Balancing Access and Quality in a Changing Market Structure: The Market Effects of Online College Education[a]

Nano Barahona Cauê Dobbin Joaquín Fuenzalida Sebastián Otero

October 28, 2024

Abstract: We examine the rapid growth of Brazils for-profit online higher education sector and its impact on market structure and enrollment. Exploiting regional and field-specific variation in online degree penetration, we find that online programs expand access for older students but divert younger students from in-person higher-quality programs. Increased competition lowers the prices of in-person programs but leads to a decline in their provision. Using an equilibrium model of college education, we quantify that in the absence of online education, the average student would experience 2.9% higher value added. While young students benefit from fewer online options, older students are disadvantaged. Targeted policies limiting online education to older cohorts have the potential to improve value added across all groups.

[a]First version: November 2024. We thank Ignacio Cuesta, Liran Einav, Matthew Gentzkow, Gautam Gowrisankaran, Caroline Hoxby, Gastón Illanes, Cristóbal Otero, Kiki Pop-Eleches, and Anna Popova for helpful conversations and suggestions, and to seminar participants at Chicago Booth, Northwestern University, University of Tokyo, Peder Sather Conference on IO, O-Lab Conference, and the Bay Area IO Fest for useful comments. We also thank Romina Quagliotti, José Tomás Feliú, Enrico Ruggeri, and Arthur Weiss for excellent research assistance. Barahona: UC Berkeley, nanobk@berkeley.edu. Dobbin: Georgetown University, caue.dobbin@georgetown.edu. Fuenzalida: UC Berkeley, jfuenzalida@berkeley.edu. Otero: Columbia University, seb.otero@columbia.edu.

1. Introduction

Over the past two decades, numerous industries have expanded their offerings to include the online delivery of goods and services, driven by significant advances in digital technology. Service sectors traditionally grounded in physical, face-to-face interactions—such as healthcare and education—are increasingly adopting hybrid or fully virtual models of operation. This shift has been particularly pronounced in post-secondary education, where colleges have increased online offerings to meet the growing demand for more affordable and flexible learning (Deming et al., 2012; Aucejo et al., 2024). In 2019, around 15% of all U.S. undergraduate students were enrolled exclusively in distance education, a figure that rose to 24% by 2022, accelerated by the COVID-19 pandemic (NCES, 2022).

Expanding education—or any service—into online formats presents both opportunities and risks. On the one hand, online education can increase access for previously excluded students, democratizing higher education access (Barrow et al., 2024). On the other hand, the shift to remote delivery can fundamentally change the nature of these services, potentially compromising their quality (Bettinger et al., 2017; Garrett et al., 2022). As a result, online education may boost overall enrollment in higher education but also risks diverting students from high-quality, in-person instruction to potentially inferior online alternatives. This is particularly problematic in higher education, where quality is difficult to assess and remains a key policy concern.

These concerns are amplified when considering the equilibrium effects of online education. The entry of online programs can increase the competitive pressure on in-person programs, initially benefiting students through lower prices. However, in an industry such as higher education, where in-person programs face substantial fixed operational costs, declining demand and prices can lead to in-person program closures, ultimately diminishing the overall quality of educational offerings. This risk is especially problematic in markets with few in-person options, leaving students with potentially lower-quality online alternatives.

This paper examines the impact of online education expansion in Brazil, the world's largest online higher education market. We leverage the differential entry of online programs across regions and fields of study to assess how the introduction of undergraduate online degrees impacts market expansion, student diversion, and in-person program availability. We integrate this analysis with an equilibrium model of college education to quantify the overall effects of online education and evaluate alternative policy designs.

Brazil provides an ideal setting for studying the expansion of online education for two reasons. First, analyzing equilibrium effects requires a setting where online education constitutes a significant share of the education market. In Brazil, fully remote programs have grown rapidly, comprising 17.4% of all undergraduate enrollments in 2010, 43.8% by 2019, and 65% post-COVID. Second, to isolate the causal impact of online education on local market outcomes from nationwide time trends, we need variation in online penetration across local markets. Brazilian regulations provide this variation by requiring online programs to establish local physical hubs, which students need to attend periodically, generating geographic differences in online market entry. Additionally, by prohibiting online education in certain majors, the regulations introduce further variation across fields of study.

For our analysis, we use several administrative datasets. First, we draw on detailed data from the Brazilian Higher Education Census to assess market shares and track the entry and exit of degree programs. We also collect tuition fee data from various sources to examine colleges' pricing strategies. Additionally, we combine university entrance exam data with the Higher Education Census and matched employer-employee records to estimate the labor market returns of specific degrees, allowing us to compute the value added of online and in-person programs. We thus have a rich window into student demand for in-person and online degrees, colleges' behavior, and overall degree quality.

Our analysis focuses on the private sector, which accounts for nearly all online programs and 82% of total enrollment. Between 2010 and 2019, private sector enrollment grew from 1.7 million to 3 million students, with 89% of this growth driven by online programs, particularly at for-profit institutions. Two main factors have driven the growth of online education: improved access for older, lower-income adults looking for affordable and flexible learning options, and a shift away from traditional in-person programs. Since 2010, in-person enrollment has stagnated overall, with significant declines in business and education programs, which have seen the largest increases in online enrollment.

We begin our analysis by comparing online degree programs to traditional in-person programs across duration, tuition fees, dropout rates, and value added. To do this, we examine equivalent programs offered by the same institutions, differing only in delivery mode. Our findings show no significant differences in program duration, consistent with regulations mandating that both formats follow the same curriculum.[1] We also find that tuition fees for online degrees are 60% lower than those for in-person programs. In terms of quality, online programs show significantly higher dropout rates and are associated with a 0.05 standard deviation lower value added, measured by labor market outcomes after accounting for students' entrance exam scores.

[1]Until 2017, all online degree programs had a corresponding in-person equivalent. After 2017, universities were allowed to launch online programs without offering an in-person counterpart.

To examine the causal effects of online program expansion in Brazil, we use a linear model that regresses changes in outcomes—such as student enrollment, market structure, and tuition—on the change in the number of online degrees from 2010 to 2019. We define our unit of analysis as the interaction of a commuting zone and a field of study. We begin by estimating the model using OLS. The validity of this approach relies on the parallel trends assumption requiring that regions and fields with lower online growth would have followed the same trend as higher-growth areas if they had experienced similar online growth. To address potential bias from unobserved shocks, we also implement a shift-share instrumental variable (SSIV) approach (Bartik, 1991; Goldsmith-Pinkham et al., 2020), combining predetermined institution headquarter locations (*shares*) with online sector growth (*shift*). Our instrument builds on the observation that institutions tend to expand their online programs in regions closer to their headquarters and assumes that a regions proximity to an institution is not correlated with unobserved region-specific shocks in fields where the institution offers in-person degrees (i.e., exogenous *shares*).

Both the OLS and SSIV approaches yield similar results: expanding online education increases online enrollment, reduces in-person enrollment, and raises overall college enrollment. This highlights the dual effect of online education. On one hand, it expands access for students who might not otherwise have attended college. On the other hand, it pulls students away from in-person programs, often toward lower-quality online alternatives. This shift heightens competition, forcing local in-person institutions to lower prices and reduce profits. As competition intensifies, in-person programs are less likely to persist in the market, accelerating the shift toward online degrees. As a result, the value added attained by the average student declines. Our cohort analysis shows that this expansion primarily benefits older students, while younger students, who are more likely to enroll in in-person programs ex-ante, experience stronger diversion to online options.

Estimating the linear model provides a clear method for recovering marginal causal effects, but it rests on a strong no-interference assumption, which implies that changes in the number of online degrees in a given region and field only affect that specific field of study. This assumption breaks down if degrees across fields are substitutes. Moreover, the linear structure may not capture out-of-sample counterfactuals effectively, where competition and supply-side responses, such as price changes and entry or exit decisions, can lead to significant non-linear effects. To address these limitations, we develop a supply and demand model for college education that accounts for equilibrium responses, allowing us to assess the impact of online education expansion under different policy scenarios.

Our equilibrium model consists of students and educational institutions. On the demand side, students decide whether to attend college and in which degree to enroll. On the supply side, institutions decide whether to enter a particular market, which degrees to offer, and what prices to charge. To operate in a market, institutions must establish either a campus for in-person degrees or a hub for online degrees. Their decision-making occurs in two stages. In the first stage, institutions simultaneously choose in which regions to operate and which degrees to offer, after observing their fixed entry costs and taking into account pre-existing offerings (Seim, 2006; Atal et al., 2022). They form expectations about their competitors' entry decisions and adjust their strategies accordingly, potentially opting out of opening an in-person campus to avoid fixed costs if they anticipate competitors will expand their online offerings. In the second stage, institutions compete on prices.

We estimate demand by leveraging different variations in the data. To estimate substitution between degrees in different areas of study and delivery modes, we use market-level differences in degree availability induced by the shift-share instrument. Our estimates suggest that in-person and online degrees are close substitutes: when an in-person degree closes, 64% of students would switch to another in-person option, 20% would move online, and 14% would opt out of college. To estimate price elasticities, we use contemporaneous prices of the same degree in other regions as proxies for cost-shifters and instrument for prices (Hausman et al., 1994). The mechanics behind this instrument is that firm-level cost variations influence prices in all markets where the degree is offered, and the key identification assumption is that, although costs for degrees from the same institution are correlated across markets, demand shocks are not. We estimate median own-price elasticities of approximately -3.5 for in-person degrees and -1.5 for online degrees, which aligns with findings from the literature (Dobbin et al., 2021; Armona and Cao, 2022).

For the supply model, we recover entry elasticities with respect to profits using two instruments that affect profits but do not influence institutions' fixed costs. First, we exploit regional variation over time in internet penetration as a demand shifter for online education. Second, we use differences in competitors' distance to various regions, which creates varying levels of competition that impact profits without affecting institutions' fixed costs. We find that in-person campuses are substantially more expensive to open than online hubs.

We use our estimated model to quantify the impact of online education expansion on student enrollment, market structure, tuition fees, and value added. To examine how supply-side equilibrium effects shape these outcomes, we simulate three progressively more flexible counterfactuals where a planner bans the supply of online education. We benchmark each of these scenarios against a baseline counterfactual that reflects the status quo, where online education is permitted.

Under the first counterfactual, we examine the effects of banning online education without accounting for supply-side responses (i.e., fixing degree offerings and prices). Of the students enrolled in online programs under the status quo, 32% exit the market, resulting in a 13.6% drop in total enrollment. Since online degrees provide higher value added than the outside option, the total value added declines by 0.004%. Under the second counterfactual, we allow institutions to respond by adjusting the prices of their degrees. As a result, tuition fees for in-person degrees increase by 14.4%, causing a further reduction in enrollment and total value added. Finally, under the third counterfactual, we allow institutions to adjust both tuition fees and degree offerings. Here, the supply of in-person degrees expands by 15.9%, attracting new students and increasing total enrollment by 4.6%. Under this scenario, total value added is 3.3% higher than under the second counterfactual, underscoring the importance of considering equilibrium responses when assessing the introduction of new technologies in markets.

We assess the distributional consequences of banning online education by evaluating its effect on value added across student cohorts. Under the first counterfactual, which ignores supply-side responses, no cohort experiences significant benefits from the ban, and older cohorts experience a decline in value added. When supply-side responses are considered, younger students benefit from increased access to in-person programs that were previously limited due to the presence of online options. However, older students, who have strong preferences for online degrees, are worse off and tend to exit the market. Our findings show that online degree expansion has uneven effects. While older cohorts benefit from a wider range of degree options and can pursue higher value added programs, the expansion reduces in-person opportunities, pushing younger students into lower-quality alternatives.

Lastly, we use insights from the model to explore government regulations that focus on expanding online degrees to groups that benefit the most. Specifically, we examine the impact of banning online education for young students. Our results indicate that this policy would increase value added for all cohorts compared to the baseline scenario representing the status quo. Our findings shed light on how new technologies in the form of online services can transform competition and market dynamics. While they expand consumer choice and reduce costs, the benefits are unevenly distributed. In markets with imperfect information, some consumers may unknowingly switch from higher-quality options to new technology, leaving them worse off. This issue is exacerbated when traditional alternatives disappear due to declining demand. To address these risks, policymakers could limit access to the new technology for groups that are negatively affected, preserving sufficient demand for incumbent providers. This would allow those who benefit most to adopt the new technology, while others retain access to traditional, higher-quality options.

This paper adds to the growing literature studying the effects of introducing more accessible, lower-quality options in educational markets. Research on community colleges shows they democratize access to higher education but also divert some students from four-year institutions (Rouse, 1995, 1998; Mountjoy, 2022). Likewise, studies on online education underscore its potential to expand access or draw students away from higher-quality alternatives (Deming et al., 2012; Goodman et al., 2019). Additional research also emphasizes the competitive pressure online degrees (i.e., the lower quality option) exert on tuition prices of traditional in-person programs (Deming et al., 2015). Our paper advances this line of work by developing an equilibrium framework that accounts for market expansion, market diversion, price changes, and endogenous degree offerings.

Our research is also related to the literature that examines the effects of online forms of education on student learning and academic progression (Figlio et al., 2013; Bettinger et al., 2017; Kofoed et al., 2024), and labor market outcomes (Deming et al., 2016; Hoxby, 2018; Fabregas and Navarro-Sola, 2024). Our findings are consistent with this literature, highlighting the role of online education as a preferable alternative to no education, albeit less favorable compared to in-person options. Furthermore, our study contributes to the expanding literature on the market effects of online services beyond education, particularly telemedicine. In line with our findings, Goetz (2023) shows that competition from telemedicine in psychotherapy can influence the pricing of high-quality providers while driving lower-quality providers out of the market. Other studies have examined the effectiveness of telemedicine in terms of patient outcomes. Zeltzer et al. (2023) show that telemedicine can reduce overall healthcare spending while maintaining care quality, without compromising diagnostic accuracy or outcomes. Additionally, evidence suggests that online healthcare services can enhance efficiency by offering faster and shorter consultations and improving the matching of doctors and patients (Dahlstrand et al., 2024; Dahlstrand, 2024).

We also contribute to the broader literature that analyzes education policy using equilibrium models of imperfect competition. These models have been used to study the effects of educational policy on pricing and quality in secondary schools (Neilson et al., 2013; Allende, 2019) and colleges (Dobbin et al., 2021; Barahona et al., 2021), instructional levels (Bau, 2022), and on institutions' decisions to participate in voucher programs (Sanchez, 2023). More closely related to our work, three papers assess the impact of competition on market structure. All these papers focus on the impact of improved public sector offerings on private institutions' entry and exit decisions. Bodéré (2023) explores the effects of higher-quality public preschools in Pennsylvania, while Dinerstein and Smith (2021) assess the impact of increased public school funding in New York. Similarly, Dinerstein et al. (2023) investigated the expansion of public schools in the Dominican Republic. Our contribution to this literature is to provide evidence of increased competition driven by the private sector itself through the introduction of a new delivery format into the market.

The paper is organized as follows. Section 2 provides institutional background on Brazil's higher education sector, discusses the data, and presents descriptive statistics on the growth of online education. Section 3 presents the results from the linear model used to estimate the causal effects of expanding online education on various outcomes. In Section 4, we introduce and estimate the equilibrium model, and in Section 5, we use it for counterfactual analysis. Finally, Section 6 concludes the paper.

