)	39.79	6.07	552,083

Source: RAIS (2016). **Notes:** wages and hours contracted of professors in each of the seven listed subjects, as well as wages and hours contracted of all employees. Prices in BRL.

Table 4: Higher education supply (2016)

	Private	Public (federal)	Public (state)	Public (municipal)
Number of Degrees	15,476	3,747	1,823	186
Number of Colleges	2,005	107	121	43
Number of Majors	79	85	74	37

Source: Higher Education Census (INEP, 2016).

The market is the set of all ENEM test takers. We drop degrees with less than 5 enrollments, or with missing data. We also drop students who chose degrees in other states, or whose data is missing.

When all restrictions are taken into account, the number of students is reduced from 5,480,902 to 3,472,119 (-36.7%) and the number of degrees is reduced from 21,232 to 8,960. Most of the reduction in the number of students comes from dropping the two largest states (30.8%) and the students who chose degrees in other markets (3.3%), meaning that the other sample restrictions have little impact.

As shown in Table (6), the degrees in the sample are similar to those in the population: even though most differences are statistically significant, they are not sizable.

Even though degrees in the population and in the sample are very close, the sample has a higher proportion of low-income and low-score students, as shown in Table (7).

Table 5: Market size (2016)

	Students	Market Share
Private	1,657,868	0.77
Public (federal)	327,818	0.15
Public (state)	144,673	0.07
Public (municipal)	13,359	0.01

Source: Higher Education Census (INEP, 2016).

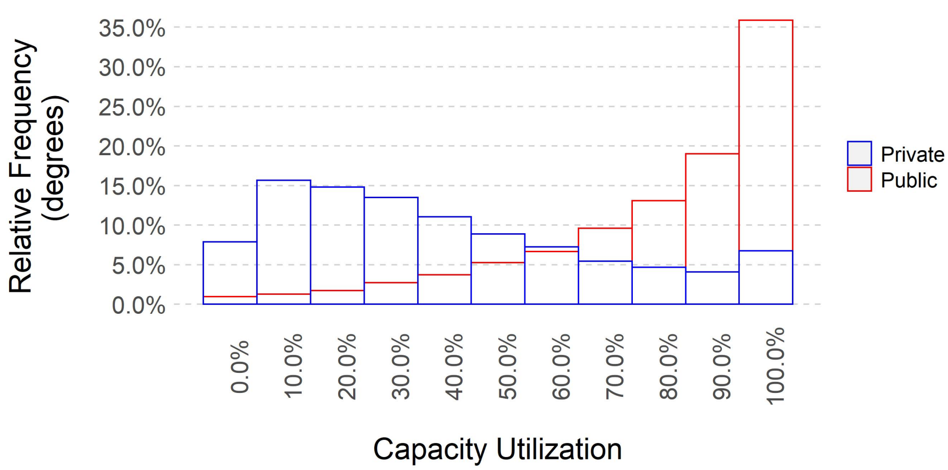
Table 6: Comparison of means: population and sample (degrees)

	Population	Sample	p-value
Prices	489.15	495.93	0.39
Proportion of public	0.27	0.34	0.00
Capacity utilization	0.51	0.56	0.00
Number of observations	21,232	8,960	

2.4 Capacity Utilization and Oversubscription in Public and Private Schools

Public schools are competitive and capacity-constrained, but private degrees display widespread idle capacity, as shown in Figure (1).

Figure 1: Capacity utilization: public vs. private (2016)



Source: Higher Education Census (INEP, 2016). **Notes:** Proportion of degrees in each capacity utilization group by ownership (public or private) status. Capacity utilization is enrollment divided by the number of reported seats.

Furthermore, the enrollment, capacity, and application data in the higher education census suggest that there are approximately 15 candidates per seat in public schools but only 1 candidate per seat in private schools.

Table 7: Comparison of means: population and sample (students)

	Population	Sample	p-value
Proportion of students in the 1st SES Quintile	0.20	0.26	0.00
Proportion of students in the 2nd SES Quintile	0.20	0.22	0.00
Proportion of students in the 3rd SES Quintile	0.20	0.19	0.00
Proportion of students in the 4th SES Quintile	0.20	0.15	0.00
Proportion of students in the 5th SES Quintile	0.20	0.18	0.00
Proportion of students in the 1st Score Quintile	0.20	0.24	0.00
Proportion of students in the 2nd Score Quintile	0.20	0.21	0.00
Proportion of students in the 3rd Score Quintile	0.20	0.18	0.00
Proportion of students in the 4th Score Quintile	0.20	0.19	0.00
Proportion of students in the 5th Score Quintile	0.20	0.18	0.00
Proportion of students not eligible to affirmative action	0.34	0.29	0.00
Proportion of students enrolling in college	0.39	0.29	0.00
Number of observations	5,480,902	3,472,119	

In other words, capacity utilization and oversubscription data suggest that public schools are selective, but private schools are not. So, we assume public schools have an active admission process, and private schools admit all students. Therefore, determining the admission process for public schools implies that we know the admission process for the entire market because private schools are not selective.[9]

2.5 Admission in Public Schools

The data suggest that most admissions in public schools are based exclusively on test scores. The higher education census data show that approximately 59% of the students used ENEM scores in their admission process, and less than 8% of all admissions are not based on test scores (ENEM or *vestibulares*). Noteworthy, ENEM is a good proxy for the admission criteria of schools that apply their tests: Estevan et al. (2018) have studied one of the most prestigious Brazilian public universities and shown that scores from *vestibulares* correlate with those from ENEM. Therefore, we use ENEM scores as a proxy for the admission criteria in all schools.

2.6 SES Gap, Subsidy Incidence, and Affirmative Action

Public schools are capacity-constrained, selective, and high-quality. Consequently, they are very attractive, and their admission requirements prevent them from enrolling poor students, who typically have lower scores.

To address this concern, since 2015 the law has required federal public degrees to reserve 50% of their seats to students from public high schools, with sub quotas for student sub-

[9]Previous papers adopted a similar hypothesis, such as Dobbin et al. (2022).

groups according to race, income, and disabilities.[10] In fact, public high schools in Brazil are low-quality and typically consumed by poorer students. By focusing on poorer students, subsidies are better targeted to credit-constrained families.

Nevertheless, there is still a sizable gap in the enrollment of high-SES and low-SES students.

To document the gap, we estimate the model in Equation (1) using 2016 data (after the affirmative action reform). In the reduced form, we explain the enrollment of student i in higher education ($Enrollment_i$) by test score dummies ($\tau_{SCORE(i)}$), SES dummies ($\tau_{SES(i)}$), location dummies ($\tau_{LOC(i)}$), and an error term (ϵ_i). We define the SES enrollment gap as the difference between the highest (5th) and lowest (1st) SES dummy estimates. Estimation results are shown in Table (8) and suggest a sizable gap between the two groups, of 19.4 percentage points.

$$Enrollment_i = \tau_{SCORE(i)} + \tau_{SES(i)} + \tau_{LOC(i)} + \epsilon_i \tag{1}$$

Table 8: SES Enrollment Gap

	Model 1	Model 2	Model 3
High-Low SES Gap	0.389	0.253	0.194
Score FE	No	Yes	Yes
Location FE	No	No	Yes

The previous exercise shows a SES-gap for overall enrollment, including private and public schools. In Figure (2) we plot the incidence of subsidies in public higher education by SES group. Results suggest that subsidies reinforce the inequalities the enrollment estimates have shown. Therefore, from a policy perspective, a stronger focus on low-SES students could improve equity in the market.

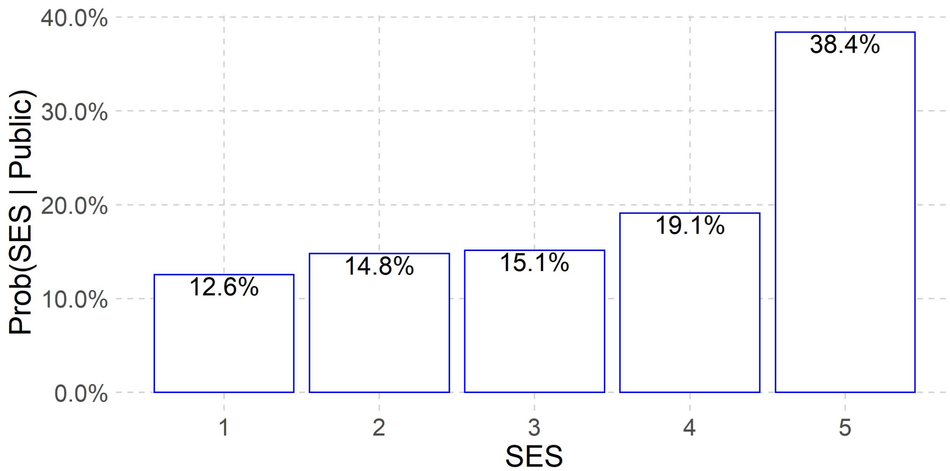
3 Model

The national market consists of a set of regional markets ($t \in T$) in which colleges ($j \in J_t$) supply degrees ($d \in D_t$) to students ($i \in I_t$). To simplify the exposition, we drop the market index and not present the model with the counterfactual subsidy that we introduce in the section that discusses the policy counterfactuals.

Schools can be private or public. Private schools do not face capacity constraints and admit all students. They hold a portfolio of degrees and choose the prices of all degrees simultaneously, maximizing profits. On the other hand, public schools do not charge tuition, face capacity constraints, and admit students using cutoff scores. Federal public schools have an affirmative action policy, so they have one cutoff for open seats and another for reserved

[10]Seats not reserved by affirmative action are called open seats, and the others are called reserved seats.

Figure 2: Subsidy incidence: public schools (2016)



Source: Higher Education Census (INEP, 2016) and ENEM (INEP, 2015).

seats. State and municipal schools do not follow affirmative action and have only one cutoff score. Cutoffs are adjusted to make demand equal to the number of available seats.

Students are high school graduates characterized by their socioeconomic status quintile (x_i^{SES}), test scores percentile (x_i^{SCORE}), their distance to all degrees in the market ($\{d_{id}\}_{\forall d}$), and whether they are eligible for reserved seats ($x_i^{AA} \in \{0, 1\}$). When $x_i^{AA} = 1$, student i is eligible for reserved seats; when $x_i^{AA} = 0$, she is not. Students choose degrees in a random utility model with heterogeneous choice sets: the choice set for each student depends on her test scores and eligibility for affirmative action. When choosing degrees, students observe the prices and cutoffs of all options available in their market.

The equilibrium is a Nash equilibrium on prices and admission scores.

3.1 Preferences

Individual's i indirect latent utility from choosing degree d (u_{id}) is shown in Equation (2). It depends on students' characteristics (X_i),[11] degree prices (p_d), unobserved degree attractiveness (ξ_d), demand parameters (θ), and an extreme value type I random variable that represents a taste shock, i.i.d. across students and degrees (ϵ_{id}).

Price sensitivity depends on SES (α_{SES}^j) and test scores (α_{SCORE}). The specification aims to capture the effect of ability (test scores) and credit constraints (SES status) on price elasticities. Students with lower SES will be more likely to be credit-constrained, so

[11]The set of individual characteristics is represented by $X_i : (x_i^{SES}, x_i^{SCORE}, x_i^{AA}, \{d_{id}\}_{\forall d})$. $x_i^{SES} \in \{1, 2, 3, 4, 5\}$, $x_i^{SCORE} \in \{1, ...100\}$, $x_i^{AA} \in \{0, 1\}$, and $\{d_{id}\}_{\forall d}$ is finite because distances are calculated between combined statistical areas. Because we restrict the market to a state, distances between students and all degrees available typically assume less than 10 values for each student. Therefore, X_i is discrete.

we expect they are more price-elastic. On the other hand, students with higher test scores will be more likely to experience higher returns from higher education, so we expect that they are less price-elastic.[12]

Students demand higher education by choosing one of the degrees available or the outside option (not going to college, indexed by $d = 0$). We assume that $p_0 = 0$, $d_{i0} = 0$, and $\xi_0 = 0$, so the utility of the outside option is given by $u_{i0} = \epsilon_{i0}$.

$$
\begin{aligned}
u_{id} &= V_d(X_i, p_d; \xi_d, \theta) + \epsilon_{id} \\
V_d(X_i, p_d; \xi_d, \theta) &= \left[-\sum_{j=1}^{5} \alpha^j_{SES} \mathbb{1}_{(x_i^{SES}=j)} - \alpha_{SCORE} x_i^{SCORE} \right] p_d - \beta d_{id} + \xi_d \qquad (2) \\
\theta &= (\{\alpha^j_{SES}\}_{j=1}^5, \alpha_{SCORE}, \beta)
\end{aligned}
$$

3.2 Choice sets

We now discuss students' choice sets.

There is a large literature about application-admission mechanisms (Gale and Shapley, 1962; Balinski and Sönmez, 1999; Agarwal and Somaini, 2018; Bó and Hakimov, 2019). Usually, these papers focus on evaluating algorithms that match students and schools. For example, the algorithm starts with students stating their preferences and colleges making initial admission decisions based on a priority index. Schools are limited by capacity and only one student is assigned per school. Because of capacity constraints, many students are not admitted to their first option, so the mechanisms' rules determine how colleges' and students' decisions are updated and the stop criteria. After some steps, the mechanism usually converges and admission decisions are fully determined.

In this paper, we are not discussing mechanisms that make the admission process converge; we assume that admissions are based on cutoff scores (λ_d), chosen by public schools, and we are looking for the equilibrium cutoff scores that maximize the utility of public schools given the prices of private schools and students' decisions. Degrees with affirmative action policy choose two cutoffs ($\lambda_d = (\lambda_d^0, \lambda_d^1)$)): λ_d^0 for open seats and λ_d^1 for reserved seats. Degrees without affirmative action policy choose only the cutoff for open seats and $\lambda_d^1 = \lambda_d^0$.[13]

To implement the admissions policy, degrees choose a function for admissions $\mu_d : \mathbb{X} \rightarrow \{0, 1\}$ that assumes value 0 if the student is not admitted by degree d and 1 otherwise. Admissions are based only on observable characteristics, and cutoffs depend on affirmative action status: admission score for open seats is λ_d^0, and for reserved seats is λ_d^1. To simplify

[12] The relationship between price sensitivity and test scores is linear to avoid identification issues that could arise from the fact that test scores are also being used to determine choice sets (Fack et al., 2019).

[13] Because $x_i^{SCORE} \in \{1, 2, ..., 100\}$, $\lambda_d^0 \in \{1, 2, ..., 100\}$ and $\lambda_d^1 \in \{1, 2, ..., 100\}$.

the notation, let us define $\lambda_d = (\lambda_d^0, \lambda_d^1)$. Therefore, the function for admissions follows Equation (3).

$$\mu_d(X_i) = \mu(X_i; \lambda_d) = \mathbb{1}_{(x_i^{SCORE} \geq \lambda_d^{x_i^{AA}})} \tag{3}$$

Therefore, the choice set of student i is given by Equation (4).

$$C(X_i) = \{d : \mu_d(X_i) = 1\} \tag{4}$$

3.3 Demand

Given admission scores and prices, observed by the students, the probability that degree d is chosen by student i (s_d) follows Equation (5).[14] [15]

$$s_d(X_i, p_d, p_{-d}, \lambda_d, \lambda_{-d}; \xi, \theta) = \frac{\mu(X_i; \lambda_d) exp(V_d(X_i, p_d; \xi_d, \theta))}{1 + \sum_k \mu(X_i; \lambda_k) exp(V_k(X_i, p_k; \xi_k, \theta))} \tag{5}$$

Defining $M(X_i)$ as the proportion of students with characteristics X_i, market shares (S_d) follow Equation (6).

$$S_d(p_d, p_{-d}, \lambda_d, \lambda_{-d}; \xi, \theta, M) = \sum_{X_i} M(X_i) s_d(X_i, p_d, p_{-d}, \lambda_d, \lambda_{-d}; \xi, \theta) \tag{6}$$

3.4 Public Supply

Following Balinkski and Sönmez (1999), we assume that public schools prefer students with higher test scores. In other words, comparing any two students (i and j), public degree preferences imply that schools prefer i to j if, and only if, $x_i^{SCORE} > x_j^{SCORE}$. We also assume that public degrees prefer any student to no student. Therefore, admission cutoffs are the lowest scores that satisfy two criteria. First, they must satisfy the capacity constraint: enrollment must not exceed degree capacity (K_d). Second, a fixed proportion of seats (Ψ_d) must be reserved for students eligible for affirmative action.

Following the standard in the literature that discusses mechanisms for assigning students and degrees, we write the model assuming that cutoff scores are calculated using the realized utility of applying students (including ϵ_{id}). This is important because capacity constraints

[14] The formulation is equivalent to heterogeneous and exogenous choice sets, following Fack et al. (2019).

[15] $\xi = \{\xi_d\}_{\forall d}$.

limit the realized (not the expected) number of enrolled students. However, when we simulate the model, we assume the market is large enough to ensure that each degree's realized number of students converges to its expected value.

In our model, we show that the only information needed for public degrees to choose the optimal cutoff scores is the set of their applicants' characteristics. Therefore, it is the only information public schools require to make their optimal admission score choices. Noteworthy, we assume that students know the cutoffs of all degrees when making the application decisions, so our model does not specify a mechanism that assigns students and degrees and students make their choices following a standard discrete choice model with heterogeneous choice sets. In other words, students only apply to the degrees they enroll in, and, consequently, we use the terms application and enrollment interchangeably.

3.4.1 Applications

Student choices are defined by n_{id}, shown in Equation (7). n_{id} equals 1 when student i chooses degree d and 0 otherwise. Because choices are linked to realized utility levels, n_{id} is also a function of ϵ_i. Students can only choose schools that would accept them.

$$n_{id}(\lambda_d, p_d, X_i, \epsilon_i; \lambda_{-d}, p_{-d}, \xi, \theta) = \mu(X_i; \lambda_d)\mathbb{1}(u_{id} > u_{iw}\,\forall w \,|\, \mu(X_i; \lambda_w) = 1) \tag{7}$$

Cutoff scores for degree d depend on the number and characteristics of students choosing it. We define the set of enrollments in degree d by A_d, the set of enrollments in degree d from students not eligible for affirmative action by A_d^0, and the set of enrollments in degree d from students eligible for affirmative action by A_d^1. The definitions are shown in Equation (8). The fact that n_{id} depends on ϵ_i implies that enrollment sets depend on a demographic variable, $\mathbb{M} = (X_i, \epsilon_i)_{\forall i}$.

$$\begin{aligned} A_d(\lambda_d, p_d; \lambda_{-d}, p_{-d}, \xi, \theta, \mathbb{M}) &= \{X_i : n_{id}(\lambda_d, p_d, X_i, \epsilon_i; \lambda_{-d}, p_{-d}, \xi, \theta) = 1\} \\ A_d^0(\lambda_d, p_d; \lambda_{-d}, p_{-d}, \xi, \theta, \mathbb{M}) &= \{X_i : X_i \in A_d \wedge x_i^{AA} = 0\} \\ A_d^1(\lambda_d, p_d; \lambda_{-d}, p_{-d}, \xi, \theta, \mathbb{M}) &= \{X_i : X_i \in A_d \wedge x_i^{AA} = 1\} \end{aligned} \tag{8}$$

Using the definitions from Equation (8), we calculate enrollment and enrollment by students eligible for affirmative action: N_d and N_d^1, respectively, as shown in Equation (9).

$$
\begin{aligned}
N_d(\lambda_d, p_d; \lambda_{-d}, p_{-d}, \xi, \theta, \mathbb{M}) &= \sum_{a \in A_d} \mathbb{1} \\
N_d^1(\lambda_d, p_d; \lambda_{-d}, p_{-d}, \xi, \theta, \mathbb{M}) &= \sum_{a \in A_d^1} \mathbb{1}
\end{aligned}
\tag{9}
$$

The capacity restriction requires that the number of enrolled students (N_d) is smaller than the number of seats (i.e., capacity, K_d). It is defined in Equation (10).

$$
N_d(\lambda_d, p_d; \lambda_{-d}, p_{-d}, \xi, \theta, \mathbb{M}) \leq K_d \tag{10}
$$

The affirmative action objective requires that all degrees reserve a fraction (Ψ_d) of their seats (K_d) for affirmative action students. It is defined in Equation (11), where ψ_d is the proportion of students eligible for affirmative action enrolled in degree d. Noteworthy, when degrees do not have to follow an affirmative action policy, $\Psi_d = 0$ and Equation (11) is satisfied for all possible parameter values. In this case, the only criteria for admitting students are the schools' preferences (given by test scores).

$$
\Psi_d \leq \psi_d(\lambda_d, p_d; \lambda_{-d}, p_{-d}, \xi, \theta, \mathbb{M}, K_d) = \frac{N_d^1(\lambda_d, p_d; \lambda_{-d}, p_{-d}, \xi, \theta, \mathbb{M})}{K_d} \tag{11}
$$

3.4.2 Cutoff scores

The higher the cutoff scores, the fewer students enroll. Therefore, we define feasible cutoffs as those for which enrollment is lower than capacity. Because public schools prefer any student to no student, their preferences imply they choose the lowest feasible cutoffs, to maximize the number of enrolling students.

Consequently, when capacity is not constrained degrees admit all students. Because the minimum score is 1 (equivalent to the first percentile of the test score distribution), degrees not capacity constrained make $\lambda_d^0 = \lambda_d^1 = \min\{x_i^{SCORE}\} = 1$.

When capacity is constrained, degrees do not admit all students, and admission scores depend on affirmative action policy. A useful way to explain the cutoff scores in each case is to discuss the lowest score a new student must have to be admitted.

When capacity is constrained and $\Psi_d = 0$, the fact that public schools prefer students with higher test scores makes scores the only admission criterion. Therefore, they make $\lambda_d^0 = \lambda_d^1$. In this case, an applying student would only be admitted if she had higher scores than the lowest score among the admitted students. After all, schools could refuse admission to the lowest-ranking student and admit her, increasing their utility. Therefore:

$\lambda_d^0 = \lambda_d^1 = \min\{x_i^{SCORE} : X_i \in A_d\}$.

When capacity is constrained and $\Psi_d > 0$, admission is based on test scores and affirmative action eligibility. First, let us consider the cutoff for reserved seats. We assume that affirmative action policy requires that an applying student eligible for affirmative action always be admitted to a degree that is capacity-constrained and has a proportion of enrolled students eligible for affirmative action (ψ_d) lower than the objective (Ψ_d). Therefore, the cutoff score for reserved seats when $\psi_d < \Psi_d$ is the lowest score possible. In other words: $\lambda_d^1 = \min\{x_i^{SCORE}\} = 1$. On the other hand, if the proportion of enrolled students eligible for affirmative action (ψ_d) is equal to or higher than the objective (Ψ_d), an applying student eligible for affirmative action must have scores higher than the lowest-ranking enrolled student to be admitted. After all, when the affirmative action objective is achieved ($\psi_d \geq \Psi_d$), the admission of another student eligible for affirmative action would still guarantee that $\psi_d \geq \Psi_d$, so schools choose students based only on test scores (their preferences). In this case, schools could deny admission to the lowest-ranking enrolled student (regardless of her affirmative action status) and grant admission to the new candidate, with higher scores, increasing its utility. Therefore, the cutoff score in this situation is $\lambda_d^1 = \min\{x_i^{SCORE} : X_i \in A_d\}$.