2. Setting and data

In this section, we describe the Brazilian higher education and online education regulatory landscape. We then describe the data sources and provide several descriptive statistics about the expansion of online education.

2.1. *Online higher education landscape in Brazil*

Brazil's higher education system has experienced significant expansion over the past decade, with new undergraduate enrollment growing from approximately 2.2 million students in 2010 to 3.6 million in 2022. A crucial factor driving this growth has been the rising popularity of online degree programs, predominantly offered by for-profit private institutions.[2] The shift towards online education has been dramatic: in 2010, 17.4% of new students chose online programs, rising to 43.8% by 2019, and soaring to 65% after the COVID-19 pandemic.

Online degree programs in Brazil, referred to as "Educao a Distncia" (EaD), offer remote versions of traditional in-person undergraduate diplomas and are required to adhere to the same curriculum and duration standards.[3] Diplomas make no distinction between whether a degree was earned online or in person, theoretically placing both modes on equal footing. However, despite their lower cost, online programs are often perceived as being of inferior quality, a concern that continues to challenge Brazilian policymakers (Bertolin et al., 2023).

Online programs are required to include in-person sessions for essential activities like

[2]Private institutions make up 95% of the total online education market, with 79% of that share held by for-profit institutions. Furthermore, this market is heavily concentrated, with seven institutions dominating 60% of the entire online education market.

[3]Students have the option to transfer credits between online and in-person formats within the same institution if they change modalities.

assessments and laboratory work, which must be conducted either at the institution's main campus or at designated local hubs. These hubs are decentralized centers created to support the face-to-face elements of online education, placing geographic limits on the reach of these programs. Despite the in-person requirements, all instruction remains fully remote. Most institutions provide live, synchronous classes to facilitate real-time interaction between students and instructors. In addition, 78% of institutions offer asynchronous resources—such as pre-recorded lectures, reading materials, and interactive quizzes—giving students more flexibility to engage with course content on their own schedule (ABED, 2018).

The growth of online education has been driven by several key factors. First, there's growing demand stemming from the flexibility these programs offer, enabling more studentsparticularly an older demographicto pursue higher education while balancing other responsibilities. In 2019, 71% of new online students were over the age of 24, compared to just 32% in traditional in-person programs. Second, the widespread improvement in internet infrastructure across Brazil has significantly facilitated access to online degrees, even in previously underserved regions. In 2010, just 40% of Brazilian households had internet access, but by 2019, this figure surged to 83%. Lastly, government reforms introduced in 2016 have streamlined the accreditation process for new online programs and granted institutions greater autonomy to establish new hubs. This regulatory evolution has made it considerably easier for educational institutions to manage, expand, and diversify their online offerings, contributing to the sector's overall growth.[4]

Certain fields of study face restrictions on being offered online. Specifically, Law, Medicine, and Psychology require special authorization from regulatory bodies such as the National Bar Association and the National Health Council. To date, no online programs in these fields have been approved. In contrast, disciplines like Business and Education have seen significant growth in online education. In 2019, these two fields accounted for 77% of total online enrollment, in contrast to their 31% share in in-person programs.

2.2. *Data*

Higher education census: This dataset encompasses several layers of information. First, it includes institution-level details, such as ownership status and the parent firm. Second, it captures program characteristics, including detailed categories of the field of study, the delivery mode (online or in-person), the required number of hours for graduation, and the year the degree was introduced. Additionally, it provides information on

[4]See Resoluao CNE/CES N1, de 11 de maro de 2016, Decreto 9.057, de 25 de maio de 2017, and Portaria Normativa 11, de 20 de junho de 2017.

the hubs associated with each online program. Third, it contains student-level data for all enrolled students, along with their demographics, enabling us to track each student's educational path.

Tuition fees: Universities are not required to report tuition fees to the supervising authority, so we rely on four distinct data sources to gather this information. The first two sources come from Brazil's government fellowship and loan programs, PROUNI and FIES. We utilize administrative records from the National Education Fund (FNDE), which track the payments made by the government to students in these programs, allowing us to estimate the tuition fees at participating institutions. The third source is a nationally representative survey conducted by Hoper, a consultancy specializing in higher education. The fourth source is administrative data from QueroBolsa, Brazil's largest degree search platform. In Appendix B.1, we outline the methodology used to combine these sources into a unified tuition price for each degree program. As a result, we are able to recover year-specific tuition prices for approximately 95% of degree-years, covering 98.5% of total enrollment.

Test scores: We have access to detailed data for all students who took the ENEM, Brazil's university entrance exam. This standardized test is high stakes, as it determines eligibility for financial aid and is used for admissions by several universities. The data include scores for each section of the exam, along with responses to a comprehensive socioeconomic background questionnaire. We focus on data from the 2010 exam, which correspond to admissions for the 2011 academic year.

Matched employer-employee records: Finally, we integrate the previously mentioned data sources with matched employer-employee annual administrative records (RAIS) from the Ministry of Labor. RAIS is regarded as a high-quality census of Brazil's formal labor market. This dataset includes detailed worker and firm-level variables, such as salaries, contracted hours, hiring and firing dates, and occupation. We use this data to construct earning profiles for each program and student, covering both online and in-person formats. These profiles allow us to calculate value added measures for each program in the system. The RAIS data we analyze spans the period from 2017 to 2022.

Additional auxiliary sources: We use data from the 2010 Brazilian Population Census to estimate the size of various age cohorts within each region, allowing us to define the potential market for higher education students. We also use administrative data from DSCOM ("Dados do Setor de Comunicaões") from the Ministry of Science and Technology to calculate rates of internet penetration across different regions over time.

2.3. *Data definitions and value added*

2.3.1. *Data definitions*: Next, we provide several definitions for the units of analysis and samples that we use throughout the draft.

*Regions and markets.*A central aspect of our analysis is defining regions that allow us to segment educational markets. Brazil consists of 5,568 municipalities, which we group into 137 meso regions—an administrative division from the Brazilian National Bureau of Statistics (IBGE) that clusters municipalities based on proximity and common features. These meso regions serve as our definition of local markets, which we refer to as *"regions"* throughout the paper. We define a *"market"* as the intersection of a region and a year. The market size is determined by the number of 18- to 45-year-old residents without a college degree living in the region for that year.

Firms and degrees. We define a firm as a company that may own multiple universities, and we use the terms "firms" and "institutions" interchangeably. Throughout this draft, we also use two different definitions of degrees. The first is *degree programs*, which are distinct undergraduate programs that share the exact same degree name within a given university but may differ in the format of instruction (either online or in-person). Its worth noting that an online degree program can be offered across multiple regions. The dataset includes approximately 52,000 unique degree programs in 2010. To reduce the data's dimensionality, we introduce a second definition of degrees by aggregating similar degree programs within the same institution. We define a *degree* as the combination of all degree programs offered by the same institution (i.e., firm), within the same field of study, and taught in the same format (either online or in-person).[5] For example, in our dataset, a degree might include all programs in the field of "Business" (such as Administration, Accounting, Marketing, and Economics) offered by Anhanguera Educacional, a for-profit educational company in Brazil, and taught in person.

Sample. Our analysis centers on the private sector, which accounts for nearly all online programs and roughly 82% of total enrollment, including both online and in-person students. This sector is predominantly non-selective and is often perceived as lower in quality compared to the public sector, which is higher quality and highly selective (Barahona et al., 2023). Consequently, there is limited substitution between the two sectors. Public institutions, which rely entirely on government fundingwhether federal or stateare also less influenced by market forces. To avoid including very small markets in our analysis, we apply the following restrictions: we exclude 27 regions with fewer than 50,000 residents aged 1845, institutions with fewer than 500 students nationwide in any given year, and

[5]In total, we have 11 fields of study based on the International Standard Classification of Education (ISCED) codes. Details on the categorization are provided in Online Appendix A, Table A.1.

degree-region pairs with fewer than five students in any year. After these exclusions, our final sample consists of 110 regions over a 10-year period, covering 470 institutions, 4,101 unique degrees, 15,174 degree-region pairs, and 92,505 degree-region-years.

Value added. An important component of our analysis is computing a quality measure for *degree programs*. We estimate proxies for program quality through value added, using student enrollment, test scores, and wage data. We do this by comparing the labor market earnings of students enrolled in different programs while controlling for their test scores and a wide range of student characteristics. We calculate value added measures, normalized relative to the outside option of not attending college. To account for the varying outside options that students face in their specific local markets, we estimate value added separately for each region.

Specifically, we estimate the following model for each region r:

$$\log(Y_i) = \sum_j \text{VA}_j^r \cdot D_{ij} + X_i'\beta + \varepsilon_i$$

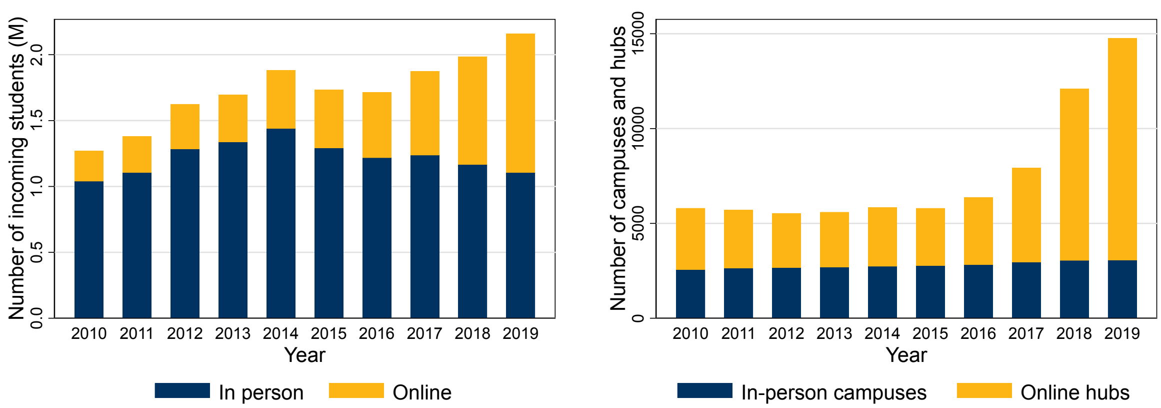
where Y_i represents the wage income of student i, $D_{ij} \in \{0, 1\}$ indicates whether student i is enrolled in degree program j, X_i includes students' characteristics such as gender, age, ENEM score, and a constant, and ε_i captures other determinants of income that are uncorrelated with school enrollment. The sample includes all potential students taking ENEM, even those who did not enroll in any program. The term VA_j^r denotes the region-specific value added of each program, reflecting how program returns vary by local market conditions and the outside option of not attending college. To improve the precision of the estimates for smaller programs, we employ an empirical Bayes procedure, shrinking the estimates following Angrist et al. (2022).

To estimate this model, we use the universe of ENEM 2010 applicants from the 2011 admissions cycle and use student labor market earnings in 2022, 12 years after their enrollment decisions. Because we use only one student cohort, our program value added measure remains constant over time. We are able to estimate value added for 20% of the program-region pairs that exist in the data, which covers 72% of enrollment across all years. Among program-region pairs that existed in 2011, we cover 85% of program-region pairs and 98% of total enrollment. In Section 4, we use the model to fill up the missing data.

2.4. *Descriptives*

2.4.1. *Trends in the Brazilian private-sector higher-education sector*: We present the trends of the overall expansion of online education from 2010 to 2019 in Figure 1. Panel

1(a) shows the enrollment trends of incoming students for both online and in-person education. Between 2010 and 2014, college enrollment steadily increased, with the share of online education remaining relatively stable, representing approximately 23% of total enrollment by 2014. However, following policy reforms around 2016, a notable shift occurred: online education saw significant growth while in-person enrollment began to decline. By 2019, over half of incoming students at private universities were enrolled in online education. Panel 1(b) illustrates similar trends in the number of in-person campuses and online hubs in Brazil. Between 2010 and 2015, the number of campuses and hubs remained comparable, with a marked rise in the number of online hubs following the policy reforms.



(a) Student enrollment

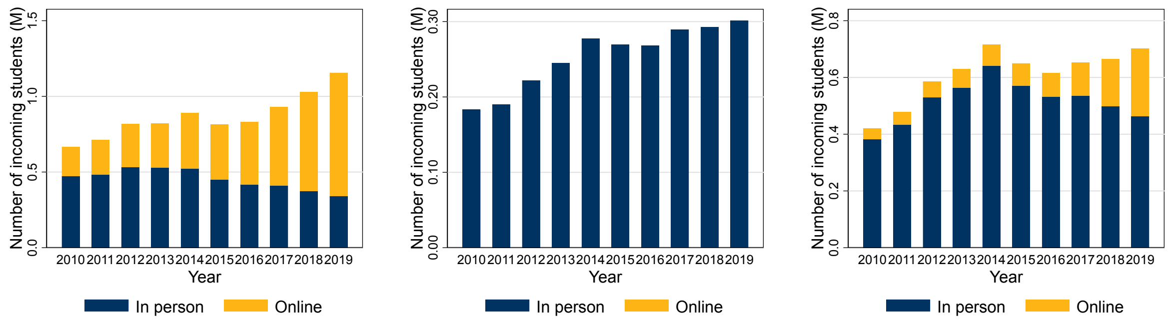
(b) Campuses and hubs

Figure 1: Expansion of online education

Notes:

After establishing the expansion of the online education sector, we highlight that this growth was uneven across different fields of study. As depicted in Figure 2, the most substantial growth occurred in programs related to Business and Education, partly due to their suitability for online delivery. In contrast, fields such as Law, Medicine, and Psychology saw no comparable expansion, as legal restrictions prevent institutions from offering these programs online.

Finally, we highlight that online degree programs are especially popular among older students. There is a noticeable difference in the age distribution of students enrolled in online versus in-person programs. In Figure 3, we track the growth of online and in-person enrollment in Education and Businessthe two fields with the largest online student populationssegmented by age group. While in-person programs are dominated by younger students, online programs attract a majority of students aged 26 and older.

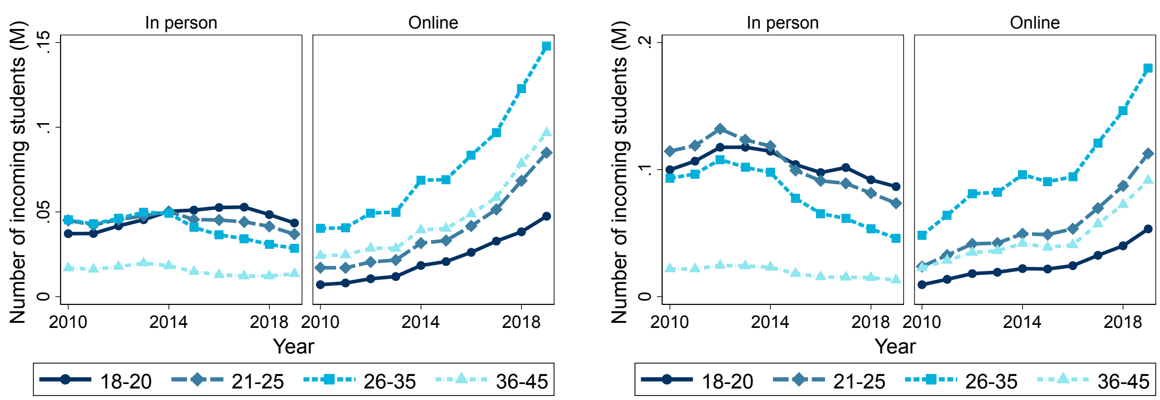


(a) Education & Business (b) Law, Medicine, & Psychology (c) Other areas

Figure 2: Expansion of online education enrollment by categories of field of study

Notes:

By 2019, around 63% of online students in these fields were aged 26 or older, compared to just 30% in in-person programs.