Now, let us consider the cutoff for open seats when capacity is constrained and $\Psi_d > 0$. We assume that when the current proportion of enrolled students eligible for affirmative action is equal to or lower than the requirement ($\psi_d \leq \Psi_d$), an applying student not eligible for affirmative action cannot take the seat of a student eligible for affirmative action, because it would bring the degree farther from the objective. So, to be admitted, she has to have scored higher than the lowest-ranking student among those not eligible for affirmative action. In other words: $\lambda_d^0 = \min\{x_i^{SCORE} : X_i \in A_d^0\}$. However, when the proportion of enrolled students eligible for affirmative action is higher than the requirement ($\psi_d > \Psi_d$), a student not eligible for affirmative action can take the seat of a student eligible for affirmative action without violating the affirmative action objective. So, in this case, to be admitted an applying student must have scored higher than the lowest-ranking student in the class: $\lambda_d^0 = \min\{x_i^{SCORE} : X_i \in A_d\}$. Noteworthy, in this case $\lambda_d^0 = \lambda_d^1$.

Table (9) summarizes the discussion about cutoff scores in each scenario. Noteworthy, schools only need to observe the applications they receive (A_d) to determine their optimal cutoffs in every scenario, and we assume they do.

Because private schools admit all students, the choice set of student i combines all private degrees available in the market with all public degrees that would admit her. So, to model students' choice sets we only need to model how public schools admit students.

Table 9: Determining cutoff scores

N_d	Ψ_d	ψ_d	λ_d^0	λ_d^1
$< K_d$			$min(x_i^{SCORE})$	$min(x_i^{SCORE})$
$= K_d$	$= 0$		$min(x_i^{SCORE} : X_i \in A_d)$	$min(x_i^{SCORE} : X_i \in A_d)$
$= K_d$	> 0	$> \Psi_d$	$min(x_i^{SCORE} : X_i \in A_d)$	$min(x_i^{SCORE} : X_i \in A_d)$
$= K_d$	> 0	$= \Psi_d$	$min(x_i^{SCORE} : X_i \in A_d^0)$	$min(x_i^{SCORE} : X_i \in A_d)$
$= K_d$	> 0	$< \Psi_d$	$min(x_i^{SCORE} : X_i \in A_d^0)$	$min(x_i^{SCORE})$

3.5 Private Supply

We assume private schools admit all students because they do not have capacity constraints: $\mu(X_i; \lambda_d) = 1 \, \forall X_i$. Marginal costs ($c_d$) are fixed, and the profit maximization problem is static, as shown in (12). Private schools choose the prices of all their degrees (D_j) simultaneously, maximizing profits.

$$\max_{\{p_d\}_{d \in D_j}} \sum_{k \in D_j} S_k(p_k, p_{-k}, \lambda_k, \lambda_{-k}; \xi, \theta, M)(p_k - c_k) \tag{12}$$

The first-order condition that arises from the profit-maximizing problem in (12) is given by Equation (13). We assume that private schools observe and take as given the admission scores and prices of other schools. Although they observe cutoff scores, private schools do not observe the application to public degrees, and choose prices maximizing expected profits calculated using the demand curve given by Equation (6). So, the behavior of private schools follows the standard procedure in the literature about price setting under monopolistic competition with heterogeneous goods. Therefore, private schools choose prices before students choose private degrees, but at the same time public schools choose admission scores, and students apply to public schools.

$$S_d(p_d, p_{-d}, \lambda_d, \lambda_{-d}; \xi, \theta, M) + \sum_{k \in D_j} (p_k - c_k) \frac{\partial S_k(p_k, p_{-k}, \lambda_k, \lambda_{-k}; \xi, \theta, M)}{\partial p_d} = 0 \tag{13}$$

3.6 Equilibrium

The equilibrium is a pure strategy Nash equilibrium on prices and admission scores. The equilibrium is a vector of prices ($\mathbb{P} = \{p_d\}_{\forall d}$) and cutoff scores ($\Lambda = \{\lambda_d\}_{\forall d}$) in which students maximize utility, as shown in (2), private schools maximize profits, as shown in (12), and public schools maximize utility taking into account capacity constraints and affirmative action requirements, choosing the cutoff scores as shown in Table (9).

The market timing is divided into three stages: first, ϵ_{id} is realized; second, private schools choose p_d, public schools choose λ_d, and students enroll in public schools, simultaneously; third, students enroll in private schools.

4 Empirical Strategy

In this section we show the strategy to estimate cutoff scores (λ_d), demand parameters (θ and ξ_d), and marginal costs (c_d).

4.1 Capacity and cutoff scores

We use the sample analog of the formulas in Table (9) to estimate cutoff scores. To do so, we must observe the sample analog of the set of the characteristics of students enrolled in each degree (A_d), whether the degree is capacity-constrained ($K_d = N_d$) or not ($K_d > N_d$), total enrollment (N_d), the proportion of enrolled students eligible for affirmative action (ψ_d), and the required proportion of students eligible for affirmative action (Ψ_d).

We observe A_d, N_d, and ψ_d. For state and municipal schools, we assume there are no reserved seats ($\Psi_d = 0.0$), but 50% of seats are reserved in federal schools ($\Psi_d = 0.5$), as stated by the affirmative action law (that imposes the restriction only to federal degrees). We also observe the number of available seats in each degree (Q_d).

We assume the existence of market frictions that allow degrees to be capacity-constrained when their capacity utilization level is below 100%. So, we now discuss the relationship between Q_d (reported capacity) and K_d (actual capacity).

We assume that a degree is capacity-constrained when there is excess demand. It is possible that capacity utilization is below 100% and there is excess demand if the mechanism matching students and degrees is not efficient enough to guarantee that all students who would like to choose a degree to which they would be admitted do enroll.

This discussion is important because oversubscription and capacity utilization data suggest different levels of capacity constraints in public schools. Even though many public degrees display capacity utilization levels very close to 100%, only one-third operate at the maximum level, and a possible explanation for this is the existence of market frictions that prevent all seats in public schools from being filled.

To analyze the likelihood of these frictions, we analyze data from the centralized system for admission in federal schools, SISU.

One of the main weaknesses of SISU is that it does not encompass all schools in the market. Therefore, many students who report a degree as their first choice may prefer a non-listed degree. Consequently, many admitted students do not enroll and college admissions depend on waitlists. However, finding waitlisted students can be challenging and generate market frictions, as (Kapor et al., 2022) documented for the Chilean case.

Table (10) reports data about the SISU admission process during the beginning of 2018.[16] Five facts can be inferred from the table. First, SISU is sizable: the number of applying students during the first semester of 2018 is close to the total number of students enrolling in college in 2016 (Table 5), and the first edition of SISU in 2018 accounts for more than 40% of enrollment in state and federal degrees in 2016. Second, most students make both first and second choices. After all, the number of applications is nearly double that of students. Third, waitlists are an important admission mechanism, accounting for more than 86% of all admissions. Fourth, oversubscription is considerable: for each admitted student,[17] a degree has, on average, 9.6 students marking it as their first option. Fifth, despite the large oversubscription, many admitted students do not enroll, and many public degrees show some idle capacity (47% of degrees do not fill all their open seats).

Table 10: SISU Statistics (1st round, 2018)

	Data
Applications	3,876,778
Students	1,990,607
Admitted (regular)	220,970
Admitted (waitlist)	1,396,168
Enrollment	193,917
Applications (only 1st choice) / Regular admissions, by degree (mean)	9.655
Proportion of degrees that do not fill all open seats	0.475

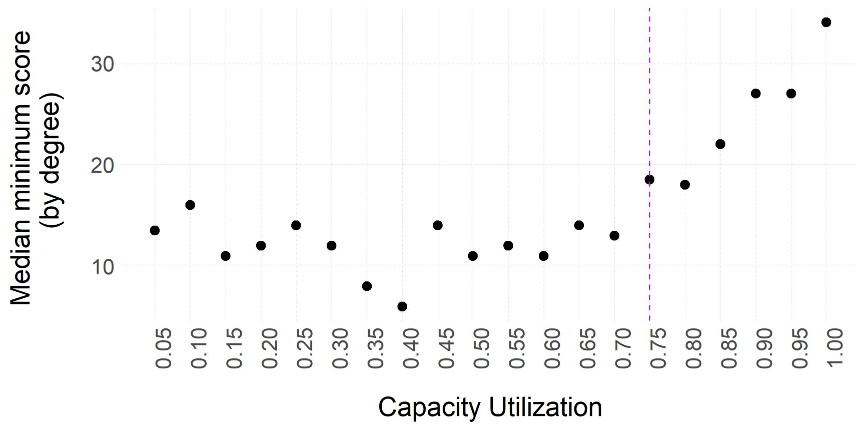
Source: Ministry of Education (2018).

Therefore, oversubscription is strong even in degrees with capacity utilization below 100%. To find a level above which it would be reasonable to assume that capacity is constrained, we explore the fact that capacity constraints increase admission cutoffs. The higher the excess demand, the higher the cutoff. The higher the cutoffs, the higher should be the lowest score among admitted students. In Figure (3), we plot the median of the minimum score among admitted students by groups of capacity utilization.[18] Figure (3) suggests that the lowest score among admitted students gets higher when capacity utilization reaches 0.75.

Therefore, we assume that market frictions make public degrees capacity-constrained when their capacity utilization level reaches 75%. To incorporate the assumption in the model, we proceed as follows. Being Q_d the capacity reported by degree d in the higher education census and N_d the number of students enrolled in d, we assume that capacity (K_d) is given by Equation (14), where κ is the level of capacity utilization above which capacity is constrained. When $N_d > \kappa Q_d$, $K_d = N_d$, meaning that capacity utilization

[16] We do not have the complete data for waitlist admissions in the 2016 and 2017 editions.
[17] Regular admissions, not counting admissions from the waitlists.
[18] Each group corresponds to a 5p.p. interval of the capacity utilization rate.

Figure 3: Median minimum scores, by degree, in each capacity utilization group



is 100% and the degree is capacity-constrained. When $\kappa = 1$, $K_d = Q_d$ and only degrees reporting full capacity utilization are capacity-constrained; when $\kappa = 0$, $K_d = N_d$ and all degrees are capacity-constraned. Therefore, we make $\kappa = 0.75$.

$$K_d(Q_d, \kappa) = \max(\kappa Q_d, N_d) \tag{14}$$

4.2 Demand parameters

The demand parameters are θ and ξ. Identification of demand parameters in a discrete choice model with cutoffs determining heterogeneous choice sets follows Fack et al. (2019) and estimation follows Berry et al. (2004): for each value of θ, there is a unique value of ξ for which the market shares from the estimated model match their sample analogs. So, we estimate θ using a two-step non-linear GMM.

Using the estimated cutoff scores, we assume that the choice set for student i is $\{d : x_i^{SCORE} \geq \hat{\lambda}_d^{x_i^{AA}}\}$.

The first set of moments used for estimation is an exclusion restriction that explores the properties of our cost shifter. To do so, we incorporate fixed effects for major $(r_d \in R)$[19] and ownership status $(h_d \in H)$[20] on degree attractiveness. We do so by calculating the residual degree attractiveness $(\Delta\xi_d)$, as shown in Equation (15).

$$\xi_d = h_d + r_d + \Delta\xi_d \tag{15}$$

Because the econometrician does not observe $\Delta\xi_d$, but schools and students do, esti-

[19] R is the set of majors.
[20] $H = \{Public, Private\}$.

mating θ requires instrumental variables that deal with the potential correlation between p_d and $\Delta\xi_d$. Our instrumental variables strategy for degree d relies on the regional variation of wages associated with occupations that are important inputs to higher education supply. We use eight cost shifters representing the average hourly wages paid by higher education institutions in the same location[21] of degree d. The first instrument uses all occupations in the sample. Each of the other seven instruments uses the hourly wages of faculty in each of the following departments: (1) Mathematics, Statistics, and Computer Science; (2) Architeture, Urbanism, Engineering, Geology, and Geophysics; (3) Biology and Health; (4) Pedagogy; (5) Language and Literature; (6) Humanities; (7) Economics, Business, and Accounting.

We include average wages by department to allow our instrument to isolate differences in the hourly cost of hiring professors in each region. We implicitly assume that those costs are driven by supply-side restrictions, such as the availability of professors in each location.

Defining Z as the matrix in which columns index cost shifters and rows index degrees, and $\Delta\xi$ as a vector for residual degree attractiveness, the moment condition for the cost shifter is summarized in Equation (16).[22]

$$E[Z^T \Delta\xi] = \overrightarrow{0} \tag{16}$$

Following Berry et al. (2004), the second set of moments used for estimation makes the model and the sample closer in several dimensions: the proportion of students in each SES group (by degree), the proportion of students eligible for affirmative action (by degree), the average students' location (by degree), and the average test scores (by degree). These dimensions are represented by g[23] and their corresponding moment conditions are shown in Equation (17).

$$\begin{aligned} E\left[\mathbb{1}_{(r_d=r)}\left[\frac{\sum_{X_i} M(X_i)s_d(X_i,p_d,p_{-d},\lambda_d,\lambda_{-d};\xi,\theta)g(X_i)}{S_d(p_d,p_{-d},\lambda_d,\lambda_{-d};\xi,\theta,M)} - \frac{\sum_{X_i\in A_d} g(X_i)}{\sum_{X_i\in A_d} 1}\right]\right] &= 0\ \forall g,r \\ E\left[\mathbb{1}_{(h_d=h)}\left[\frac{\sum_{X_i} M(X_i)s_d(X_i,p_d,p_{-d},\lambda_d,\lambda_{-d};\xi,\theta)g(X_i)}{S_d(p_d,p_{-d},\lambda_d,\lambda_{-d};\xi,\theta,M)} - \frac{\sum_{X_i\in A_d} g(X_i)}{\sum_{X_i\in A_d} 1}\right]\right] &= 0\ \forall g,h \end{aligned} \tag{17}$$

There are two major threats to our identification strategy. The first is that local wages may not be strongly correlated with degree prices. The second is that our instrument may be correlated with demand shocks. For example, if some regions have higher total factor productivity, they may simultaneously have higher wages and higher education demand. In this case, $E[Z^T\Delta\xi]$ would be positive (instead of zero).

[21] Combined statistical area.

[22] $\overrightarrow{0}$ is a vector of zeros.

[23] $g(X_i)$: $\mathbb{1}_{(x_i^{SES}=j)}$, x_i^{AA}, x_i^{SCORES}, or d_{id}

We estimate a homogeneous demand model to evaluate these threats. To do so, we reestimate the model assuming $u_{id} = \alpha p_d + h_d + r_d + \Delta\xi_d + \epsilon_{id}$. We do so for two reasons: first, the homogeneous demand model can be written in a reduced form that allows the comparison of OLS and 2SLS results; second, first-stage estimates provide useful information about the inclusion restrictions. Results (reported in the Appendix) show that the OLS estimate for the price sensitivity component is not statistically significant, but the 2SLS estimate is negative and statistically different from zero, ruling out the possibility that $E[Z^T \Delta\xi] > 0$. Moreover, first-stage results suggest that our instrument satisfies the inclusion restriction.

4.3 Marginal costs

Marginal costs are estimated using the demand curve and the first-order condition of private degrees, shown in Equation (13). We do not estimate marginal costs for public schools, assuming that the costs for the public sector depend on capacity, not enrollment. Consequently, the cost for public schools is the same in the baseline and all counterfactual simulations, since our counterfactual policies do not change public school capacity.[24]

$$S_d(p_d, p_{-d}, \lambda_d, \lambda_{-d}; \xi, \theta, M) + \sum_{k \in D_j} (p_k - c_k) \frac{\partial S_k(p_k, p_{-k}, \lambda_k, \lambda_{-k}; \xi, \theta, M)}{\partial p_d} = 0 \tag{18}$$

5 Results

5.1 Preferences

Table (11) shows demand estimates. All coefficients are statistically significant and have the expected signs. As expected, the price sensitivity coefficient becomes closer to zero as the socioeconomic status increases (as shown by $\hat{\alpha}^1_{SES}$ to $\hat{\alpha}^5_{SES}$). Even though the point estimates suggest that the fourth socioeconomic status ($\hat{\alpha}^4_{SES}$) has a smaller price sensitivity than the fifth ($\hat{\alpha}^5_{SES}$), the differences are not statistically significant. Higher scores also imply a lower price sensitivity ($\hat{\alpha}_{SCORE}$). Moreover, the distance between students and degrees strongly negatively affects indirect utility ($\hat{\beta}$).

[24]Some of our simulations assume public schools have no capacity constraints, but they are only used to decompose the overall effect of counterfactual policies.

Table 11: Demand estimates

	Estimates
$\hat{\alpha}^1_{SES}$	6.1389^{***}
	(0.0361)
$\hat{\alpha}^2_{SES}$	5.4576^{***}
	(0.0365)
$\hat{\alpha}^3_{SES}$	5.1869^{***}
	(0.0405)
$\hat{\alpha}^4_{SES}$	4.3425^{***}
	(0.0396)
$\hat{\alpha}^5_{SES}$	4.4627^{***}
	(0.0388)
$\hat{\alpha}_{SCORE}$	-2.4111^{***}
	(0.0407)
$\hat{\beta}$	19.4714^{***}
	(0.1604)
N	8960

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors in parantheses. Test Scores vary between 0 and 1. Distances measured in thousand kilometers. Subsample: 25 states (out of 27)

Equation (19) shows the own-price elasticity of student i (η_d). Figure (4) shows the mean of the price elasticities of students at each point of the SES distribution. Even though $\hat{\alpha}^4_{SES}$ is slightly lower than $\hat{\alpha}^5_{SES}$, the average own-price elasticity of students from the 4^{th} SES group is higher than that of the students from the 5^{th}, as expected.

$$\eta_d(X_i, p_d, p_{-d}, \lambda_d, \lambda_{-d}; \xi, \theta) = \alpha^{x_i^{SES}}_{SES}\, p_d(1 - s_d(X_i, p_d, p_{-d}, \lambda_d, \lambda_{-d}; \xi, \theta)) \tag{19}$$

5.2 Markups

Markup for degree d (b_d) follows Equation (20), and Figure (5) shows markup estimates. The median markup is 0.44.

$$b_d = \frac{p_d - c_d}{p_d} \tag{20}$$

Figure 4: Mean own-price elasticity by SES group

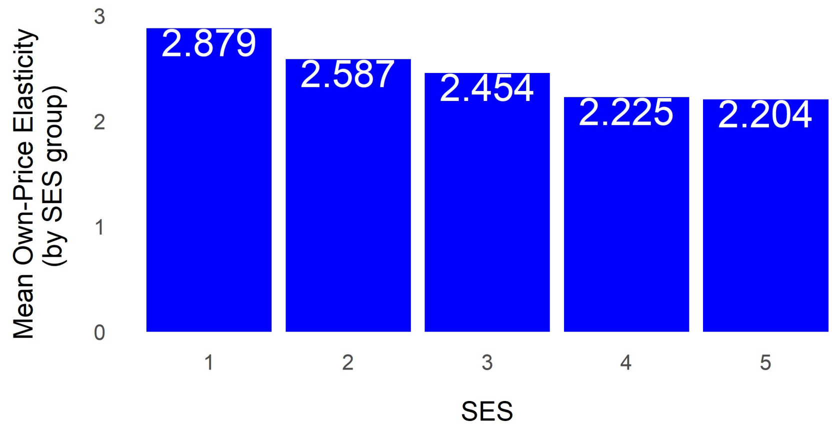
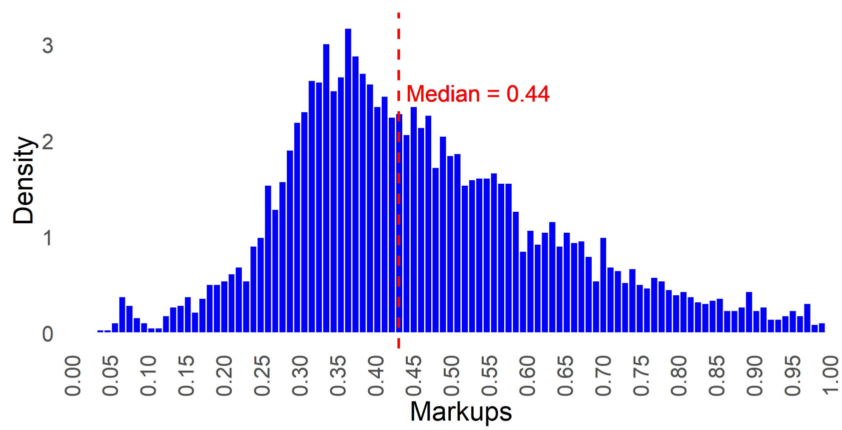


Figure 5: Histogram of markup estimates (by degree)



6 Counterfactuals

Using the model, we simulate two types of counterfactual policies: conditional and unconditional subsidies. Both target low-income students to increase the enrollment of low-SES students in higher education. We assume that the government does not have reliable information on income, so the feasible way to target low-income students is to target those eligible for affirmative action. Moreover, subsidies are flat to reduce the pass-through of unconditional subsidies to private prices (Sahai, 2023).

A subsidy design w is defined by the vector $\Theta_w = (T_w, P_w, t_w)$. T_w is the stipend value, P_w is the set of subsidized schools, and t_w is the price of public schools. Unlike T_w, which affects only students eligible for affirmative action policies, t_w affects all students.

Under the subsidy design w, degree d charges $\tilde{p}_{dw}$.[25] On the other hand, the price student i pays to choose degree d depends on her individual characteristics (X_i) according to $\tilde{p}^{Std}_{dw}(X_i)$. Prices chosen by schools do not depend on students' characteristics because we assume no price discrimination. The relationship between the price students pay and the price schools receive is given by Equation (21).

The stipend depends on three components. First, the subsidy value: T_w, fixed for all recipients. Second, the eligibility status for students: only students eligible for affirmative action receive the scholarship. Third, the eligibility status for degrees: when subsidies are unconditional, the set of subsidized degrees is the set with all degrees in the market ($P_w = U = \{d : \forall d\}$); when subsidies are conditional, only public degrees receive the subsidy ($P_w = C = \{d : h_d = Public\}$). In the baseline, $w = 0$ ($T_0 = 0, P_0 = C, t_0 = 0$).

$$\tilde{p}^{Std}_{dw}(X_i, \tilde{p}_{dw}) = \tilde{p}_{dw} - T_w \mathbb{1}_{(x^{AA}_i = 1 \wedge d \in P_w)} + t_w \mathbb{1}_{d \in Public} \tag{21}$$

Market share for degree d under subsidy design w is given by Equation (22).

$$\tilde{S}_{dw}(\tilde{p}_{dw}, \tilde{p}_{-dw}, \tilde{\lambda}_{dw}, \tilde{\lambda}_{-dw}; \xi, \theta, M) = \sum_{X_i} M(X_i) \tilde{s}_{dw}(X_i, \tilde{p}_{dw}, \tilde{p}_{-dw}, \tilde{\lambda}_{dw}, \tilde{\lambda}_{-dw}; \xi, \theta)$$

Where:

$$\tilde{s}_{dw}(X_i, \tilde{p}_{dw}, \tilde{p}_{-dw}, \tilde{\lambda}_{dw}, \tilde{\lambda}_{-dw}; \xi, \theta) = s_d(X_i, \tilde{p}^{Std}_{dw}(X_i, \tilde{p}_{dw}), \tilde{p}^{Std}_{-dw}(X_i, \tilde{p}_{-dw}), \tilde{\lambda}_{dw}, \tilde{\lambda}_{-dw}; \xi, \theta) \tag{22}$$

To evaluate equity in counterfactual equilibria, we use the high-low SES gap ($\tilde{G}_w$), which is the difference in enrollment between students in the highest SES quintile and students in the lowest SES quintile, as shown by Equation (23).

$$\begin{aligned} \tilde{G}_w = &\frac{\sum_{X_i : x^{SES}_i = 5} M(X_i) \tilde{s}_{dw}(X_i, \tilde{p}_{dw}, \tilde{p}_{-dw}, \tilde{\lambda}_{dw}, \tilde{\lambda}_{-dw}; \xi, \theta)}{\sum_{X_i : x^{SES}_i = 5} M(X_i)} - \\ &\frac{\sum_{X_i : x^{SES}_i = 1} M(X_i) \tilde{s}_{dw}(X_i, \tilde{p}_{dw}, \tilde{p}_{-dw}, \tilde{\lambda}_{dw}, \tilde{\lambda}_{-dw}; \xi, \theta)}{\sum_{X_i : x^{SES}_i = 1} M(X_i)} \end{aligned} \tag{23}$$

Following Small and Rosen (1981), we use Equation (24) to compare the consumer surplus changes from each policy reform ($\Delta \tilde{C}_w$).