(a) Education (b) Business

Figure 3: Expansion of online education enrollment in Education and Business by age group

Notes:

2.4.2. *Comparison between online and in-person programs*: To examine the practical differences between online and in-person programs, we compare equivalent *degree programs* offered by the same institution, differing only in their mode of delivery. We estimate the following regression:

$$Y_{jrt} = \beta \cdot o_j + \delta_j + \delta_r + \varepsilon_j,$$

where Y_{jrt} is an outcome in degree program j in region r in year t, $o_j \in \{0, 1\}$ indexes whether the program is online, δ_{jt} is a degree program and year-specific fixed effect (e.g., Economics at the Universidade Norte do Paraná in 2011), and δ_r is a region fixed effect. We examine four primary outcomes: (1) the log-number of hours that are mandated to complete the program, (2) the log of the tuition price, (3) dropout rates, and (4) value added as defined in Section 2.3.1. Since our value-added measure does not vary over time, we estimate the corresponding regression using a cross-section of the data for $t = 2011$. All regressions are weighted by the number of students enrolled in program j.

The results are presented in Table 1. Column (1) reports the findings for the log of the required number of hours to complete the degree, showing that both online and in-person programs require a similar number of hours, which is expected as it is mandated by law. In column (2), we observe that online programs are 0.6 log-points less expensive than their in-person counterparts. Column (3) shows that online programs have dropout rates that are XX percentage points higher (over a base of XX) than their in-person counterparts. Lastly, column (4) reveals that the value added of online programs is, on average, 5.7 percentage points lower than that of in-person programs.

Table 1: Comparison between online and in-person programs

	log of total hours (1)	log prices (2)	value added (3)
Online	-0.003	-0.601	-0.057
	(0.005)	(0.072)	(0.007)
Obs.	544920	429412	17540
Mean dep var. (levels)	3220	4430	0.14
Region FE	Yes	Yes	Yes
University × Program name × Year FE	Yes	Yes	Yes

Notes:

These findings shed light on the potential trade-offs between in-person and online education. On one hand, online education can be delivered at a significantly lower cost, resulting in more affordable tuition. On the other hand, it is associated with lower persistence and completion rates, as well as reduced value added compared to in-person programs. Despite these differences, both online and in-person degrees, on average, offer higher value added compared to the alternative of not attending college (see Figure A.1 in the Online Appendix A).

3. The Effects of Online Entry on Market Outcomes

In this section, we estimate the medium-term effects of introducing an additional online degree on various market outcomes using a linear model. To achieve this, we compare changes in outcomes between 2010 and 2019 across regions and fields of study with varying levels of exposure to the growth of online degrees. Throughout this section, we use the *"degree"* definition as outlined in Section 2.3.1.

Specifically, we estimate the following structural equation:

$$\Delta y_{ra} \quad = \quad \phi \Delta N^o_{ra} + \varepsilon_{ra}, \tag{1}$$

where ΔN^o_{ra} denotes the change in the number of online degrees offered in region r and field of study a between 2010 and 2019, and Δy_{ra} represents the change during the same period in the following outcomes of interest: (i) the number of online students relative to market size, (ii) the number of in-person students relative to market size, (iii) the total number of in-person degrees, (iv) the average price of in-person degrees, and (v) the average value added of existing degrees. The error term ε_{ra} captures unobserved shocks at the region-field level that influence the trend of y_{ra}. The parameter ϕ is the key parameter of interest.[6]

Estimating a linear model presents both advantages and limitations. On the one hand, it offers simplicity and transparency, helping to illustrate important correlations that inform the direction and magnitude of relevant causal effects. On the other hand, it relies on strong assumptions. In particular, it imposes a no-interference assumption implying that changes in the number of online degrees offered in region r and field of study a only impact the outcomes in that particular region and field of study, without influencing outcomes in other regions or fields—a condition known as the Stable Unit Treatment Value Assumption (SUTVA). This assumption may be violated if degrees across different fields are close substitutes. In Section 4, we extend our model to account for potential market-level interactions.

We propose two alternative estimation strategies for the linear model, each based on different assumptions for identifying the parameter of interest. First, we outline the assumptions required for a causal interpretation of ϕ when estimating Equation (1) using OLS. Next, we introduce a shift-share instrumental variable approach to address potential threats to the identification in the OLS regression. Finally, we present and compare the results from both estimation strategies.

[6] Equation (1) can be derived by taking differences between 2019 and 2010 of the following structural equation: $y_{rat} = \phi N^o_{rat} + \delta_{ra} + \delta_t + \varepsilon_{rat}$.

3.1. *Ordinary Least-Squares Regression*

The parameters estimated through OLS can be interpreted as causal under the conditional independence assumption $\mathbb{E}[\varepsilon_{ra}|\Delta N^o_{ra}] = 0$. Since Equation (1) is written in differences, this assumption is implied by a strong parallel trends assumption, which states that the trajectory of outcomes for regions and fields of study with lower online degree growth must represent the outcomes in higher-growth regions and fields had they experienced lower online growth (Callaway et al., 2024). Unfortunately, standard tests for this assumption, based on evaluations of pre-trends, are not feasible, as online education was already widespread and growing at the start of our sample.

As an alternative to a pre-trend analysis, it is helpful to examine the differential growth of online education separately across both fields of study and regions. First, fields of study with high online growth should have followed trends similar to those with low online growth. To evaluate this, we take advantage of changes in online growth rates in certain fields of study driven by the 2016 policy reforms and compare their outcomes to areas where online education was prohibited. In Online Appendix C.1, we show that, prior to the policy reforms, all fields of study exhibited similar trends across various outcomes. However, following the reform, some fields expanded more rapidly, leading to divergent trends. These results partially validate the OLS assumptions. Second, we require that institutions choose to enter regions that would have otherwise followed trends similar to those where no entry occurred. This condition may be violated if institutions can anticipate shocks ε_{ra} and strategically enter regions where demand for a particular field is expected to grow. To address this concern, we implement a shift-share instrumental variable approach.

3.2. *Shift-share instrumental variables*

We tackle the issue of endogeneity in institutions' entry decisions by exploiting quasi-random variation in regions' exposure to online entry. Following Goldsmith-Pinkham et al. (2020), we implement a shift-share instrumental variables (SSIV) design—or "Bartik instruments"—with exogenous shares.[7] Our *shares* variable is constructed from a combination of three factors: (i) differences in regions' exposure to potential entrants based on the distance between the regions and institutions' headquarters (ii) institutions' specialization across different fields of study based on their 2010 in-person offerings, and (iii) an indicator variable that captures whether online education is allowed in a given field

[7] Goldsmith-Pinkham et al. (2020) implement SSIV under the assumption that the shares are exogenous. In contrast, Borusyak et al. (2021) and Borusyak and Hull (2023) provide a framework for SSIV where identification is achieved through an exogenous shift.

of study.[8] For the *shift*, we use the rapid expansion of online degree programs, driven by a combination of factors, including growing demand, advancements in online technology, and policy changes.

Building on this, we define the following shift-share instrument:

$$\begin{aligned} z_{ra} &= \sum_f \underbrace{\Delta N_f^o}_{\text{Shift}} \cdot \underbrace{z_{fra}}_{\text{Share}} \\ &\equiv \sum_f \Delta N_f^o \cdot (z_{fr} z_{fa} z_a), \end{aligned} \tag{2}$$

where ΔN_f^o represents the *shift*, capturing the total number of online degrees introduced by institution f between 2010 and 2019. We allow this shift to be correlated with the distribution of shocks ε_{ra}. The variable $z_{fra} \equiv z_{fr} z_{fa} z_a$ corresponds to the exogenous *shares* that predicts the region r and field of study a where institution f is likely to expand. These shares are derived from three sources: z_{fr}, z_{fa}, and z_a, which we detail next. A visualization of each component of the instrument is provided in Figure 4, with each panel explained as the components are introduced.

3.2.1. *Exposure to potential entrants (z_{fr})*: We leverage the differential exposure of regions to institutions' online expansion, driven by their distance from each institution's headquarters. Our empirical analysis shows that institutions are more likely to establish online hubs—and consequently offer degrees—in regions closer to their headquarters, likely due to the lower costs of launching and maintaining nearby operations.[9] We document this pattern through the following regression analysis, using data from 2010:

$$\text{Entered}_{fr} = g(d_{fr})\gamma + \delta_f + \delta_r + \eta_{fr}, \tag{3}$$

where Entered_{fr} equals 1 if institution f had entered region r by 2010 and 0 otherwise; d_{fr} represents the distance between the headquarters of institution f and region r. We define $g(d_{fr}) = [\log(1 + d_{fr}), H_{fr}]$, where $H_{fr} \in \{0, 1\}$ indicates whether the headquarters of institution f are located in region r (i.e., $d_{fr} = 0$), to account for cases where the distance is zero and the log transformation of the distance would be undefined. Our results show a negative and significant relationship between distance and the probability

[8]As discussed in Section 2.1, regulation prohibits the fields of Law, Medicine, and Psychology from offering remote education.

[9]These lower costs are primarily due to easier management and coordination enabled by proximity, reduced setup and operational expenses, and a deeper understanding of local needs and challenges near the headquarters.

of entry. Specifically, a 10% increase in the distance between a region and an institution's headquarters reduces the probability of entry by 0.88%. More detailed results can be found in Online Appendix A, Table A.2.

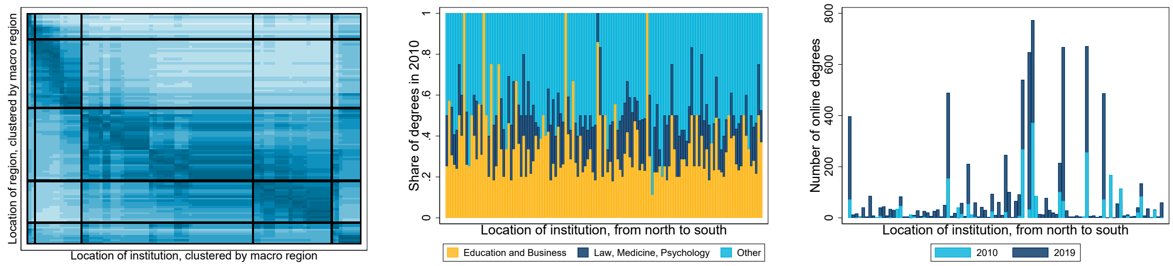
We use the estimates from Equation (3) to determine the likelihood that institution f will open an online hub in region r, based solely only the distance between the institution and the region. This likelihood is calculated by predicting entry using distance while excluding fixed effects, thereby removing potentially endogenous variation from unobserved regional characteristics. We then normalize the likelihood so that it sums to one across all regions.[10] Our measure of exposure is defined as:

$$z_{fr} \quad = \quad \frac{g(d_{fr})\hat{\gamma}}{\sum_r g(d_{fr})\hat{\gamma}}. \tag{4}$$

Figure 4(a) illustrates the exposure instrument, z_{fr}. The y-axis represents each of the 110 regions in our sample, organized by the five macro-regions defined by the IBGE, with black lines separating each macro-region. Within each macro-region, regions are arranged from northwest to southeast, based on the IBGE encoding system. On the x-axis, each column corresponds to one of the 93 institutions that expanded their online presence between 2010 and 2019, ordered by the location of their headquarters, using the same strategy as for the regions. Each region-institution pair is shaded in blue, with lighter shades indicating lower exposure, or a lower likelihood that institution f will enter region r, and darker shades indicating higher exposure. We observe that exposure tends to be higher for region-institution pairs that are geographically close and within the same macro-region.

3.2.2. *Institutions' propensity to expand in different fields of study (z_{fa})*: The previous subsection focuses on estimating the likelihood that each institution enters a given region. However, different institutions might have different likelihoods of expanding into different fields of study. To address this, we use institutions' 2010 offerings to predict the fields in which they are most likely to expand. The rationale is that some institutions may specialize in certain areas of study, making it easier for them to expand in those fields compared to others. Since many institutions offered few or no online degrees in 2010, we use the intensity of their in-person instruction to predict where they are most likely to expand online. We estimate the likelihood of expansion in each field based on the share of in-person degrees offered in that field in 2010. Specifically, we calculate:

[10] Equation (4) can be interpreted as the probability that institution f opens an online hub in region r, assuming it opens exactly one hub, with the probability depending only on the distance between the institutions headquarters and the region.



(a) Likelihood that institution f opened an online hub in region r by $t = 2010$

(b) Share of in-person degrees across field of study by firm

(c) Total number of online degrees by institution in 2010 and 2019

Figure 4: Expansion of online education enrollment by categories of field of study

Notes:

$$z_{fa} \quad = \quad \frac{\sum_r N^{\iota}_{fra,2010}}{\sum_a \sum_r N^{\iota}_{fra,2010}}, \tag{5}$$

where $N^{\iota}_{fra,2010}$ takes the value of 1 if institution f offers a degree from field of study a in region r in 2010.

Figure 4(b) illustrates the resulting shares by institution, aggregated into broader groups of study areas. On the x-axis, each column represents one of the 93 institutions that expanded their online presence between 2010 and 2019, ordered by the location of their headquarters, as in Panel 4(a). We observe that some institutions concentrated more on fields like Education and Business, while others prioritized areas such as Engineering or Math. Our instrument predicts more intense online expansion in field of study a in regions with nearby institutions that were already specializing in that area of study in 2010.

3.2.3. *Areas of study regulatory constraints (z_a)*: Throughout our sample, regulations prohibit online education in Law, Medicine, and Psychology. We use this constraint to build an indicator variable, z_a, which equals 1 for areas of study allowed to expand online and 0 for those that are not. We incorporate z_a into Equation (2), ensuring that our instrument predicts zero online growth in areas of study such as Law, Medicine, and Psychology.

3.2.4. *Institutions' expansion of online degrees (ΔN^o_f)*: The final component of the instrument in Equation (2) is the shift term. We estimate ΔN^o_f by calculating the total number of online degrees that each institution opened in any region or field of study between 2010 and 2019:

$$\Delta N_f^o = \sum_r \sum_a \left(N_{fra,2019}^o - N_{fra,2010}^o \right), \tag{6}$$

where $N_{fra,t}^o$ takes the value of 1 if institution f is offers a degree in field of study a in region r in year t.[11]

Figure 4(c) illustrates the variation in overall online degree expansion across institutions. On the x-axis, each column represents one of the 93 institutions that expanded online between 2010 and 2019, organized by the location of their headquarters, similar to Figure 4(a). There is significant heterogeneity across institutions. Our instrument predicts more intense online expansion in regions located near institutions that expanded more aggressively.

3.3. *Identification*

As is standard in linear models using instrumental variables, two assumptions must hold for the correct interpretation of causal effects. First, the instrument needs to be relevant, meaning it must have predictive power over the endogenous variable. We test this by conducting a first-stage regression between the SSIV, z_{ra}, and the endogenous variable, $\Delta N_{ra}^0$the change in the number of online degrees offered in region r and field $a$yielding an F-statistic of 307. The coefficient from this regression is presented in Table 2, Panel B, Column (1). Second, the exclusion restriction must hold. Specifically, the exposure shares, z_{fra}, must be uncorrelated with the structural error term, ε_{ra}. Formally, this requires that $\mathbb{E}[z_{fr} z_{fa} z_a \varepsilon_{ra}] = 0$ for all f where $\Delta N_f^o \neq 0$.

To build intuition around the identification assumption, it helps to define $\bar{\epsilon}_{fr} = \mathbb{E}[z_{fa} z_a \varepsilon_{ra} | fr]$, which represents the propensity of institution f to expand in region r due to potentially anticipated shocks, ε_{ra}. For instance, if institution f specializes in business degrees in 2010, and region r is expected to see an increase in demand for online business degrees, institution f would be more likely to offer online degrees in that region. Using the law of iterated expectations, the identification assumption can be reformulated as $\mathbb{E}[z_{fr} z_{fa} z_a \varepsilon_{ra}] = \mathbb{E}[z_{fr} \bar{\epsilon}_{fr}] = 0$. This condition holds if the distance between regions and a given institution is uncorrelated with region-specific unobserved demand shocks, ε_{ra}, related to the field of study in which the institution specializes in in-person education.

In Online Appendix C.2, we explore the validity of the identifying assumption by

[11] Our results are robust to using a leave-one-out estimate of ΔN_f^o. Because our identification strategy relies on having exogenous shares, a leave-one-out estimator is not necessary in this setting.

estimating the model on different sub-samples. We show that the strength of the first stage varies between early and later periods. Nevertheless, the coefficient of the reduced-form regression scales proportionally to the first stage, consistent with a situation where the instrument affects the outcome only through changes in the endogenous variable.

3.4. *Results*

Table 2 presents the results for both identification strategies. Panel A displays the OLS results, while Panel B shows the SSIV results. We find that the OLS and IV strategies deliver very similar results. We interpret this as evidence that endogeneity concerns are minimal, as institutions might have difficulty anticipating future demand shocks when deciding to expand. This suggests that most expansion decisions are driven by other factors, such as costs—which are captured by our shift-share instrument—and further influenced by the 2016 policy reforms that facilitated rapid expansion to low-cost locations. Given the similarity of the results, we focus our discussion on the SSIV, though the conclusions are the same for OLS.

3.4.1. *Average effects*: We begin by analyzing the impact of introducing an additional online degree in a specific field and region on both online and in-person enrollment. These results are presented in Columns (2) and (3), respectively, with outcomes reported relative to market size, defined by the number of individuals aged 18 to 45 without a college degree in that region. Column (2) shows that each additional online degree introduced between 2010 and 2019 increased online enrollment by 0.372 students per 1,000 individuals in the market, which is equivalent to 52% of the average 2010 online enrollment, a year when there were, on average, 2.82 online degrees per region and field of study. Column (3) reveals that the increase in online degrees led to a reduction in in-person enrollment by 0.2 students per 1,000 individuals, or 6.5% of the average 2010 in-person enrollment. Combining these findings, we observe that total enrollment rose by 0.172 students per 1,000 individuals in the relevant market, which represents a 4.5% increase relative to the 2010 baseline enrollment.

These results reveal two opposing forces at play: online degrees expand the market by attracting new students to college, while simultaneously diverting students away from in-person programs. Our findings show that for each additional online student, 46% are new to higher education, while 54% would have otherwise enrolled in an in-person degree. These two forces create ambiguous effects on total value added. Market expansion increases value added, as newly enrolled students enter programs with positive value compared to the outside option of no college, which we normalize to have zero value added. However, market diversion reduces value added by shifting students from higher value-added in-person programs to lower value-added online alternatives.

Using the average value added differences between online and in-person degrees, a back-of-the-envelope calculation suggests that each additional online degree increases value added by 0.0076 percentage points per student enrolled in college in 2010 due to the rise in online enrollment, However, the corresponding decline in in-person enrollment reduces total value added by 0.0086 percentage points per student enrolled in college in 2010. These effects cancel out, resulting in a reduction in total value added of 0.0010 percentage points for each student enrolled in college in 2010.[12]

Table 2: Effects of introducing an additional online degree

	Δ in online degrees (1)	Δ in online students (2)	Δ in in-person students (3)	Δ in in-person degrees (4)	Δ in log-price of in-person degrees (5)
Panel A: OLS regression					
Δ in online degrees		0.344 (0.033)	-0.185 (0.026)	-0.133 (0.029)	-0.009 (0.002)
Panel B: IV regression					
shift-share instrument	1.785 (0.102)				
Δ in online degrees		0.372 (0.035)	-0.200 (0.028)	-0.157 (0.039)	-0.012 (0.002)
Panel C: Average value of the dependant variable in levels in 2010 and 2019					
2010	2.82	0.72	3.07	10.83	5.66
2019	9.48	2.79	2.94	12.18	6.22
Obs.	1210	1210	1210	1210	954

Notes:

We next analyze the consequences of the online expansion on the availability and pricing of in-person degrees. In Column (4), we show that the number of in-person degrees decreased in regions and fields with larger online growth. Specifically, for each additional online degree, there are 0.157 fewer in-person degrees. These effects stem from both an increase in degree exits and entry deterrence, which exacerbates the diversion from in-person degrees to online alternatives. As a result, even students who prefer to enroll in a specific in-person program may be forced to switch to an online alternative when

[12] We calculate changes in total value added as follows: $\Delta \text{VA} = \frac{\beta^o \text{VA}^o + \beta^\iota \text{VA}^\iota}{s^o + s^\iota}$, where β^o and β^ι are the coefficients from Columns (2) and (3) of Table 2, Panel B. Here, VA^o and VA^ι represent the average value added of online and in-person degrees in 2010, given by 0.078 and 0.164 percentage points, respectively. Finally, s^o and s^ι denote the average number of online and in-person students attending college in 2010 for every 1,000 individuals aged 18 to 45 living in that region, as reported in Columns (2) and (3) of Table 2, Panel C.

their preferred in-person option exits the market. Finally, in Column (5), we show that the average price of in-person degrees dropped by 1.2% for each additional online degree introduced in that region and field of study. This supports the idea that online degrees intensify local competition, driving down prices and deterring new in-person program entry.

3.4.2. *Effects by age groups*: The expansion of online education may affect age groups differently. Evidence shows that online programs tend to attract older individuals, who often face greater challenges and costs in attending in-person classes (Goodman et al., 2019; Aucejo et al., 2024). In our setting, by 2010, individuals aged 36-45 were 90% less likely to attend an in-person degree and 21% more likely to attend an online degree than those aged 18-20. This suggests online education can expand access for older cohorts with minimal risk of diverting them from in-person alternatives. Younger cohorts, however, are more likely to already attend in-person programs. By 2010, 6.3 per 1,000 individuals aged 18-20 were enrolled in in-person programs, ten times the number enrolled online. For these younger groups, online options may pose a greater risk of diversion, as they offer an attractive, low-cost alternative. If the quality of online degrees is lower and hard to assess, this shift could be detrimental to them.

In Table 3, we analyze enrollment changes across different age cohorts. We find high rates of market diversion among younger cohorts. For students aged 18-21, each additional online enrollment (Column 1, Panel B) leads to nearly a one-to-one substitution of students leaving the in-person sector (Column 5, Panel B), indicating that online education acts as a substitute rather than expanding the market. For students aged 21-25, 0.57 students leave in-person programs for every new online student (ratio of the coefficient in Column 6 to Column 2, Panel B). For older cohorts (ages 26-35 and 36-45), this substitution drops to 0.43 and 0.29 students, respectively (calculated similarly using Panel B coefficients). Therefore, online education proves to be an effective tool for expanding the market and increasing college access for older cohorts.

Overall, the results from Tables 2 and 3 indicate that online education has the potential to expand access to higher education, especially among older cohorts. However, online education also diverts students from in-person alternatives. Increased competition drives down in-person tuition, reduces new program entry, and increases program exits. Consequently, the average value added declines, and students who would prefer in-person learning may find themselves pushed into online alternatives as their preferred programs disappear.

This section's analysis relies on a linear model that assumes no interaction between

Table 3: Effects of introducing an additional online degree on enrollment by age groups

	Δ in online students				Δ in in-person students			
	18-20	21-25	26-35	36-45	18-20	21-25	26-35	36-45
Panel A: OLS regressions by cohort								
Δ in online degrees	0.333 (0.044)	0.414 (0.043)	0.370 (0.036)	0.263 (0.024)	-0.325 (0.057)	-0.229 (0.036)	-0.167 (0.030)	-0.079 (0.012)
Panel B: IV regressions by cohort								
Δ in online degrees	0.394 (0.056)	0.462 (0.046)	0.414 (0.036)	0.301 (0.024)	-0.395 (0.071)	-0.263 (0.034)	-0.179 (0.028)	-0.088 (0.012)
Panel C: Average value of the dependant variable in levels in 2010 and 2019								
2010	0.67	0.75	0.96	0.81	6.31	2.67	1.32	0.63
2019	3.50	3.08	3.05	2.55	8.19	2.59	1.01	0.55
[0.5em] Obs.	1210	1210	1210	1210	1210	1210	1210	1210

Notes:

degrees in different fields within the same region and overlooks supply-side equilibrium responses. To address cross-field substitution, we need a more flexible model that accounts for substitution across areas of study. Moreover, the linear approach is unsuitable for out-of-sample counterfactuals, where competition, price changes, and entry/exit decisions can have nonlinear effects. In the next section, we develop an equilibrium model of demand and supply to capture substitution between online and in-person degrees, estimate entry and exit elasticities, and conduct counterfactual analysis on the impact of online education expansion in Brazil and possible policy interventions to enhance total value added.

4. Model

We now develop and estimate an equilibrium model for the Brazilian in-person and online college education market. Our goal is three-fold. First, we measure substitution patterns between in-person and online degrees and understand how the proliferation of online degrees affects prices and market structure. Second, we use the model to learn how online education can affect students' expenditure and education value-added. On one hand, online degrees are more flexible and cheaper. On the other hand, some of them may provide lower value-added, affecting students' choices associated with different long term outcomes. Finally, we use the model to run counterfactual analysis and study what would have happened if the government did not allow for the expansion of online education or allowed it further in fields in which it is still forbidden.

4.1. *Setup*

The model consists of institutions (i.e., firms) $f \in \mathcal{F}$ that offer undergraduate college education in Brazil. Students can choose to major in a particular field of study $a \in \mathcal{A}$ by enrolling in either an in-person or online degree. A degree is defined as the combination of the firm offering it, the field of study it belongs to, and whether it is in-person or online, and is represented by $j \in \{\mathcal{F} \times \mathcal{A} \times \{0, 1\}\}$. Geographical regions are denoted by $r \in \mathcal{R}$ and years by $t \in \mathcal{T}$. A single degree j can be available in multiple regions, and a region may offer multiple degrees. Each firm offers a product bundle $\mathcal{J}_{frt} \in \mathcal{B}_f$ in each region and year, where $\mathcal{J}_{frt} = 0 \in \mathcal{B}_f$ for all f represents the option of not offering any degree.

We model the equilibrium as a static game. At the beginning of the period, institutions know the pre-existing market structure $\mathcal{J}_{frt_0}$. The timing of the game is as follows. First, institutions learn the fixed costs of offering different bundles in a given region FC_{fr}, which depend on the bundles composition, the current offerings, the distance between the institution's headquarters and the region, and a private, idiosyncratic shock. Next, institutions decide which bundle to offer in each region or whether to offer one at all, $\mathcal{J}_{frt}$. Afterward, demand and supply shocks are realized, and institutions set tuition prices for each degree in their bundle. Finally, students decide where to enroll. The timeline is summarized in Figure 5.

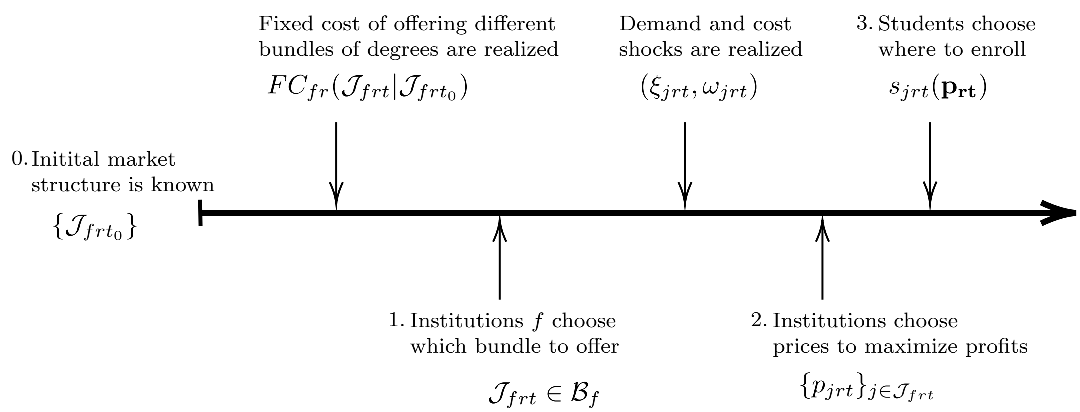


Figure 5: Model timeline

In the following sections, we present the components of the model in reverse order. First, we outline the demand model. Second, we discuss the institutions' pricing decisions. Third, we detail the institutions' optimal bundle choice.

4.2. *Students' choice of where to enroll*

In each year t and region r, a potential student, i, decides whether to enroll in a degree $j \in \mathcal{J}_{rt}$, where $\mathcal{J}_{rt}$ are the degrees available in year t and region r, or not to enroll at all. A degree j is characterized by the institution f that offers it, the area of study a it belongs to, and a vector of characteristics $x_j = [x_j^{(1)}, x_j^{(2)}]$, where $x_j^{(1)} = [\iota_j, o_j]$ contains indicator variables for in-person and online instruction, and $x_j^{(2)}$ include a constant, the age of the program, the average score of incoming students, the average wages of graduate students, the length of the program (in number of hours required), and the degree's stem load.[13]

The utility student i gets from enrolling in degree j in region r in year t is:

$$u_{ijrt} \quad = \quad -\alpha_i p_{jrt} + x_j^{(1)} \beta_i + \delta_{jrt} + \epsilon_{ijrt}, \tag{7}$$

where p_{jtr} denote the tuition price and $x_j^{(1)}$ indicates whether the degree is online or in-person. The associated coefficients vary with individual characteristics, assuming $[\log(\alpha_i), \beta_i]' \sim \mathcal{N}(\mu_{b(i)}, \Sigma)$, where $\mu_{b(i)}$ depends on the student's age bin $b(i)$. The term δ_{jrt} captures degree-market-specific characteristics constant across individuals, while ϵ_{ijrt} represents a consumer-specific demand shock following a generalized extreme value distribution, consistent with a nested logit model. The nests are defined by the degree's area of study, and the intra-nest correlation is denoted by ρ.