[25] Variables with a tilde are those calculated under a counterfactual subsidy design.

$$
\begin{aligned}
\Delta\tilde{C}_w &= \sum_{X_i} \frac{M(X_i)}{\alpha(X_i;\theta)} \Delta\tilde{c}_w(X_i) \\
\Delta\tilde{c}_w(X_i) &= ln\left(\sum_d \mu(X_i;\tilde{\lambda}_{dw}) \exp(V_d(X_i, \tilde{p}^{Std}_{dw}(X_i, \tilde{p}_{dw}); \xi_d, \theta))\right) - \\
&\quad - ln\left(\sum_d \mu(X_i;\tilde{\lambda}_{d0}) \exp(V_d(X_i, \tilde{p}^{Std}_{d0}(X_i, \tilde{p}_{d0}); \xi_d, \theta))\right)
\end{aligned} \tag{24}
$$

6.1 Capacity Constraints in the Baseline

The first set of counterfactual exercises evaluates the effect of removing capacity constraints in the baseline. The exercise is not a counterfactual policy, but, as we will show, it is an important tool for understanding the role of capacity constraints in the model.

We decompose the impact of subsidies on equilibrium outcome y for degree d under subsidy w ($\tilde{y}_{dw}$) into a direct and an indirect effect, using Equation (25). Variables with a *NoCC* superscript are calculated assuming public degrees do not face capacity constraints.[26] The decomposition allows us to understand the role of two mechanisms behind counterfactual results.

The first of the two components is the direct effect (D^y_{dw}), which measures changes in the outcome y for degree d under subsidy w comparing counterfactual equilibrium outcomes without capacity constraints in public degrees to a counterfactual baseline in which public schools do not have capacity constraints. Therefore, the direct effect measures the pure impact of the subsidy, without capacity constraint effects.

The second is the indirect effect. To understand the indirect effect, it is important to define I^y_{dw}, which is how capacity constraints change equilibrium outcome y for degree d under subsidy w $\left(I^y_{dw} = \frac{\tilde{y}_{dw}}{\tilde{y}^{\text{NoCC}}_{dw}}\right)$. The indirect effect is given by $\frac{I^y_{dw}}{I^y_{d0}}$.

$$
\begin{aligned}
(1+\text{Effect}) &= \frac{\tilde{y}_{dw}}{\tilde{y}_{d0}} = D^y_{dw}\frac{I^y_{dw}}{I^y_{d0}} \\
D^y_{dw} &= \frac{\tilde{y}^{\text{NoCC}}_{dw}}{\tilde{y}^{\text{NoCC}}_{d0}} = 1 + \text{Direct} \\
\frac{I^y_{dw}}{I^y_{d0}} &= 1 + \text{Indirect} \\
I^y_{dw} &= \frac{\tilde{y}_{dw}}{\tilde{y}^{\text{NoCC}}_{dw}} \\
\text{Effect} &= (1+\text{Direct})(1+\text{Indirect}) - 1
\end{aligned} \tag{25}
$$

[26] We show a multiplicative decomposition, but a straightforward additive decomposition can also be used.

Removing capacity constraints from the baseline is equivalent to finding I^y_{d0}. Noteworthy, I^y_{d0} is the same for every possible subsidy w. Therefore, understanding the role of capacity constraints in the baseline is important to understanding the indirect effects from subsidy changes. Results from the simulation that removes capacity constraints in public degrees are shown in Table (12).

Table 12: Effect of Removing Capacity Constraints

Enrollment	23.42%
SES Gap	5.49pp
Consumer Surplus	25.84
Public Share	9.09pp
Private Share	-2.29pp
Outside Share	-6.80pp
Public Enrollment	113.63%
Private Enrollment	-10.88%
Average Market Prices	-26.43%
Average Private Prices	1.88%
Profit	-9.00%

Notes: Consumer surplus in BRL per student in the market. All results compared to baseline levels.

As expected, removing capacity constraints in public degrees significantly increases public school enrollment (by 114%). Furthermore, many students switch from the outside option, causing total enrollment in higher education to increase by 23%. The expansion of public school capacity also increases public subsidies because public schools do not charge tuition (the amount of the subsidy per student would remain the same, but many more students would benefit). Therefore, it is unsurprising that the enrollment gap for students in the highest and lowest SES groups falls, and that average market prices fall.

Even though public subsidies would be much higher without capacity constraints in public degrees, private prices would rise by 1.9%. The result suggests that the composition effect determines price dynamics: more subsidies to public schools increase the average price elasticity of students attending public degrees and reduce the average price elasticity of those attending private schools. Therefore, more subsidies to public degrees would increase private schools' prices (even though their profit falls by 9.0%). Although the markup of private schools increases, their profits fall, suggesting that capacity constraints in public schools, as expected, increase the profit of their private counterparts.

Results also show an important impact of capacity constraints on consumer surplus. The result is driven by the reduction in average market prices and the increase in students' choice sets (capacity restrictions in public schools were lifted, so now all students are admitted). Therefore, policies that reduce subsidies in public schools would have two main effects: the first, a lower consumer surplus due to higher average prices; the second, a higher consumer surplus due to a smaller level of capacity constraints in public degrees. In the next subsection, we explore this result and analyze the effects of counterfactual policies.

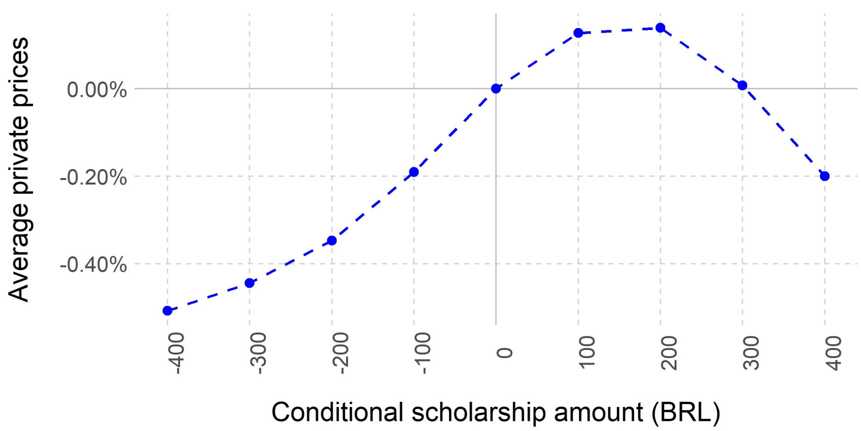
6.2 Counterfactual Policy Outcomes

The second set of counterfactual policies analyzes the impact of varying T_w (the scholarship amount) for both types of subsidy: conditional and unconditional. We focus on the price outcomes. In the simulations shown in this subsection, we make $t_w = 0$ (public school prices vary only in the budget-neutral exercises).

6.2.1 Conditional subsidies

First, we focus on conditional subsidies ($P_w = C$). Figure (6) shows the impact of conditional subsidies on private prices, varying T_w. All results are compared to the baseline unless stated otherwise. Two main facts can be inferred from the graph: the impact of changing the subsidy is very close to zero and non-monotonic: public subsidies increase private prices until private prices start going down.

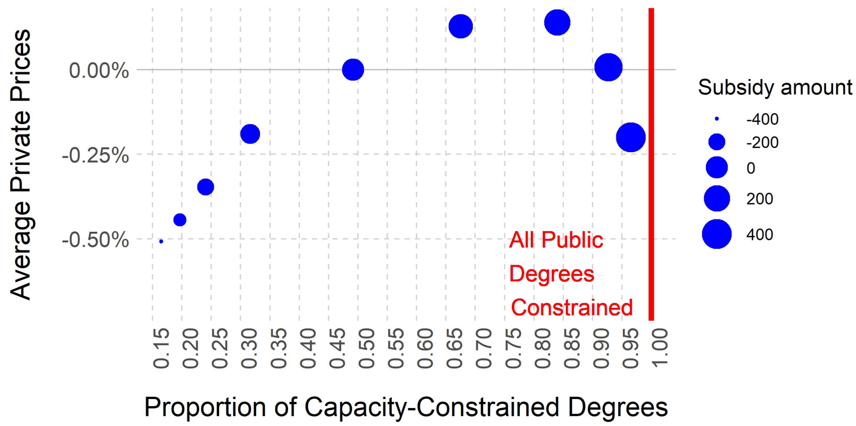
Figure 6: Conditional subsidies and private prices



Notes: median impact on private prices in each counterfactual simulation for conditional subsidies. In the x-axis, we show several possible values for the monthly scholarship; in the y-axis, we show the increase in average private prices.

To better understand this effect, Figure (7) shows the relationship between capacity constraints, conditional subsidies, and private prices. Increasing public subsidies increases private prices when the proportion of capacity-constrained degrees is low. However, the relationship is reversed as the proportion of capacity-constrained degrees becomes closer to its maximum (100%). In other words, capacity constraints are crucial in explaining why the relationship between public subsidies and private prices is not monotonic.

Figure 7: Conditional subsidies, private prices, and capacity constraints



Notes: change in average prices in each counterfactual simulation for conditional subsidies. In the x-axis, we show several possible values for the proportion of capacity-constrained public degrees; in the y-axis, we show the increase in average prices paid by the students in each scenario; the size of the bullets represents the subsidy amount.

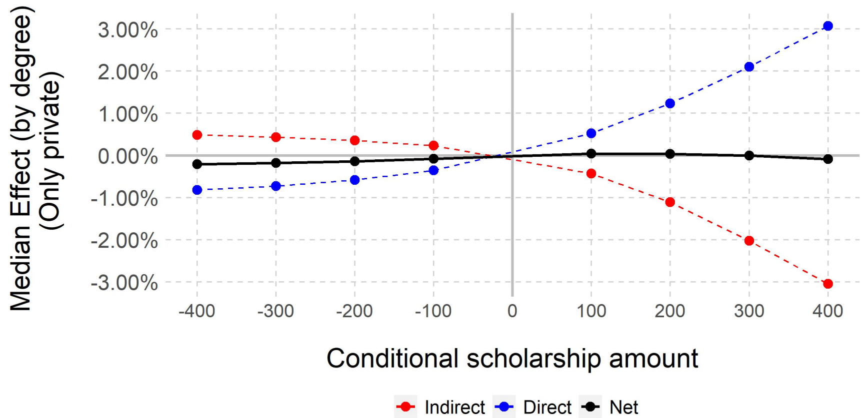
The decomposition of price effects into direct and indirect effects can cast more light on the mechanics of the equilibrium results. Results from the decomposition are shown in Figure (8). The direct effect (when public schools have no capacity constraints) from increasing conditional subsidies makes private prices go up. However, the indirect effect goes in the opposite direction. The mechanics behind the result are as follows: more subsidies in public schools increase the level of capacity constraints, raising admission scores and forcing low-score students to switch from public to private schools; because students with lower scores are more price-elastic, this increases the price elasticity of students attending private degrees and reduces prices. Noteworthy, when subsidies to public schools are high, the indirect effect dominates the direct effect and changes the median pass-through direction.