We further decompose the degree-market-specific utility as follows:

$$\delta_{jrt} \quad = \quad \psi w_{jrt} + \delta_j + \delta_{ra} + \delta_{ta} + \delta_{to_j} + \xi_{jrt}, \tag{8}$$

where w_{jrt} is a demand shifter based on region r's internet penetration in year t interacted with whether degree j is in person or online. The term δ_j capture degree-specific components of utility, δ_{ra} captures area of study and region-specific factors, δ_{ta} represents area of study and year-specific components, δ_{to_j} account for online-year factors, allowing for differential yearly demand shifts between online and in-person education. Finally, ξ_{jrt} denotes a degree-region-year-specific idiosyncratic demand shock.

We denote s_{jrt} as the share of potential students from region r in year t who choose to enroll in degree j, where

$$s_{jrt}(\mathbf{p_{rt}}) \quad = \quad \int_{i \in I_{rt}} \mathbb{1}\{u_{ijrt} \geq u_{ikt}, \;\; \forall k \in \mathcal{J}_{rt}\} di, \tag{9}$$

[13]Degrees are either in person or online but not both (i.e., $x_j^{(1)} \in \{[1,0], [0,1]\}$).

with $\mathcal{J}_{rt}$ the set of degrees available in region r and year t, and $\mathbf{p_{rt}}$ as the market vector of degree prices.

4.2.1. *Identification and estimation*: We estimate the demand model using the generalized method of moments introduced by Berry et al. (1995), combining both instrumental variables and micro-moments to identify the model parameters.[14]

Instruments for prices. A key challenge in demand estimation is price endogeneity, as institutions may set prices in response to unobserved demand shocks ξ_{jrt}. Ideally, we would rely on cost-shifters that affect pricing decisions but are independent of demand shocks. When these shifters are difficult to observe, proxies serve as a practical alternative. Following Hausman et al. (1994), we use the contemporaneous prices of the same degree in other regions as a proxy. The intuition behind this instrument is that variation in firm-level costs impacts prices across all markets where the degree is offered. The primary identification assumption is that while costs for degree j are correlated across regions, demand shocks are not.

Instruments for substitution patterns. The demand model incorporates individual heterogeneity through random coefficients on α_i and β_i and the nesting parameter ρ. To identify these parameters, we use instruments that shift the competitor mix—understood as firms offering degrees in a given market—but are unrelated to demand. To construct these instruments, we employ a shift-share design, as outlined in Section 3. We build exposure measures for firm-level entry of new hubs and campuses, following the method in Section 3.2.1, and denote these as z^o_{fr} for hubs and z^ι_{fr} for campuses. We measure institutions' propensity to expand across areas of study following Section 3.2.2, and denote it by z_{fa}. We construct firm-level expansion measures by mode of delivery, as outlined in Section 3.2.4, denoted by N^o_{ft} for online and N^ι_{ft} for in-person. These measures are used to develop three sets of market-level instruments:

$$z^o_{rt} = \sum_a \sum_f z^o_{fr} z_{fa} z_a N^o_{ft} \;\; , \;\; z^\iota_{rt} = \sum_a \sum_f z^\iota_{fr} z_{fa} N^\iota_{ft} \;\; , \;\; z^a_{rta} = \sum_f (z^o_{fr} z_{fa} z_a N^o_{ft} + z^\iota_{fr} z_{fa} N^\iota_{ft}).$$

The first two instruments, z^o_{rt} and z^ι_{rt}, shift the total number of online and in-person degrees available in region r and year t, respectively, providing variation to identify the random coefficient on $x^{(1)}_j$, the indicator for online or in-person degrees. The third instrument, z^a_{rta}, shifts the total number of degrees offered in each area of study, generating variation to identify the nesting parameter ρ. Finally, following Gandhi and Houde (2019),

[14]For a comprehensive review of related literature, see Berry and Haile (2016), and for a detailed guide on best practices, including those we adopt, refer to Conlon and Gortmaker (2020).

we construct differentiation instruments by applying functions of our price instruments. These generate variation that helps identify the variance of the random coefficient on price.

Micro-moments for age heterogeneity. We incorporate additional micro-moments to discipline the age heterogeneity parameters, $\mu_{b(i)}$, by matching moments predicted by the model with their empirical counterparts. In total, we use eight moments, defined by $\mathbb{E}[i$ attends an in-person degree$|i \in b]$ and $\mathbb{E}[i$ attends an online degree$|i \in b]$, where b are the four age bins described in Section 3.

Identification of components in the linear part of the utility. The estimator proposed by Berry et al. (1995) produces consistent estimates of $\mu_{b(i)}$, Σ, ψ, and δ_{jrt}, which together recover the distribution of α_i, β_i, market shares, and price elasticities. However, the estimator treats the components of δ_{jrt} as nuisance parameters and does not provide consistent estimates for them individually. While this is typically not an issue—since market shares and price elasticities only depend on the nuisance parameters through δ_{jrt}—in our setting, we are interested in recovering market shares and price elasticities of degree that may not be offered in certain markets in the data. This requires estimates for each component of δ_{jrt}. To estimate them, we impose additional structure and estimate a mixed-effects Bayesian hierarchical model. Specifically, for each of the components in Equation 8, we assume $\delta_j \sim \mathcal{N}(0, \sigma_j^2)$, $\delta_{ra} \sim \mathcal{N}(0, \sigma_{ra}^2)$, $\delta_{ta} \sim \mathcal{N}(0, \sigma_{ta}^2)$, $\delta_{to} \sim \mathcal{N}(0, \sigma_{to}^2)$, and $\xi_{jrt} \sim \mathcal{N}(0, \sigma_\xi^2)$, and δ a non-random parameter. We estimate the model via maximum likelihood to recover unbiased estimates of each component of δ_{jrt} that we use to project utility into degree-markets that are offered and not offered in the data.

4.2.2. *Results*: We present the estimated parameters in Online Appendix A, Table A.3. To summarize our results, we present median own-price elasticities and diversion ratios for 2010 and 2019 in Table 4. By 2019, we find that when the price of an in-person degree marginally increases, 64 % of students who leave that degree would enroll in another in-person program, 20 % would switch to an online program, and 14 % would exit higher education altogether. Additionally, we find that younger cohorts show stronger preferences for attending college relative to the outside option, and that such age gradient is stronger for in-person degrees (see Online Appendix A, Table A.3). Finally, our findings indicate that higher internet penetration increases demand for both online and in-person degrees.

Table 4: Elasticity and diversion ratios

	$t = 2010$		$t = 2019$	
	In-person	Online	In-person	Online
Median own-price elasticity:	-3.32	-1.52	-3.53	-1.49
Median diversion ratios:				
To in-person:	0.76	0.35	0.64	0.15
To online:	0.04	0.47	0.2	0.7
To outside good:	0.15	0.15	0.14	0.15

Notes:

4.3. *Institutions' pricing decision*

Once institutions decide their degree offering and demand shocks ξ_{rt} and supply shocks ω_{rt} shocks are realized, they set tuition prices to maximize profits.[15] The profits of institution f, given its degree offerings $\mathcal{J}_{frt}$, are represented by:

$$\pi_{frt}(\mathcal{J}_{frt}) = \max_{\{p_{jrt}\}_{j \in \mathcal{J}_{frt}}} \sum_{j \in \mathcal{J}_{frt}} (p_{jrt} - c_{jrt}) \cdot s_{jrt}(\mathbf{p}_{rt}). \tag{10}$$

The log of marginal cost c_{jrt} is given by:

$$\log(c_{jrt}) = \gamma_z z_{jrt} + \underbrace{x_j^{(2)} \gamma_x + \gamma_j + \gamma_{ra} + \gamma_{ta} + \gamma_{to} + \omega_{jrt}}_{\gamma_{jrt}}, \tag{11}$$

where z_{jrt} are a set of cost-shifters that include the log-distance between institution f main campus and region r, an indicator on whether region r corresponds to the location of the main campus, and other price instruments used for demand estimation, $x_j^{(2)}$ are the degree characteristics from Equation (??) that include a constant, the age of the program, the average score of incoming students, the average wages of graduate students, the length of the program (in number of hours required), and the degree's stem load, γ_j are degree specific components of the cost function, γ_{ta} are area of study-year specific components, γ_{to} are online-year specific components that allow for differential trends in demand for online and in-person education, and ω_{jrt} is a degree-region-year specific idiosyncratic supply shock. Institutions compete a-la-Bertrand.[16]

[15] We consider the assumption that firms maximize profits to be reasonable in this context, especially since many of them are publicly traded.

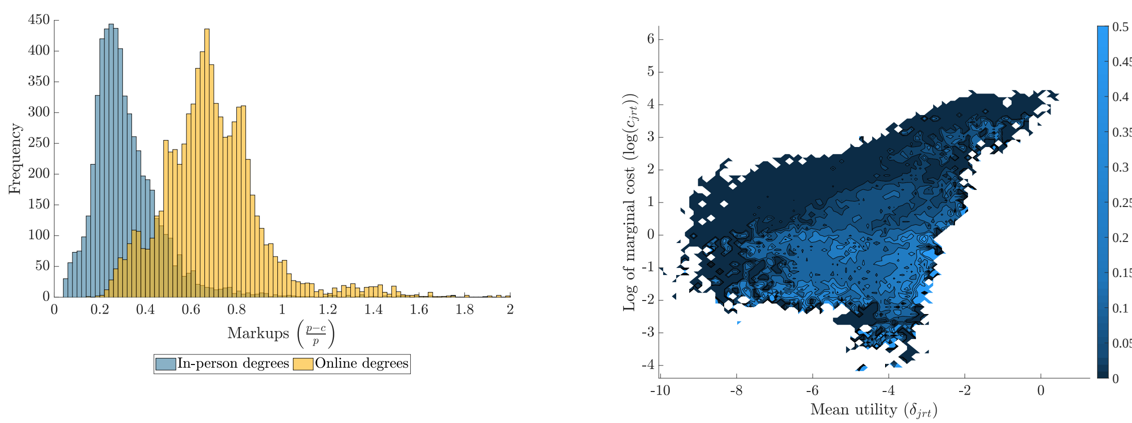
[16] Note that profits also depend implicitly on other institutions' offerings of degrees, $\mathcal{J}_{f'rt}$.

4.3.1. *Estimation*: We estimate the cost parameters in three steps. First, we recover c_{jrt} for all degrees offered in the data by inverting the firms' first order conditions. Second, we estimate γ_z via a fixed effects OLS model. Third to recover the components of γ_{jrt} we estimate a mixed-effects Bayesian hierarchical model where

$$\gamma_{jrt} \quad = \quad x_j^{(2)}\gamma + \gamma_j + \gamma_{ra} + \gamma_{ta} + \gamma_{to} + \omega_{jrt}, \tag{12}$$

with $\gamma_j \sim \mathcal{N}(0, \varsigma_j^2)$, $\gamma_{ra} \sim \mathcal{N}(0, \varsigma_{ra}^2)$, $\gamma_{ta} \sim \mathcal{N}(0, \varsigma_{ta}^2)$, $\gamma_{to} \sim \mathcal{N}(0, \varsigma_{to}^2)$, and $\xi_{jrt} \sim \mathcal{N}(0, \varsigma_{\omega}^2)$, and γ a non-random parameter. We estimate the model via maximum likelihood to recover unbiased estimates of each component of γ_{jrt} that we use to project utility into degree-markets that are offered and not offered in the data.

4.3.2. *Results*: We present the distribution of markups—defined as the ratio of price minus marginal cost to price—for in-person and online degrees in Figure 6(a). We present the corresponding estimated parameters in Online Appendix **??**, Table A.4. We find that institutions price their online degrees at a more inelastic part of the demand curve making their markups higher. However, despite having larger markups, online degrees' prices are lower due to lower marginal costs. The average price of in-person and online degrees is 5.662 and 1.822 thousand dollars per year, respectively, and the average estimated marginal cost is 4.213 and 0.666 thousand dollars per year, respectively.



(a) Distribution of markups

(b) Probability of a degree being offered in each region in 2019

Figure 6: Estimated parameters from demand and pricing decision

Notes:

We also find that degrees that face higher demand—through a higher value of δ_{jrt} from Equation (**??**)—and a lower marginal cost, are more likely to be offered in multiple regions in 2019. Figure 6(b) shows a heatmap with the relationship between degrees' demand, marginal cost, and their empirical probability of being offered in 2019. This finding suggests that institutions are choosing more profitable degrees when deciding which bundles to offer.

4.4. *Institutions' choice of degrees' offerings*

Firm f's fixed cost of offering a given bundle bundle, $\mathcal{J}_{frt}$ in region r in $t = 2019$ depends on three key components: (i) the composition of the bundle and whether degrees offered are existing degrees or new degrees, (ii) the infrastructure investments required to offer the degrees in the bundle that depend on the requirements of having an in-person campus or an online hub, on whether such infrastructure existed already, and the distance between firm f's headquarters and region r, and (iii) a private information cost shock only known by firm f. We parametrize the cost function by:

$$FC_{fr}(\mathcal{J}_{frt}|\mathcal{J}_{frt_0}) = \text{Degrees}(\mathcal{J}_{frt}|\mathcal{J}_{frt_0}) + \text{Infrastructure}(\mathcal{J}_{frt}|\mathcal{J}_{frt_0}) + \sigma_\varepsilon \varepsilon_{fr\mathcal{J}_{frt}} \tag{13}$$

where

$$\text{Degrees}(\mathcal{J}_{frt}|\mathcal{J}_{frt_0}) \;\;=\;\; \kappa_0 + \sum_{k\in\mathcal{J}_{frt}\cap\mathcal{J}_{frt_0}} \kappa_k^{old} + \sum_{k\in\mathcal{J}_{frt}\setminus\mathcal{J}_{frt_0}} \kappa_k^{new} \tag{14}$$

is the cost associated to keeping degrees open and opening new degrees that vary at the area of study level.

$$\text{Infrastructure}(\mathcal{J}_{frt}|\mathcal{J}_{frt_0}) \;\;=\;\; \sum_{k\in\{\text{campus},\text{hub}\}} New_{frk} f_k(d_{fr}) \tag{15}$$

is the cost associated to opening a new campus or hub necessary to host a given degree. New_{frk} for $k \in \{\text{campus}, \text{hub}\}$ is an indicator variable that takes the value of 1 if $\mathcal{J}_{frt}$ contains an in person or online degree, respectively, but $\mathcal{J}_{frt_0}$ does not, and $f_k(d_{fr}) = \chi_k^0 + \chi_k^1 \mathbb{1}\{d_{fr} = 0\} + \chi_k^2 \log(1 + d_{fr})$ is a function of the distance between firm f's heardquarters and region r, d_{fr}. Finally, $\varepsilon_{fr\mathcal{J}_{frt}}$ is a firm-region-bundle specific idiosyncratic shock that is only observed by firm f and that we assume to follow an extreme value type I distribution. We restrict κ_k^{old}, κ_k^{new}, and $f_k(d_{fr})$ to be all greater than zero so that offering degrees is more expensive than not offering them, and building new infrastructure is more expensive than using existing one.

Before observing the demand and supply shocks, ξ_{jrt} and ω_{jrt}, but after learning the region-bundle specific idiosyncratic shocks, $\varepsilon_{fr\mathcal{J}_{frt}}$, firms make their choices of what degrees to offer. The firm's problem is given by

$$\Pi_{frt} \quad = \quad \max_{\mathcal{J}_{frt} \in \mathcal{B}_f} \quad \mathbb{E}[\pi_{frt}(\mathcal{J}_{frt})] - FC_{fr}(\mathcal{J}_{frt}|\mathcal{J}_{frt_0}) \tag{16}$$

where $\pi_{frt}(\mathcal{J}_{frt})$ is given by Equation (10) and the expectation is taken over the distribution of all other firms' idiosyncratic shocks, $\varepsilon_{f'r\mathcal{J}_{f'rt}}$, and over the demand and supply shocks, ξ_{jrt} and ω_{jrt}.

4.4.1. *Estimation*: To estimate the fixed cost parameters, we exploit the firms' bundle choices across different markets. We follow Sweeting (2009) and estimate the parameters in two steps. In the first step, we compute choice probabilities directly from the data. Specifically, we estimate a multinomial choice logit model using all own and competitors' drivers of variable profits, which include the demand mean utilities, demand shifters, and marginal costs components of goods belonging to each bundle, $\{\delta_{jrt}, w_{jrt}, \gamma_{jrt}, z_{jrt}\}_{j\in\mathcal{J}_{frt}}$, and the determinants of fixed costs from Equation (13). We use the multinomial logit model to estimate the probability that each bundle is offered in the data. We denote the resulting estimated probability distribution by $\hat{\Phi}$.

We use the estimated probability distribution $\hat{\Phi}$ to estimate expected variable profits, $\mathbb{E}[\pi_{frt}(\mathcal{J}_{frt})]$, from Equation (16). We randomly draw own and competitors' bundle choices, $\{\mathcal{J}_{frt}, \mathcal{J}_{f'rt}\}$, from the estimated probability distribution, $\hat{\Phi}$, and demand and supply shocks, ξ_{jrt} and ω_{jrt} from their respective empirical distributions. For each draw, we compute variable profits solving for the static equilibrium game given by Equation (10). We do this 10,000 times. Then, we compute expected variable profits for each bundle choice by integrating over all simulation draws. Finally, we fit a random forest prediction model to reduce noise and get stable predictions for bundles that have low probability of being chosen.

In the second step, we use the estimated expected variable profit functions and the equilibrium conditions on entry probabilities to recover the fixed cost parameters. From Equation (16), the probability that bundle $\mathcal{J}_{frt}$ is chosen is given by

$$\phi_{\mathcal{J}_{frt}} = \frac{\exp(\frac{1}{\sigma_\varepsilon}\hat{\mathbb{E}}[\pi_{frt}(\mathcal{J}_{frt})] - \text{Degrees}(\mathcal{J}_{frt}|\mathcal{J}_{frt_0}) - \text{Infrastucture}(\mathcal{J}_{frt}|\mathcal{J}_{frt_0}))}{\sum_{\mathcal{J}\in\mathcal{B}_f}\exp(\frac{1}{\sigma_\varepsilon}\hat{\mathbb{E}}[\pi_{frt}(\mathcal{J})] - \text{Degrees}(\mathcal{J}|\mathcal{J}_{frt_0}) - \text{Infrastucture}(\mathcal{J}|\mathcal{J}_{frt_0}))}. \tag{17}$$

We estimate the model using Maximum Likelihood. To keep the problem tractable, we assume that the set of bundles that a firm can ever offer, $\mathcal{B}_f$, is given by the union of all bundles that we see the firm ever offering in the data.

4.4.2. *Results*: We present the estimated parameters in Online Appendix **??**, Table A.5. We find that adding a new degree to the bundle has larger cost than offering an existing degree. For many areas of study, offering an existing degree does not have an additional fixed cost, which explains why, conditional on keeping their campus or hub open, institutions keep offering in 2019 degrees they were already offering in 2010. We also find important costs of building new infrastructure that are increasing in the distance between regions and firms' headquarters. For in-person campuses, the cost can go up to \$39 per capita, while for online hubs it goes up to \$18 per capita. We show the full distribution of fixed costs associated with opening a new campus or hub in every region as a function of distance in Online Appendix **??**, Figure A.2.

4.5. *Value-added*

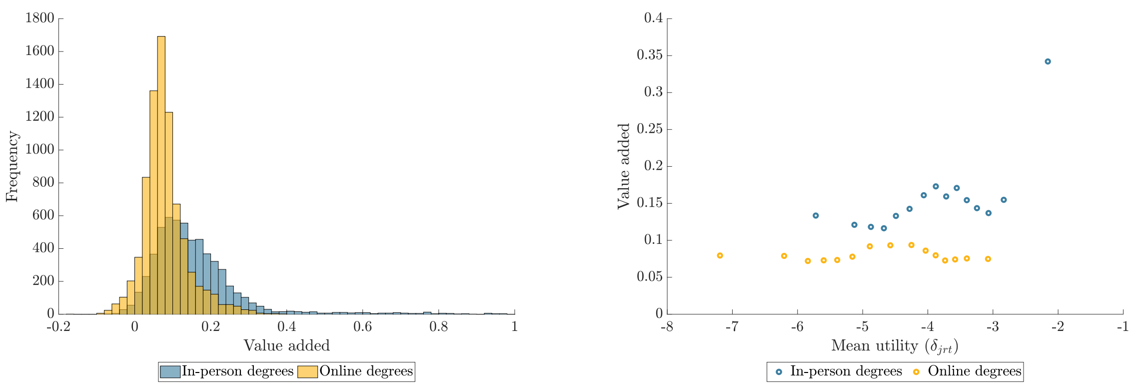
A social planner might be interested in understanding how the expansion of online education changes labor market outcomes of potential students. To do so, we estimate a simple model of value added using the value-added measures presented in Section 2.3.1. We denote it by VA_{jr}. In our model, each degree-region pair is associated with a specific value added given by

$$VA_{jr} = x_j^{(1)}\nu_o + x_j^{(2)}\nu_x + \nu_j + \nu_{ra} + \varsigma_{jr} \tag{18}$$

where $x_j^{(1)} = [\iota_j, o_j]$ contains indicator variables for in-person and online instruction, $x_j^{(2)}$ include a constant, the age of the program, the average score of incoming students, the average wages of graduate students, the length of the program (in number of hours required), and the degree's stem load, ν_j are degree specific components of value added, ν_{ra} are area of study-region specific components, and ς_{jr} is a degree-region specific idiosyncratic shock. The outside option is normalized to be zero in each region. We use this model to impute value added on degree-regions that are and are not offered in the data.

4.5.1. *Estimation*: Following the previous sections, we estimate the value added model using mixed-effects Bayesian hierarchical model where with $\nu_j \sim \mathcal{N}(0, \vartheta_j^2)$, $\nu_{ra} \sim \mathcal{N}(0, \vartheta_{ra}^2)$, and $\varsigma_{jr} \sim \mathcal{N}(0, \vartheta_\varsigma^2)$, and (ν_o, ν_x) non-random parameters. We estimate the model via maximum likelihood to recover unbiased estimates of each component of the value added model that we use to project utility into degree-markets that are offered and not offered in the data.

4.5.2. *Results*: We present the estimated parameters in Online Appendix **??**, Table A.6. In Figure 7(a), we show the distribution of value added for online and in-person degrees. We find that in-person and online degrees have an average value added of 0.157 and 0.08, respectively. Finally, in Figure 7(b), we show that students' preferences over degrees have a positive but weak correlation with value added. The correlation is particularly low for online degrees, which tend to be newer and students might be less informed about their quality.



(a) Distribution of value added

(b) Relationship between preferences and value added

Figure 7: Estimated parameters from fixed costs and value added model

Notes:

5. Counterfactuals

We use the model to assess the impact of introducing online education, isolating the effects of demand, pricing, and program offerings. Additionally, we explore policies that restrict online education for specific groups, examining how these measures could enhance in-person programs while ensuring online access for those who benefit most.

5.1. *The effects of a ban on online education: the role of supply and demand*

Using our model, we compare the 2019 status quo, referred to as *Baseline* (BL), where online education is permitted, with increasingly restrictive counterfactuals that isolate demand and supply effects. In the first scenario, *No supply-side responses* (NS), we remove online degrees while restricting institutions from adjusting tuition or program offerings. Value added under NS depends on two factors: (i) the value added differences between online degrees, in-person degrees, and the outside option, and (ii) the extent to which students switch from online degrees to in-person programs or to the outside option. In the second scenario, *Price responses* (PR), institutions are allowed to set prices optimally, though their program offerings stay constant. This scenario helps measure the effects of reduced competition, which may lead to increased prices and potentially displace students from high-quality in-person degrees. In the final counterfactual *Equilibrium* (EQ), institutions have the flexibility to decide which degrees to offer in each market, corresponding to the equilibrium model outlined in Section 4. Under EQ, the availability of in-person degrees is expected to expand, potentially boosting enrollment in these programs and increasing total value added. The four counterfactuals are summarized in Table 5.

Table 5: Counterfactual exercises

Counterfactual	**Description**
Baseline (BL)	Free entry of online education
No supply-side responses (NS)	No online education & no supply-side responses
Price responses (PR)	(NS) + firms choose prices
Equilibrium (EQ)	(NS) + (PR) + firms choose degree offerings

Notes:

We analyze four outcome categories:

- Degree Offerings: We assess the average number of online and in-person degrees across markets for each counterfactual scenario.

- Total Enrollment: We report national college enrollment, including in-person and online degree enrollments.

- Total Value Added: Calculated as the percentage increase in log-income students gain from attending their preferred degree, normalized by the share of students attending college under the Baseline (BL) scenario:

$$\text{Total value added} = \frac{\sum_r M_r \sum_j s_{jr} VA_{jr}}{\sum_r M_r \sum_j s_{jr}^{BL}},$$

where M_r is the market size for region r and the denominator is constant across counterfactuals. (*Note*: Define s_{jr} for clarity and consider an appendix showing this metric's relevance for social planning.)

- Price and Profit Metrics: We provide (a) the average price of in-person degrees, weighted by in-person enrollment in the Baseline scenario, and (b) total variable profits per capita at Baseline:

$$\text{Avg. price of in-person degrees} = \frac{\sum_r M_r \sum_{j:\iota_j=1} s_{jr}^{BL} p_{jr}}{\sum_r M_r \sum_{j:\iota_j=1} s_{jr}^{BL}}$$

$$\text{Total profits per capita} = \frac{\sum_r M_r \sum_j s_{jr}(p_{jr} - c_{jr})}{\sum_r M_r \sum_j s_{jr}^{BL}}.$$

We focus on 4 sets of outcomes. First, we show what happens to the average number of online and in-person degrees across markets offered under each counterfactual. Second, we show nation-wide college enrollment levels as well as enrollment in in-person and online degrees. Third, we show total value added, calculated as the percentage gain in log-income that students perceive from attending their preferred degree, normalized by the share of students attending college under the Baseline counterfactual (BL), given by

$$\text{Total value-added} = \frac{\sum_r M_r \sum_j s_{jr} VA_{jr}}{\sum_r M_r \sum_j s_{jr}^{BL}}, \tag{19}$$

where M_r is the market size of region r and the denominator is kept fixed across counterfactuals. Finally, we present the average price of in-person degrees weighted by in-person enrollment at Baseline and total variable profits per enrollment at Baseline, given by

$$\text{Av. price of in-person degrees} = \frac{\sum_r M_r \sum_{j:\iota_j=1} s_{jr}^{BL} p_{jr}}{\sum_r M_r \sum_{j:\iota_j=1} s_{jr}^{BL}} \tag{20}$$

$$\text{Total profits per capita} = \frac{\sum_r M_r \sum_j s_{jr}(p_{jr} - c_{jr})}{\sum_r M_r \sum_j s_{jr}^{BL}}. \tag{21}$$

We present results in Table 6. In the presence of online education (BL), there is an average of 44.5 online and 37.7 in-person degrees taught in any given market. Total college enrollment is 5.7, with 3.3% of students attending in person and 2.5% attending online. We find that under BL, total value-added per student enrolled is 0.131 log-points. On the supply side, the average price of in-person degrees is close to \$5900 dollars per year and average industry profits per capita are \$82.68.

When online degrees are not allowed in the market anymore but institutions can not respond by changing prices or their degree offerings (i.e., under NS), total enrollment goes down by 13.6% (from 5.7% to 5%). The decrease of total enrollment highlights the role of online education in expanding the market to students that would not attend college otherwise. Because online education can also divert students from in-person degrees, when restricting online education supply, enrollment in in-person degrees goes up by

Table 6: The effects of banning online education supply

	BL (1)	NS (2)	PR (3)	EQ (4)	$\Delta_{\text{EQ/BL}}$ (5)
Number of online degrees:	44.5	0	0	0	—
Number of in-person degrees:	37.7	37.7	37.7	43.7	15.9%
Total college enrollment:	5.7%	5%	4.9%	5.1%	-11.4%
In-person enrollment:	3.3%	5%	4.9%	5.1%	54.8%
Total value-added:	0.131	0.131	0.129	0.135	2.9%
Av. price of in-person degrees:	5904	5904	6752	6179	4.7%
Total profits per capita:	82.68	73.83	77.55	78.78	-4.7%

Notes: $\Delta_{\text{EQ/BL}} = \frac{EQ-BL}{BL}$

50.9% (from 3.3% to 5%). The decrease in enrollment comes with a subsequent decrease in profits per capita of 10.7%. In terms of value added, restricting the supply of online education has ambiguous effects. On the one hand, because online degrees have larger value added that not attending college, total value added can go down due to the decrease of enrollment in tertiary education. On the other hand, because online degrees have, on average, lower value added than in-person alternatives, total value added can go up due to the increase in in-person enrollment. On average, we find small decreases in value added from 0.131 log-points to 0.131 log-points..

When we allow institutions to respond to the ban of online degrees by adjusting prices of their in-person alternatives (i.e., under PR), we see an overall increase in tuition prices of 14.4%. The increase in prices displaces students from high-quality degrees and total value added goes down to 0.129 log-points.