We can apply a similar procedure to decompose the effect of conditional subsidies on the average price sensitivity ($PS_{d,w}$) of students attending private degrees.[27] The average price sensitivity of students attending degree d under subsidy design w is given by Equation (26). The objective of understanding the behavior of the average price sensitivity under each counterfactual subsidy simulation is to understand to which extent the price effects are being driven by the composition effect. In the literature about the equilibrium effects of subsidies, there are two important drivers of the effect of subsidies on prices: the individual effect (how the price elasticity of the consumers of good A is changed by policy reform) and the composition effect (how more or less price-elastic consumers consume more or less of good A after the reform).

If the composition effect is an important driver of our price responses, we will see the

[27] Now, the decomposition is additive.

Figure 8: Pass-through decomposition: conditional subsidies



Notes: median impact on private prices in each counterfactual simulation for conditional subsidies. In the x-axis, we show several possible values for the monthly scholarship; in the y-axis, we show the median price increase in each scenario. The figure reports the direct effect (when public schools do not face capacity constraints) and the indirect effect (from changes in the level of capacity constraints in public schools). The direct and indirect effects are defined in Equation (25).

price sensitivity of students attending private degrees and private prices move in opposite directions. Until now, we discussed our equilibrium effects emphasizing the relevance of the composition effect, and the price sensitivity decomposition offers evidence that the composition effect is indeed the main driver of the price equilibrium responses.

$$PS_{d,w} = \frac{\sum_{X_i} M(X_i)\tilde{s}_{dw}(X_i, \tilde{p}_{dw}, \tilde{p}_{-dw}, \tilde{\lambda}_{dw}, \tilde{\lambda}_{-dw}; \xi, \theta)\alpha(X_i;\theta)}{\tilde{S}_{dw}(\tilde{p}_{dw}, \tilde{p}_{-dw}, \tilde{\lambda}_{dw}, \tilde{\lambda}_{-dw}; \xi, \theta, M)}$$

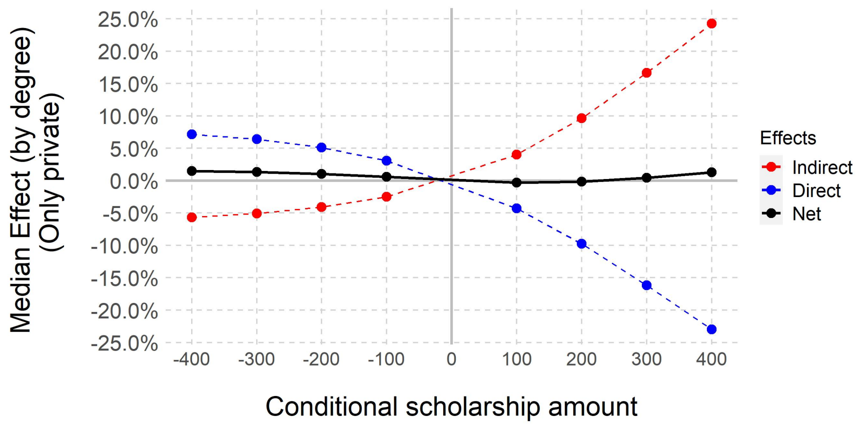
Where:

$$\alpha(X_i;\theta) = \sum_{j=1}^{5} \alpha^j_{SES}\mathbb{1}_{(x_i^{SES}=j)} + \alpha_{SCORE}x_i^{SCORE} \tag{26}$$

As shown in Figure (9) the direct and indirect effects from increasing conditional subsidies impact the price sensitivity of students in private degrees in opposite ways. On the one hand, the direct effect of increasing subsidies makes prices go up and the price sensitivity of students attending private degrees go down, providing evidence that the composition effect is an important driver of the price responses our model generates. Subsidizing public schools incentivizes students with higher price sensitivity to switch from private to public degrees; those who still choose private degrees after the higher subsidy in public schools have lower price sensitivity.

On the other hand, the indirect effect of increasing subsidies makes prices go down and the price sensitivity of students attending private degrees go up, reinforcing the evidence that the composition effect is an important driver of the price responses our model generates. Results suggest that capacity constraints play a vital role in understanding the overall composition effect under conditional subsidies: the sign of the net effect on the average price sensitivity comes from the indirect component in the simulations with higher scholarship amounts. Moreover, the changes in the average price sensitivity of students attending private degrees explain the direction of the changes in private prices.

Figure 9: $PS_{d,w}$ decomposition for each w: conditional subsidies



Notes: median impact on the average price sensitivity of students attending private degrees in each counterfactual simulation for conditional subsidies. In the x-axis, we show several possible values for the monthly scholarship; in the y-axis, we show the median price sensitivity changes in each scenario. The figure reports the direct effect (when public schools do not face capacity constraints) and the indirect effect (from changes in the level of capacity constraints in public schools). The direct and indirect effects are defined in Equation (25), but are additive (instead of multiplicative).

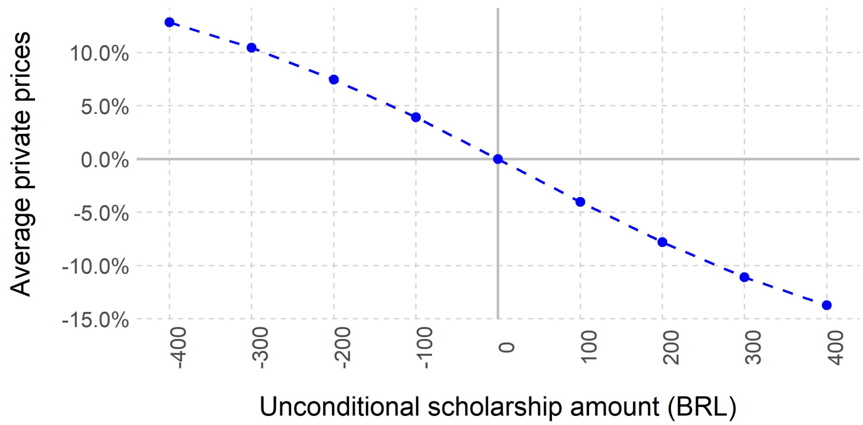
6.2.2 Unconditional subsidies

The effects of unconditional subsidies ($P_w = U$) on average private prices are shown in Figure (10). When subsidies are unconditional, they reduce private prices. For all simulated policies, the relationship between these two variables is monotonic.

As shown in Figure (11), the higher subsidy levels are not enough to make all degrees capacity-constrained.

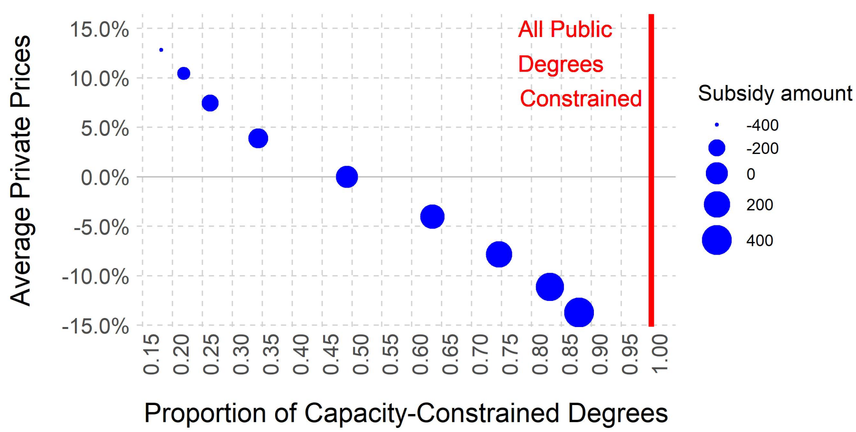
In Figure (12), we decompose the effect of subsidies on the median pass-through using Equation (25). Similarly to what was observed under conditional subsidies, the indirect effect of increasing subsidies reduces private prices. On the other hand, when subsidies increase, the direct effect reduces private prices, opposing the result of the conditional subsidy simulations. In a world without capacity constraints, unconditional subsidies reduce

Figure 10: Unconditional subsidies and private prices



Notes: median impact on private prices in each counterfactual simulation for unconditional subsidies. In the x-axis, we show several possible values for the monthly scholarship; in the y-axis, we show the increase in average private prices in each scenario.

Figure 11: Unconditional subsidies, private prices, and capacity constraints

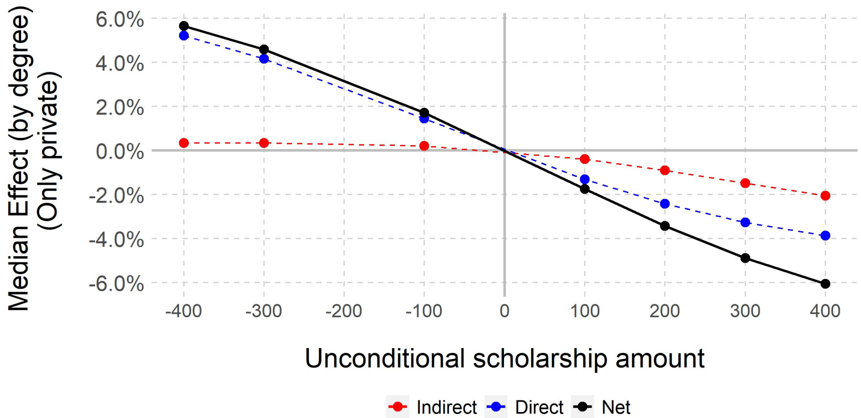


Notes: change in average prices in each counterfactual simulation for unconditional subsidies. In the x-axis, we show several possible values for the proportion of capacity-constrained public degrees; in the y-axis, we show the increase in average prices paid by the students in each scenario; the size of the bullets represents the subsidy amount.

private prices and conditional subsidies increase them. Noteworthy, the effect of capacity constraints becomes sizable as the subsidy grows.

Using the definition of average price sensitivity level in Equation (26) and the additive version of the decomposition outlined in Equation (25), we decompose the average price sensitivity of students attending private degrees under unconditional subsidies. Results are shown in Figure (13). The direct effect of increasing unconditional subsidies is an increase

Figure 12: Pass-through decomposition: unconditional subsidies



Notes: median impact on private prices in each counterfactual simulation for unconditional subsidies. In the x-axis, we show several possible values for the monthly scholarship; in the y-axis, we show the median price increase in each scenario. The figure reports the direct effect (when public schools do not face capacity constraints) and the indirect effect (from changes in the level of capacity constraints in public schools). The direct and indirect effects are defined in Equation (25).

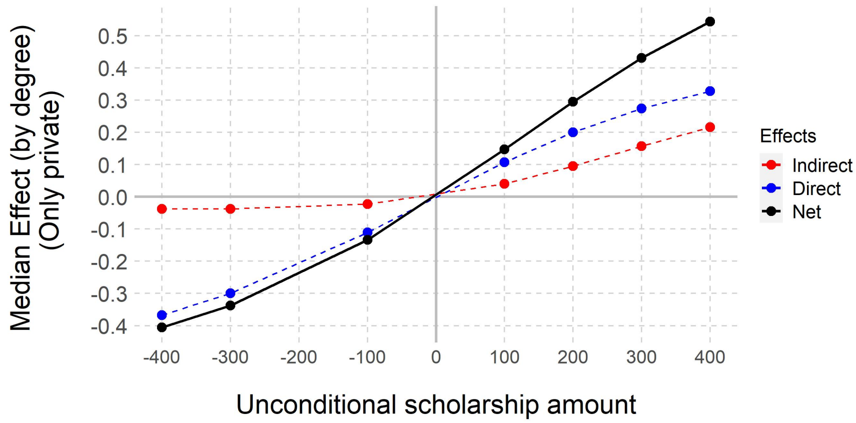
in the price sensitivity of students attending private schools, in line with the fact that the same effect makes prices go down and reinforcing the role of the composition effect to understanding the price responses our model generates. On the other hand, the indirect effect of increasing unconditional subsidies also increase the price sensitivity of students attending private degrees because of the impact of subsidies on admission requirements, forcing low-score students to switch from public to private schools. Results suggest that capacity constraints play a key role to estimating the composition effect under unconditional subsidies.

All in all, results show that a flat, targeted unconditional subsidy does not increase private prices. Such a possibility reduces the concerns about the pass-through of private subsidies to private prices. Since these simulations are not budget-neutral, analyzing their welfare impacts in equilibrium would require knowledge about the marginal costs and benefits of increasing or reducing taxation. So, we redesign previous experiments to make them budget-neutral and look for welfare-improving policies that increase unconditional subsidies: the next set of counterfactuals charge tuition in public schools and use the tuition revenue to pay a scholarship to low-income students.

6.3 Counterfactual Budget-Neutral Policies

The third set of counterfactuals are the budget-neutral policies that increase unconditional subsidies charging tuition in public schools.

Figure 13: $PS_{d,w}$ decomposition for each w: unconditional subsidies



Notes: median impact on the average price sensitivity of students attending private degrees in each counterfactual simulation for unconditional subsidies. In the x-axis, we show several possible values for the monthly scholarship; in the y-axis, we show the median price sensitivity changes in each scenario. The figure reports the direct effect (when public schools do not face capacity constraints) and the indirect effect (from changes in the level of capacity constraints in public schools). The direct and indirect effects are defined in Equation (25), but are additive (instead of multiplicative).

To simulate budget-neutral policies, $t_w > 0$ varies exogenously for public degrees. In each simulation, all public degrees charge the same tuition.

The budget-neutrality condition adjusts T_w to guarantee that the revenue from tuition in public schools matches the aggregate scholarship expenditures. The budget-neutral T_w, $\tilde{T}_w(t_w)$, is shown in Equation (27)

$$\tilde{T}_w(t_w) = \frac{\sum_{d:h_d=Public} \tilde{S}_{dw}(\tilde{p}_{dw}, \tilde{p}_{-dw}, \tilde{\lambda}_{dw}, \tilde{\lambda}_{-dw}; \xi, \theta, M) t_w}{\sum_d \sum_{X_i} M(X_i) \tilde{s}_{dw}(X_i, \tilde{p}_{dw}, \tilde{p}_{-dw}, \tilde{\lambda}_{dw}, \tilde{\lambda}_{-dw}; \xi, \theta) \mathbb{1}_{(x_i^{AA}=1)}} \tag{27}$$

Results from budget-neutral simulations are shown in Table (13). Interestingly, the level of public tuition that maximizes the subsidy is below average private prices in the baseline. The only simulation that raises enrollment is the one in which tuition is small: BRL 300. It allows a monthly stipend much closer to the one provided by charging a BRL 450 tuition but reduces enrollment in public schools much less. When tuition is low, aggregate enrollment rises and the enrollment gap between high-SES and low-SES students falls. The small tuition scenario is the only one in which consumer surplus rises, which is a surprising result given that average market prices rise by more than 22%: most of the effect comes from the reduction in overcapacity in public schools (enrollment in public schools falls significantly in all scenarios), increasing the choice set of low-score students. In all simulations, private enrollment and profits rise. However, private prices decrease because the price elasticity of students attending private degrees rises.

Rising tuition reduces capacity constraints in public degrees, making the positive effects of reducing admission scores in public schools outweigh the negative effects of higher prices in the market.

Table 13: Budget-neutral simulations (unconditional subsidies)

Public tuition	300.00	450.00	650.00	850.00
Monthly subsidy (BRL)	70.26	71.99	61.36	47.09
Enrollment	0.97%	-2.08%	-6.62%	-10.43%
SES Gap	0.25pp	0.00pp	-0.23pp	-0.33pp
Consumer Surplus	0.55	-3.20	-8.80	-13.49
Average Market Prices	22.54%	28.67%	33.39%	35.78%
Public Enrollment	-47.63%	-65.52%	-80.92%	-89.45%
Private Enrollment	19.45%	22.04%	21.63%	19.61%
Profit	15.24%	17.51%	17.64%	16.49%
Average Private Prices	-3.07%	-3.20%	-2.75%	-2.11%

Notes: Subsidy and tuition in monthly BRL. Consumer surplus in BRL per student in the market. All results compared to baseline levels.

Therefore, increasing subsidies to private schools can improve the market equilibrium even when public schools are of high quality. Results show that a budget-neutral reform that charges a small tuition fee in public schools and provides an unconditional scholarship can improve market efficiency by increasing enrollment, consumer surplus, and profits. The same reform can increase equity, reducing the gap between low-SES and high-SES students. So, the reform has positive effects on the higher education market, improving the welfare of its participants, and its effects have the potential to expand the positive externalities of higher education.

7 Conclusion

We estimated and simulated a model to compare the effects of conditional and unconditional subsidies on higher education. Our model incorporates degree and student heterogeneity, highlighting the role of capacity constraints and admission scores in the market equilibrium.

The effect of any reform on the price elasticity of the students attending a private degree is an empirical question: even if the price elasticity of every consumer in the market falls, the average price elasticity of consumers attending some degrees may rise depending on the switching behavior of students with higher and lower price elasticities (the composition effect). Because of the composition effect, private and public subsidies can reduce or increase private prices, depending on the context.

We show that the price elasticity of students attending private degrees is lower when public schools are heavily subsidized, increasing private prices. Moreover, if public schools are capacity-constrained and selective, heavily subsidizing public degrees makes them artificially selective, increasing admission requirements and reducing the admissions of low-score students. Since low-score students typically have higher price sensitivity, capacity constraints affect the average price elasticity of students attending private degrees. Simulations showed that capacity constraints change the direction of the impact of public subsidies on private prices: without capacity constraints, heavily subsidizing public schools raises private prices; with capacity constraints, subsidies to public schools pass through admission scores and reduce private prices.

Unconditional subsidies reduce private prices because price-elastic students switch to private degrees, and admission scores in public schools rise. So, market characteristics reduce the concern that private schools would raise prices if subsidized, and it may be possible to design a reform that increases subsidies to private schools and improves welfare.

Then, we simulate budget-neutral policies that charge tuition in public schools and use the tuition revenue to give an unconditional scholarship to low-income students. Those simulations found that charging small tuition in public schools (BRL 300 monthly) and distributing unconditional stipends of approximately BRL 70 monthly raises enrollment by approximately 1%, reducing the enrollment gap between high-SES and low-SES students. The same reform also raises consumer surplus, even though average prices rise more than 22%. Profits rise by 15%, private prices fall 3%, and private enrollment rises by 19%.

References

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

Atal, J. P., Cuesta, J. I., González, F., and Otero, C. (2022). The Economics of the Public Option: Evidence from Local Pharmaceutical Markets.

Balinkski, M. and Sönmez, T. (1999). A Tale of Two Mechanisms: Student Placement. *Journal of Economic Theory*, 84.

Banerjee, A., Hanna, R., Kyle, J., Olken, B., and Sumarto, S. (2019). Private Outsourcing and Competition: Subsidized Food Distribution in Indonesia. *Journal of Political Economy*, 127.

Berry, S., Levinsohn, J., and Pakes, A. (2004). Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market. *Journal of Political Economy*, 112.

Brazilian Ministry of Education (2018). Regulation 1,095.

Busso, M. and Galiani, S. (2019). The Causal Effect of Competition on Prices and Quality: Evidence from a Field Experiment. *American Economic Journal: Applied Economics*, 11.

Bó, I. and Hakimov, R. (2019). Iterative versus standard deferred acceptance: experimental evidence. *The Economic Journal*, 130.

Cellini, S. (2009). Crowded Colleges and College Crowd-Out: The Impact of Public Subsidies on the Two-Year College Market. *American Economic Journal: Economic Policy*, 1.

Cellini, S. and Goldin, C. (2014). Does Federal Student Aid Raise Tuition? New Evidence on For-Profit Colleges. *American Economic Journal: Economic Policy*, 6.

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

Cohodes, S. and Goodman, J. (2014). Merit Aid, College Quality, and College Completion: Massachussetts' Adams Scholarship as an In-Kind Subsidy. *American Economic Journal: Applied Economics*, 6.

Curto, V., Einav, L., Finkelstein, A., Levin, J., and Bhattacharya, J. (2019). Health Care Spending and Utilization in Public and Private Medicine. *American Economic Journal: Applied Economics*, 11.

Dinerstein, M. and Smith, T. (2021). Quantifying the Supply Response of Private Schools to Public Policies. *American Economic Review*, 111(10):3376–3417.

Dobbin, C., Barahona, N., and Otero, S. (2022). The Equilibrium Effects of Subsidized Student Loans.

Dynarski, S., Page, L., and Scott-Clayton, J. (2022). College Costs, Financial Aid, and Student Decisions.

Epple, D., Romano, R., Sarpça, S., and Sieg, H. (2017). A general equilibrium analysis of state and private colleges and access to higher education in the us. *Journal of Public Economics*, 155.

Estevan, F., Gall, T., and Morin, L.-P. (2018). Redistribution without distortion: evidence from an affirmative action programme at a large Brazilian university. *The Economic Journal*, 129.

Fack, G., Grenet, J., and He, Y. (2019). Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions. *American Economic Review*, 109.

Fonseca, J. and Matray, A. (2022). Financial Inclusion, Economic Development, and Inequality: Evidence from Brazil.

Gale, D. and Shapley, L. (1962). College Admissions and the Stability of the Marriage. *The American Mathematical Monthly*, 69(1):9–15.

Handbury, J. and Moshary, S. (2021). School Food Policy Affects Everyone: Retail Responses to the National School Lunch Program.

INEP (2015). Exame Nacional do Ensino Médio.

INEP (2016). Censo da Educação Superior.

Jiménez-Hernández, D. and Seira, E. (2022). Should the Government Sell You Goods? Evidence from the Milk Market in Mexico.

Kapor, A., Karnani, M., and Neilson, C. (2022). Aftermarket Frictions and the Cost of Off-platform Options in Centralized Assignment Mechanisms.

Kargar, M. and Mann, W. (2022). The Incidence of Student Loan Subsidies: Evidence from the PLUS Program. *The Review of Financial Studies*, 36.

Levy, D. (2018). Global private higher education: an empirical profile of its size and geographical shape. *Higher Education*, 76.

Long, B. (2004). Does the Format of a Financial Aid Program Matter? The Effect of State In-Kind Tuition Subsidies. *Review of Economics and Statistics*, 86.

Lucca, D., Nadauld, T., and Shen, K. (2018). Credit Supply and the Rise in College Tuition: Evidence from the Expansion in Federal Student Aid Programs. *The Review of Financial Studies*, 32.

Peltzman, S. (1973). The Effect of Government Subsidies-in-Kind on Private Expenditures: The Case of Higher Education. *Journal of Political Economy*, 81.

Sahai, H. (2023). The Welfare Implications of School Voucher Desing: Evidence from India.

Singell, L. and Stone, J. (2007). For whom the Pell tolls: The response of university tuition to federal grants-in-aid. *Economics of Education Review*, 26.

Small, K. and Rosen, H. (1981). Applied Welfare Economics with Discrete Choice Models. *Econometrica*, 49.

Turner, L. (2017). The Economic Incidence of Federal Student Grant Aid.

Turner, N. (2012). Who benefits from student aid? The economic incidence of tax-based federal student aid. *Economics of Education Review*, 31.

A Appendix

A.1 SES and Affirmative Action Eligibility Data

Per capita income is the quotient between total family income and the number of family members. The number of family members is observable (the sixth question in ENEM's socioeconomic questionnaire, varying from 1 to 20). However, total family income (the fifth question in ENEM's socioeconomic questionnaire) is measured using 17 possible income ranges. Since the range does not allow for a direct calculation of *per capita* income, we assume family income is in the midpoint of each range. The last range has no maximum, so we extrapolate the income growth from the third to the second to last ranges and assume the last range is equivalent to 21.5 times the minimum wage $(17.5 + (17.5 - 13.5))$, as shown in Table (14).

Table 14: Total family income

Range (ENEM's questionnaire)	Imputation
No income	0.00
Up to 1.0 minimum wage	0.50
From 1.0 to 1.5 minimum wage	1.25
From 1.5 to 2.0 minimum wages	1.75
From 2.0 to 2.5 minimum wages	2.25
From 2.5 to 3.0 minimum wages	2.75
From 3.0 to 4.0 minimum wages	3.50
From 4.0 to 5.0 minimum wages	4.50
From 5.0 to 6.0 minimum wages	5.50
From 6.0 to 7.0 minimum wages	6.50
From 7.0 to 8.0 minimum wages	7.50
From 8.0 to 9.0 minimum wages	8.50
From 9.0 to 10.0 minimum wages	9.50
From 10.0 to 12.0 minimum wages	11.00
From 12.0 minimum wages to 15.0 minimum wages	13.50
From 15.0 minimum wages to 20.0 minimum wages	17.50
More than 20.0 minimum wages	21.50

Affirmative action eligibility comes from students' answers to the 47th question in ENEM's socioeconomic questionnaire. Only students who answered that they were continuously enrolled in public high schools were eligible for affirmative action policies.

A.2 Instruments

Two major threats to our identification strategy exist. The first is that our instrument, local wages, has little variation and is not strongly correlated with prices. The second is that our instrument is correlated with demand shocks. For example, if some regions have higher

Table 15: Homogeneous demand

	OLS	OLS (FE)	2SLS	2SLS (FE)
Price	0.108^{***}	-0.024	-2.070^{***}	-2.385^{***}
	(0.018)	(0.028)	(0.232)	(0.272)
Major FE	No	Yes	No	Yes
Ownership FE	No	Yes	No	Yes
Num. obs.	8960	8960	8960	8960

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors in parantheses.

total factor productivity, they may simultaneously have higher wages and higher education demand. In this case, $E[Z^T \Delta\xi] > 0$.

We estimate a demand model without consumer heterogeneity to discuss the relevance of these two threats by comparing the results of OLS and 2SLS estimates and reporting first-stage results. The estimated model assumes that $\alpha(X_i, \theta) = -\alpha$, resulting in the reduced form shown in Equation (28).

$$log(s_d) - log(s_0) = \alpha p_d + h_d + r_d + \Delta\xi_d \tag{28}$$

OLS and 2SLS estimates are shown in Table (15). In our preferred specification (with fixed effects), OLS estimates are statistically insignificant and very close to zero, while 2SLS estimates are statistically significant and negative. These results are robust to the model without fixed effects. If $E[Z^T \Delta\xi] > 0$, the 2SLS estimates would be higher than the OLS estimates, meaning that our results seem to reduce our concerns with the identification threats we presented.

Moreover, we use the simplified model to evaluate the correlation between prices and our instruments using the first stage of the 2SLS procedure. Results are shown in Table (16). Because public school prices are always zero (exogenously determined), we also estimate a first-stage regression including only private degrees in the sample. The F-statistic for the regression with all schools is 21.6 and 61.0 when only private schools are included, suggesting that regional wages are correlated with degree prices.

The signs of some estimates are negative because of the correlation between our explanatory variables. We estimate alternative versions of our first-stage results, including only one instrument at a time in a sample that has only private degrees. Results are reported in Table (17).[28] All estimates have positive signs and are statistically significant. Furthermore, in Table (18) we report the correlation between our instruments: many of them are higher than 0.50.

[28] Group 1: Mathematics, Statistics, and Computer Science. Group 2: Architecture, Urbanism, Engineering, Geology, and Geophysics. Group 3: Biology and Health. Group 4: Pedagogy. Group 5: Language and Literature. Group 6: Humanities. Group 7: Economics, Business, and Accounting.

Table 16: First-stage results (homogeneous demand)

	First-stage	Only private degrees
Mathematics, Statistics, and Computer Science	-0.011	-0.289^{***}
	(0.082)	(0.081)
Architecture, Urbanism, Engineering, Geology, and Geophysics	0.371^{***}	0.709^{***}
	(0.116)	(0.112)
Biology and Health	-0.159	-0.409^{***}
	(0.114)	(0.109)
Pedagogy	0.677^{***}	1.072^{***}
	(0.077)	(0.074)
Language and Literature	0.002	-0.049
	(0.095)	(0.091)
Humanities	0.049	0.319^{***}
	(0.101)	(0.099)
Economics, Business, and Accounting	0.144^{***}	0.204^{***}
	(0.053)	(0.049)
All employees	-0.227	-0.133
	(0.158)	(0.157)
Major FE	Yes	Yes
Ownership FE	Yes	Yes
Sample	All degrees	Only private
R^2 (full model)	0.6272	0.7695
R^2 (projection model)	0.0191	0.0768
F-statistic (full model)	165.77	272.28
F-statistic (projection model)	21.63	61.04
Num. obs.	8960	5944

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors in parantheses. Cost shifters are the wages of faculty, by department, and the wages of all college employees. Statistics reported for the full (projection) model do (not) include major dummies.

Table 17: Robustness for first stage (homogeneous demand)

	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
All	1.445^{***}							
	(0.125)							
Group 1		0.407^{***}						
		(0.054)						
Group 2			0.606^{***}					
			(0.056)					
Group 3				0.627^{***}				
				(0.059)				
Group 4					1.025^{***}			
					(0.055)			
Group 5						0.396^{***}		
						(0.051)		
Group 6							0.745^{***}	
							(0.056)	
Group 7								0.247^{***}
								(0.038)
Major FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (full)	0.6218	0.6217	0.6227	0.6221	0.6249	0.6216	0.6225	0.6210
R^2 (proj)	0.0050	0.0049	0.0073	0.0058	0.0131	0.0046	0.0069	0.0029
F (full)	175.81	175.78	176.47	176.04	178.13	175.69	176.37	175.22
F (proj)	44.39	43.55	65.23	51.69	117.76	40.69	62.05	25.82
Num. obs.	5944	5944	5944	5944	5944	5944	5944	5944

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$. Standard errors in parantheses. Regression explaining private prices from the cost shifter and major dummies. Cost shifters are the wages of faculty, by department, and the wages of all college employees. Statistics reported for the full (projection) model do (not) include major dummies.

Table 18: Correlation matrix: cost shifters

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	All
Group 1	1.00	0.72	0.58	0.20	0.63	0.54	0.39	0.33
Group 2	0.72	1.00	0.71	0.15	0.81	0.59	0.52	0.30
Group 3	0.58	0.71	1.00	0.32	0.63	0.80	0.44	0.36
Group 4	0.20	0.15	0.32	1.00	0.07	0.39	-0.05	0.53
Group 5	0.63	0.81	0.63	0.07	1.00	0.54	0.61	0.31
Group 6	0.54	0.59	0.80	0.39	0.54	1.00	0.45	0.39
Group 7	0.39	0.52	0.44	-0.05	0.61	0.45	1.00	0.31
All	0.33	0.30	0.36	0.53	0.31	0.39	0.31	1.00

A.3 SES and Test Score Distributions

Table 19: Minimum, maximum, and mean per capita income in each SES quintile

SES quintile	Min	Max	Mean
1	0.000	0.250	0.123
2	0.250	0.417	0.319
3	0.417	0.625	0.524
4	0.625	1.037	0.799
5	1.037	21.500	2.054

Table 20: Minimum, maximum, and mean test scores in each test score percentile

Test score percentile	Min	Max	Mean
1	258.520	371.940	346.756
2	371.940	390.620	382.508
3	390.620	401.200	396.270
4	401.200	408.840	405.178
5	408.840	414.960	411.999
6	414.960	420.160	417.628
7	420.160	424.740	422.497
8	424.740	428.860	426.836
9	428.860	432.600	430.751
10	432.600	436.100	434.355
11	436.100	439.360	437.747
12	439.360	442.400	440.897
13	442.400	445.280	443.867
14	445.280	448.040	446.676
15	448.040	450.680	449.379
16	450.680	453.180	451.945
17	453.180	455.620	454.413
18	455.620	457.960	456.790
19	457.960	460.260	459.112
20	460.260	462.500	461.379
21	462.500	464.640	463.565
22	464.640	466.740	465.683
23	466.740	468.800	467.765
24	468.800	470.820	469.794
25	470.820	472.720	471.752
26	472.720	474.640	473.671
27	474.640	476.480	475.566
28	476.480	478.260	477.392
29	478.260	480.000	479.179

30	480.000	481.780	480.888
31	481.780	483.340	482.556
32	483.340	485.000	484.228
33	485.000	486.720	485.864
34	486.720	488.120	487.446
35	488.120	489.780	488.956
36	489.780	491.120	490.468
37	491.120	492.800	491.969
38	492.800	494.000	493.420
39	494.000	495.560	494.823
40	495.560	497.000	496.290
41	497.000	498.300	497.642
42	498.300	499.820	499.050
43	499.820	501.000	500.458
44	501.000	502.440	501.779
45	502.440	503.920	503.137
46	503.920	505.000	504.465
47	505.000	506.460	505.768
48	506.460	508.000	507.208
49	508.000	509.060	508.531
50	509.060	510.580	509.869
51	510.580	512.000	511.254
52	512.000	513.120	512.536
53	513.120	514.700	513.932
54	514.700	516.000	515.370
55	516.000	517.340	516.671
56	517.340	518.920	518.097
57	518.920	520.000	519.423
58	520.000	521.520	520.799
59	521.520	523.000	522.291
60	523.000	524.500	523.677
61	524.500	526.000	525.299
62	526.000	527.580	526.718
63	527.580	529.140	528.408
64	529.140	531.000	530.094
65	531.000	532.520	531.694
66	532.520	534.160	533.378
67	534.160	536.000	535.141

68	536.000	537.940	536.883
69	537.940	539.800	538.763
70	539.800	541.600	540.613
71	541.600	543.480	542.482
72	543.480	545.520	544.486
73	545.520	547.760	546.591
74	547.760	549.980	548.789
75	549.980	552.220	551.076
76	552.220	554.700	553.485
77	554.700	557.120	555.911
78	557.120	559.900	558.493
79	559.900	562.640	561.177
80	562.640	565.520	564.043
81	565.520	568.620	567.040
82	568.620	571.820	570.166
83	571.820	575.180	573.496
84	575.180	579.000	577.072
85	579.000	582.860	580.816
86	582.860	587.000	584.913
87	587.000	591.320	589.157
88	591.320	596.000	593.631
89	596.000	601.160	598.548
90	601.160	606.980	604.010
91	606.980	613.180	610.002
92	613.180	619.960	616.499
93	619.960	627.440	623.612
94	627.440	636.020	631.624
95	636.020	645.920	640.874
96	645.920	657.320	651.511
97	657.320	671.400	664.130
98	671.400	689.160	679.852
99	689.160	715.320	701.113
100	715.320	866.480	742.437