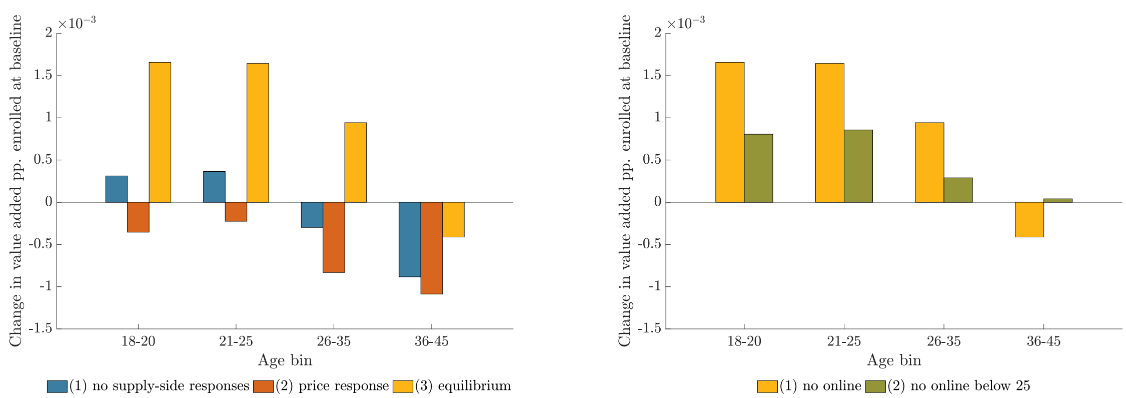
Finally, when institutions are allowed to change their degree offerings (i.e., under EQ), they increase their supply of in-person alternatives by 15.9%, from an average of 37.7 degrees to 43.7 degrees per market. Overall, from the baseline counterfactual to the full equilibrium counterfactual, total enrollment goes down by close to 11.4% (from 5.7% to 5.1%) and in-person enrollment goes up by 54.8% (from 3.3% to 5.1%). Under the equilibrium counterfactual, total value added is similar but higher than at Baseline, with an increase of 2.9% reading 0.135 log-points. Finally, we find that the average tuition price of in-person degrees goes up by 4.7% relative to the Baseline counterfactual. The increase in markups, however, is not large enough to offset the profits losses from not offering online degrees. We find that, on average, institutions decrease their profits by

4.7%.

5.2. *Unpacking value-added changes*

The small changes we see in total value added mask important heterogeneity for winners and losers. Most of the benefits from the expansion of online education tend to be accrued by older individuals who would not attend college if online education would not exist. Younger individuals, on the other hand, are more likely to be diverted from higher-quality in-person alternatives. Diversion from those degrees can be particularly detrimental if online education induces exit or deters entry of in-person degrees that young individuals would prefer even in the presence of online options.

We explore value-added heterogenous effects by cohort in Figure 8(a), where we compare changes in total value added relative to the Baseline counterfactual for each age group. We find that upon forbidding online education and absence any supply side responses, cohorts 18-25 years old see mild increases in value added while cohorts 25-45 see large decreases in value added. Once we allow for institutions to respond by changing prices, every cohorts experiments reductions in value added. It is only once we solve for the equilibrium model and allow for endogenous degrees offerings that cohorts 18-25 years old see substantial gains in total value added.



(a) Value added changes decomposition

(b) Targeted ban to 18-25 years old students

Figure 8: Changes in total value added by age group under each counterfactual

Notes:

5.3. *Targeted policies*

In this subsection, we study hypothetical policies that can take advantage of the benefits that online education can bring to older students without decreasing value added for those younger cohorts. Specifically, we study a policy that limits access to online education to students between 18-25 years old. By forcing young cohorts to attend in-person, the regulator leaves enough profits for in-person degrees on that table, which allows them to co-exist with online alternatives. As a consequence, younger cohorts are able to attend in-person degrees with high value-added, while older cohorts can still benefit from online education's market expansion. We present our main results in Figure 8(b). We find that relative to the Baseline counterfactual (BL), this targeted policy increases value-added for all cohorts.

6. Conclusion

In this paper, we study the equilibrium effects of a rapid expansion of the online for-profit tertiary education sector in Brazil. Two key findings arise from our empirical analysis. First, two opposing forces are driven by the expansion of online degrees. It expands the market by enabling new students to attend college but it diverts students from attending in-person alternatives. Second, online degrees increase competition, putting pressure on prices of in-person degrees and deterring entry.

We develop and estimate an equilibrium model of supply and demand for tertiary education and use it to calculate the effects of the online education expansion on total value added. We find that, in terms of gains of value added, banning online education benefit young cohorts and detriment older ones. Overall, our findings highlight the heterogeneous effects of the online education expansion across the population. While older cohorts benefit from a broader set of available degrees and are able to pursue positive value-added programs, the expansion reduces the availability of in-person options, diverting younger students towards lower-quality alternatives. We find that a policy banning online education for young students would increase value added for all cohorts compared to the baseline scenario where online education is allowed.

Our analysis shows how a theoretical framework combined with data can shed light on how new technologies can reshape competition and market structure. While innovations can offer consumers more choices at lower costs, the benefits are unevenly distributed. In settings with imperfect information, some consumers may unknowingly switch from higher-quality options to the new technology, leaving them worse off. This concern is amplified when traditional alternatives exit the market due to lack of demand. To mitigate these adverse effects, policymakers might consider restricting access to the new technology for negatively impacted groups, ensuring enough demand for incumbent providers to stay in the market.

Online education is a new technology with the potential to reshape education markets. While this paper covers important features of the recent expansion of online education, several unanswered questions remain. First, it focuses on the choices of students and educational institutions and uses value added as an outcome that is invariant to policy. However, the rapid expansion of online degrees can saturate labor markets, changing the returns to college education in unintended ways. Moreover, as online education expands, new degrees might enter the market with an unknown level of quality. Second, future technological developments might improve the effectiveness and lower the cost of online education even further, potentially changing some of the conclusions of this study. Finally, many of the benefits from the expansion of online education accrue to older cohorts that did not have the chance to attend college when younger. As more people attend college right after high school, some of the benefits might vanish.

References

ABED (2018). Censo ead.br: relatrio analitico da aprendizagem a distncia no brasil 2017. *National Center for Education Statistics*.

Allende, C. (2019). Competition under social interactions and the design of education policies. *Job Market Paper*.

Angrist, J., P. Hull, and C. R. Walters (2022). Methods for measuring school effectiveness.

Armona, L. and S. Cao (2022). Redesigning federal student aid in sub-baccalaureate education. *Available at SSRN 4300755*.

Atal, J. P., J. I. Cuesta, and M. Sæthre (2022). Quality regulation and competition: Evidence from pharmaceutical markets. Technical report, National Bureau of Economic Research.

Aucejo, E. M., A. S. Perry, and B. Zafar (2024). Assessing the costs of balancing college and work activities: The gig economy meets online education. Technical report, National Bureau of Economic Research.

Barahona, N., C. Dobbin, H. Ho, S. Otero, and C. Yannelis (2021). Skin in the game: colleges financial incentives and student outcomes. *Occasional Paper. Retrieved from https://Theb. stanford. edu/˜ sotero/papers/skiningame. pdf*.

Barahona, N., C. Dobbin, and S. Otero (2023). Affirmative action in centralized college admissions systems. Technical report, Columbia University.

Barrow, L., W. T. Morris, and L. Sartain (2024). The expanding landscape of online education: Who engages and how they fare. *Journal of Labor Economics 42*(S1), S417–S443.

Bartik, T. J. (1991, September). *Who Benefits from State and Local Economic Development Policies?* W.E. Upjohn Institute.

Bau, N. (2022). Estimating an equilibrium model of horizontal competition in education. *Journal of Political Economy 130*(7), 1717–1764.

Berry, S. and P. Haile (2016). Identification in differentiated products markets. *Annual review of Economics 8*, 27–52.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile Prices in Market Equilibrium. *Econometrica 63*(4), 841–890.

Bertolin, J., T. McCowan, and H. Bittencourt (2023). Expansion of the distance modality in brazilian higher education: Implications for quality and equity. *High Educ Policy 36*, 231–249.

Bettinger, E. P., L. Fox, S. Loeb, and E. S. Taylor (2017, September). Virtual Classrooms: How Online College Courses Affect Student Success. *American Economic Review 107*(9), 2855–2875.

Bodéré, P. (2023). Dynamic spatial competition in early education: An equilibrium analysis of the preschool market in pennsylvania. *Job Market Paper*.

Borusyak, K. and P. Hull (2023). Nonrandom exposure to exogenous shocks. *Econometrica 91*(6), 2155–2185.

Borusyak, K., P. Hull, and X. Jaravel (2021, 06). Quasi-Experimental Shift-Share Research Designs. *The Review of Economic Studies 89*(1), 181–213.

Callaway, B., A. Goodman-Bacon, and P. H. Sant'Anna (2024). Difference-in-differences with a continuous treatment. Technical report, National Bureau of Economic Research.

Conlon, C. and J. Gortmaker (2020). Best practices for differentiated products demand estimation with pyblp. *The RAND Journal of Economics 51*(4), 1108–1161.

Dahlstrand, A. (2024). Defying distance? the provision of services in the digital age. Technical report, University of Zurich.

Dahlstrand, A., N. Le Nestour, and G. Michaels (2024). Online versus in-person services: Effects on patients and providers. Technical report, Institute of Labor Economics (IZA), Bonn.

Deming, D. J., C. Goldin, and L. F. Katz (2012, February). The for-profit postsecondary school sector: Nimble critters or agile predators? *Journal of Economic Perspectives 26*(1), 139–64.

Deming, D. J., C. Goldin, L. F. Katz, and N. Yuchtman (2015, May). Can online learning bend the higher education cost curve? *American Economic Review 105*(5), 496–501.

Deming, D. J., N. Yuchtman, A. Abulafi, C. Goldin, and L. F. Katz (2016, March). The value of postsecondary credentials in the labor market: An experimental study. *American Economic Review 106*(3), 778806.

Dinerstein, M., D. Morales, C. Neilson, and S. Otero (2023). The equilibrium effects of public provision in education markets..

Dinerstein, M. and T. D. Smith (2021). Quantifying the supply response of private schools to public policies. *American Economic Review 111*(10), 3376–3417.

Dobbin, C., N. Barahona, and S. Otero (2021). The equilibrium effects of subsidized student loans. Technical report, Working paper, Stanford University.

Fabregas, R. and L. Navarro-Sola (2024, July 25). Broadcasting education at scale: Long-term labor market impacts of television-based schools. Working paper.

Figlio, D., M. Rush, and L. Yin (2013). Is it live or is it internet? experimental estimates of the effects of online instruction on student learning. *Journal of Labor Economics 31*(4), 763–784.

Gandhi, A. and J.-F. Houde (2019). Measuring substitution patterns in differentiated-products industries. *NBER working paper* (w26375).

Garrett, R., B. Simunich, R. Legon, and E. E. Fredericksen (2022). Chloe 7: Tracking online learning from mainstream acceptance to universal adoption. Technical report, Retrieved from the Quality Matters website: https://www.qualitymatters.org/qa-resources/resource- center/articles-resources/CHLOE-project.

Goetz, D. (2023). Telemedicine competition, pricing, and technology adoption: Evidence from talk therapists. *International Journal of Industrial Organization 89*, 102956.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020). Bartik instruments: What, when, why, and how. *American Economic Review 110*(8), 2586–2624.

Goodman, J., J. Melkers, and A. Pallais (2019, January). Can Online Delivery Increase Access to Education? *Journal of Labor Economics 37*(1), 1–34. Publisher: The University of Chicago Press.

Hausman, J., G. Leonard, and J. D. Zona (1994). Competitive analysis with differenciated products. *Annales d'Economie et de Statistique*, 159–180.

Hoxby, C. M. (2018). Online postsecondary education and labor productivity. In C. R. Hulten and V. A. Ramey (Eds.), *Education, Skills, and Technical Change: Implications for Future US GDP Growth*, Volume 2, pp. 401–460. Chicago: University of Chicago Press. Conference held in October 2015, published in December 2018.

Kofoed, M. S., L. Gebhart, D. Gilmore, and R. Moschitto (2024, September). Zooming to class? experimental evidence on college students' online learning during covid-19. *American Economic Review: Insights 6*(3), 32440.

Mountjoy, J. (2022, August). Community colleges and upward mobility. *American Economic Review 112*(8), 25802630.

NCES (2022). Digest of education statistics 2022, table 311.15. Technical report, U.S. Department of Education.

Neilson, C. et al. (2013). Targeted vouchers, competition among schools, and the academic achievement of poor students. *Job Market Paper 48*.

Rouse, C. E. (1995). Democratization or diversion? the effect of community colleges on educational attainment. *Journal of Business & Economic Statistics 13*(2), 217–224.

Rouse, C. E. (1998). Do two-year colleges increase overall educational attainment? evidence from the states. *Journal of Policy Analysis and Management 17*(4), 595–620.

Sanchez, C. (2023). Equilibrium consequences of vouchers under simultaneous extensive and intensive margins competition. Technical report, Working Papers.

Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. *The RAND Journal of Economics 37*(3), 619–640.

Sweeting, A. (2009). The strategic timing incentives of commercial radio stations: An empirical analysis using multiple equilibria. *The RAND Journal of Economics 40*(4), 710–742.

Zeltzer, D., L. Einav, J. Rashba, and R. D. Balicer (2023, 05). The Impact of Increased Access to Telemedicine. *Journal of the European Economic Association 22*(2), 712–750.

Online Appendix for:

The Effects of Online Education on Market Structure and Students' Enrollment

(Not for publication)

Nano Barahona Cauê Dobbin Joaquin Fuenzalida Sebastián Otero

October 28, 2024

Appendix A: Additional Figures and Tables

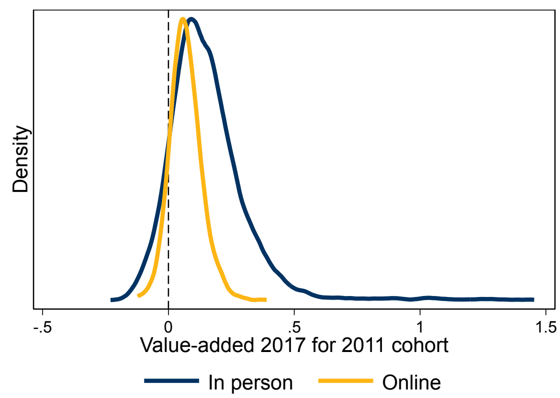


Figure A.1: Value-added distribution for in-person and online degrees

Notes:

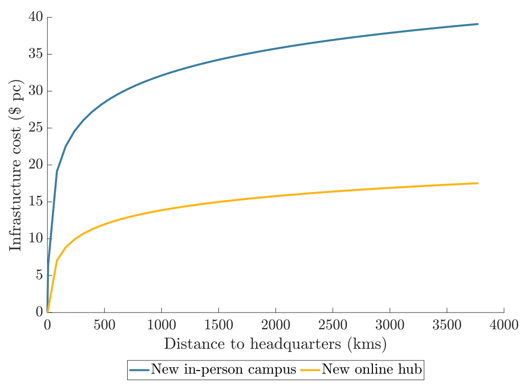


Figure A.2: Cost of building infrastructure as a function of distance

Notes:

Table A.1: Areas of Study and Degrees

Area of Study	**Type**	**Degrees**
Arts, Humanities, and Social Sciences	Bachelor	Library, Political Science, Social Communication, Design, Journalism, International Relations
	Technical	Interior Design, Fashion Design, Product Design, Graphic Design, Photography, Digital Marketing, Audiovisual Production, Multimedia Production
Business	Bachelor	Administration, Accounting, Economics, Public Administration, Social Communication, Advertisement and Marketing, Public Relations, Executive Secretariat (Secretariado Executivo)
	Technical	Foreign Trade, Entrepreneurship, Commercial Management, Quality Management, Information Techonology Management, Human Resources Management, Security Management, Public Health Management, Legal Management, Information System Management, Financial Management, Hospital Management, Public Management, Logistics, Marketing, Real Estate, Management Processes, Secretariat (Secretariado)
Education	Bachelor	Visual Arts, Biological Sciences, Physical Education, Philosophy, Geography, History, Mathematics, Psychopedagogy, Chemistry
	?	Visual Arts, Biological Sciences, Social Sciences, Physical Education, Geography, History, (Letras), Mathematics, Pedagogy, Chemistry, Sociology
Engineering	Bachelor	Agronomy, Architecture and Urbanism, Civil Engineering, Computer Engineering, Software Engineering, Electric Engineering, Mechanical Engineering, Chemical Engineering
	Technical	Agribusiness, Agribusiness Management, Industrial Mechatronics (Mecatronica Industrial), Mechanical Manufacturing
Health Sciences	Bachelor	Nursing, Pharmacy, Physiotherapy, Veterinary Medicine, Speech Therapy, Nutrition, Occupational Therapy
	Technical	Radiology
Services	Bachelor	Aeronautical Sciences, Social Service, Tourism
	Technical	Beauty and Personal Image, Aesthetics and Cosmetics, Gastronomy, Environmental Management, Tourism Management, Public Security, Juridic Services
Others	Bachelor	Environmental Engineering, Production Engineering, Theology
	Technical	Industrial Automation, Social Educator, Management of Industrial Production
Law	Bachelor	Law
Medicine	Bachelor	Medicine, Odontology
Psychology	Bachelor	Psychology
Math, Computer and Natural Sciences	Bachelor	Biomedicine, Computer Science, Information Systems
	Technical	Analysis and System Development, *Banco de Dados*, *Jogos Digitais*, *Redes de Computadores*, Information Security, *Sistemas para Internet*

Table A.2: Relationship between distance and probability of opening an online degree for selected years

	entered_{fr} (1)
$\log(1+d_{fr})$	-0.088
	(0.008)
H_{fr}	-0.146
	(0.076)
Obs.	4070
Regions	110
Firms	37

Notes:

Table A.3: Demand parameters

Panel A: BLP parameters									
Price coef.		Price RC		Online RC		In-person RC		Nest param.	
$\bar{\alpha}$	-1.976	σ_α	0.295	σ^o	0.017	σ^p	0.016	ρ	0.85
	(0.494)		(2.05)		(21.902)		(27.876)		(0.125)
Dem. shifter		Dem. shifter		Age 21-25		Age 26-35		Age 36-45	
ψ^0	0.178	ψ^o	-0.032	μ_2^p	-0.953	μ_3^p	-1.48	μ_4^p	-1.987
	(0.024)		(0.026)		(3.571)		(3.361)		(12.549)
Age 21-25		Age 26-35		Age 36-45					
μ_2^o	-0.659	μ_3^o	-0.935	μ_4^o	-1.285				
	(2.294)		(7.092)		(3.335)				
Panel B: Mixed model parameters									
Degree r.e.		Region-area r.e.		Year-area r.e.		Year-online r.e.		Demand sock	
σ_j	0.255	σ_{ra}	0.619	σ_{ta}	0.854	σ_{to}	0.276	σ_ξ	0.321
	(0.004)		(0.013)		(0.061)		(0.052)		(0.001)
Hours f.e.		Stem f.e.		Degree age f.e.		Av. score f.e.		Av. wages f.e.	
δ_1	0.132	δ_2	0.033	δ_3	0.049	δ_4	0.007	δ_5	0.024
	(0.007)		(0.024)		(0.007)		(0.0002)		(0.023)
Const. f.e.									
δ_0	-10.949								
	(0.199)								

Notes:

Table A.4: Marginal cost parameters

Panel A: BLP parameters									
Demand instrument		log-distance		same-region					
γ_{z_1}	0.101	γ_{z_2}	0.133	γ_{z_3}	0.545				
	(0.101)		(0.133)		(0.545)				
Panel B: Mixed model parameters									
Degree r.e.		Region-area r.e.		Year-area r.e.		Year-online r.e.		Cost sock	
σ_j	0.414	σ_{ra}	0.267	σ_{ta}	0.177	σ_{to}	0.782	σ_ξ	0.441
	(0.006)		(0.007)		(0.028)		(0.127)		(0.001)
Hours f.e.		Stem f.e.		Degree age f.e.		Av. score f.e.		Av. wages f.e.	
δ_1	0.13	δ_2	-0.119	δ_3	0.026	δ_4	0.006	δ_5	0.186
	(0.01)		(0.034)		(0.011)		(0.0003)		(0.036)
Const. f.e.									
δ_0	-5.96								
	(0.311)								

Notes:

Table A.5: Fixed cost parameters

Panel A: Infrastructure costs											
		New campus						New hub			
χ_c^0	-0.763	χ_c^1	0.763	χ_c^2	0.959	χ_c^0	-0.945	χ_c^1	0.945	χ_c^2	0.503
	(NaN)		(NaN)		(NaN)		(NaN)		(NaN)		(NaN)
Panel B: Degree costs											
κ_0	0.924	$\kappa_{\iota 1}^{old}$	0.763	κ_{o1}^{old}	0	$\kappa_{\iota 2}^{old}$	0.171	κ_{o2}^{old}	0	$\kappa_{\iota 3}^{old}$	0
	(NaN)		(NaN)		(NaN)		(NaN)		(NaN)		(NaN)
κ_{o3}^{old}	0	$\kappa_{\iota 4}^{old}$	0	κ_{o4}^{old}	0	$\kappa_{\iota 5}^{old}$	0	κ_{o5}^{old}	0	$\kappa_{\iota 6}^{old}$	0
	(NaN)		(NaN)		(NaN)		(NaN)		(NaN)		(NaN)
$\kappa_{\iota 7}^{old}$	0	κ_{o7}^{old}	0	$\kappa_{\iota 8}^{old}$	0	$\kappa_{\iota 9}^{old}$	0	κ_{o9}^{old}	0	$\kappa_{\iota 10}^{old}$	0
	(NaN)		(NaN)		(NaN)		(NaN)		(NaN)		(NaN)
$\kappa_{\iota 11}^{old}$	0	κ_{o11}^{old}	0	$\kappa_{\iota 1}^{new}$	0	κ_{o1}^{new}	0.743	$\kappa_{\iota 2}^{new}$	0	κ_{o2}^{new}	0.975
	(NaN)		(NaN)		(NaN)		(NaN)		(NaN)		(NaN)
$\kappa_{\iota 3}^{new}$	0	κ_{o3}^{new}	0.71	$\kappa_{\iota 4}^{new}$	0	κ_{o4}^{new}	0	$\kappa_{\iota 5}^{new}$	0	κ_{o5}^{new}	0
	(NaN)		(NaN)		(NaN)		(NaN)		(NaN)		(NaN)
$\kappa_{\iota 6}^{new}$	0	$\kappa_{\iota 7}^{new}$	0.207	κ_{o7}^{new}	0.449	$\kappa_{\iota 8}^{new}$	0	$\kappa_{\iota 9}^{new}$	0	κ_{o9}^{new}	0
	(NaN)		(NaN)		(NaN)		(NaN)		(NaN)		(NaN)
$\kappa_{\iota 10}^{new}$	0	$\kappa_{\iota 11}^{new}$	0	κ_{o11}^{new}	0	σ_ε	0.005				
	(NaN)		(NaN)		(NaN)		(NaN)				

Notes:

Table A.6: Value added parameters

Panel B: Mixed model parameters									
Degree r.e.		Region-area r.e.		VA shock		Const. f.e.		Online f.e.	
σ_j	0.057	σ_{ra}	0.068	σ_{ta}	0.068	δ_0	-1.236	δ_o	-0.041
	(0.002)		(0.003)		(0.001)		(0.056)		(0.007)
Hours f.e.		Stem f.e.		Degree age f.e.		Av. score f.e.		Av. wages f.e.	
δ_1	0.066	δ_2	-0.026	δ_3	-0.016	δ_4	0.001	δ_5	0.102
	(0.003)		(0.007)		(0.004)		(0.0001)		(0.008)

Notes:

Table A.7: Estimation summary by area of study

	In person			Online		
	c_{jrt}	$\hat{c}_{jrt}$	p_{jrt}	c_{jrt}	$\hat{c}_{jrt}$	p_{jrt}
	δ_{jrt}	$\hat{\delta}_{jrt}$	va_{jr}	δ_{jrt}	$\hat{\delta}_{jrt}$	va_{jr}
	$\frac{p_{jrt}-c_{jrt}}{p_{jrt}}$	κ_a^{old}	κ_a^{new}	$\frac{p_{jrt}-c_{jrt}}{p_{jrt}}$	κ_a^{old}	κ_a^{new}
Arts,	3.228	3.111	4.67	0.672	0.725	1.951
Humanities and	-5.037	-4.94	0.069	-6.397	-6.447	0.03
Social Sciences	0.373	0	0.004	0.71	0.001	0
	2.471	2.3	3.714	0.487	0.48	1.589
Business	-2.997	-3.013	0.116	-3.733	-3.727	0.076
	0.386	0	0.005	0.729	0	0
	2.269	2.197	3.506	0.456	0.461	1.562
Education	-3.315	-3.307	0.077	-3.836	-3.839	0.071
	0.399	0	0.004	0.754	0	0
	4.422	4.54	5.81	1.663	1.613	2.823
Engineering	-3.993	-3.96	0.183	-5.105	-5.138	0.144
	0.259	0	0	0.449	0	0
	4.149	4.11	5.549	1.82	1.75	3.247
Health Sciences	-3.476	-3.474	0.216	-4.31	-4.317	0.181
	0.274	0	0	0.501	0	0
	4.057	4.181	5.501	—	—	—
Law	-3.775	-3.775	0.168	—	—	—
	0.283	0	0.001	—	—	—
Math,	3.05	2.99	4.436	0.583	0.544	1.726
Computer and	-4.54	-4.59	0.129	-5.659	-5.628	0.068
Natural Sciences	0.354	0	0.002	0.699	0	0
	17.373	18.49	19.986	—	—	—
Medicine	-2.909	-2.909	0.486	—	—	—
	0.158	0	0	—	—	—
	3.942	3.307	5.297	0.662	0.758	1.918
Other	-5.222	-5.174	0.193	-6.174	-6.204	0.127
	0.286	0	0	0.723	0	0
	3.904	4.005	5.562	—	—	—
Psychology	-4.665	-4.665	0.106	—	—	—
	0.328	0	0	—	—	—
	2.546	2.461	3.874	0.601	0.575	1.758
Services	-4.73	-4.823	0.058	-5.374	-5.342	0.05
	0.399	0	0	0.717	0	0

Notes:

Appendix B: Data build

In this Appendix, we describe how we construct each of the variables we use in our analysis:

B.1. *Tuition Fees*

To construct program-level prices, we combine these data sources as follows: First, we use the information from all sources to recover an average posted price by program, hub, and year. This is done by regressing log-prices on program-hub-year and source fixed effects. Controlling for source fixed effects helps account for persistent differences across data sources and allows us to recover a program-hub-year price. In cases where information for a certain year is missing, we run a regression for the predicted price on program-hub and year fixed effects, imputing the missing values based on the sum of the coefficients. Finally, in situations where information on both year and hub is unavailable, we regress the predicted price on program and year fixed effects, filling the missing values using the sum of these coefficients.

Appendix C: Tests for identifying assumptions

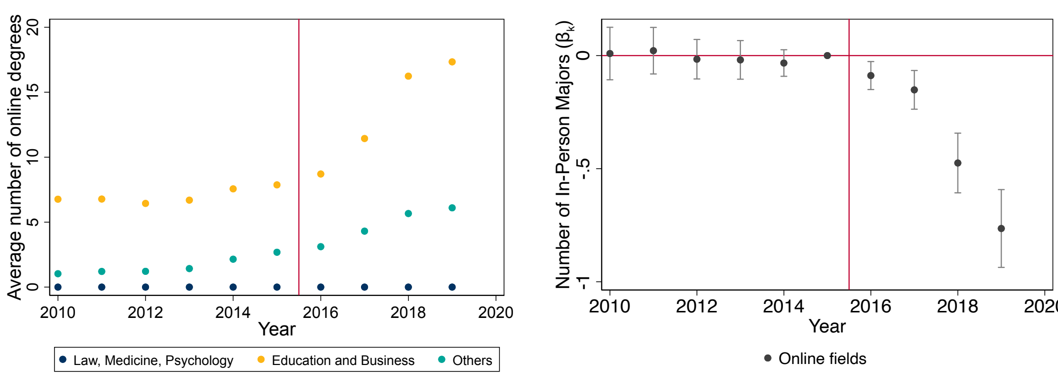
C.1. *Parallel trends across areas of study*

In Figure 1(a), we show that starting in 2016, there was a change in the trend of the number of online majors that were offered in each region belonging to education and business. Other majors, such as Arts, Humanities and Social Sciences, or Math, Computer, and Natural Sciences, also saw a large increase. Regulated majors like Law, Medicine, or Psychology, on the other hand, didn't see any expansion.

To assess how the expansion of online degrees affected the supply of in-person degrees, we run the following regression:

$$\text{in-person}_{rta} = \sum_k \beta_k T_a \mathbb{1}\{t=k\} + \delta_r + \delta_t + \delta_a + \varepsilon_{rta}$$

where in-person$_{rtf}$ is the number of in-person majors offered in region r in year t that belong to area of study a, T_f gets the value of 1 if the field of study a is unregulated, and δ_r, δ_t, and δ_a are region, year, and field of study fixed effects, respectively. We are interested in the coefficient β_k, which shows the relative difference in the number of in-person majors between unregulated and regulated fields of study. We present results in Figure 2(d). Standard errors are clustered at the region level.



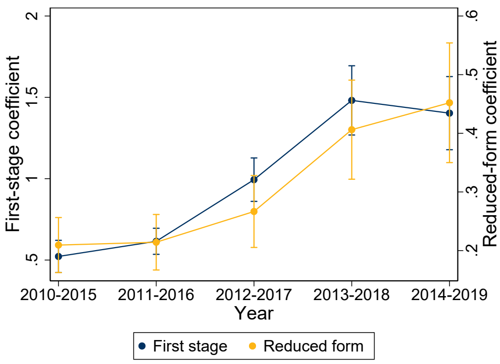
(a) Online Majors

(b) In-person Majors

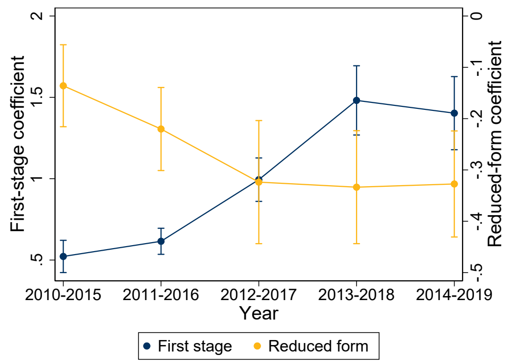
Figure C.1: Average number of online and in-person majors offered in each region

Notes:

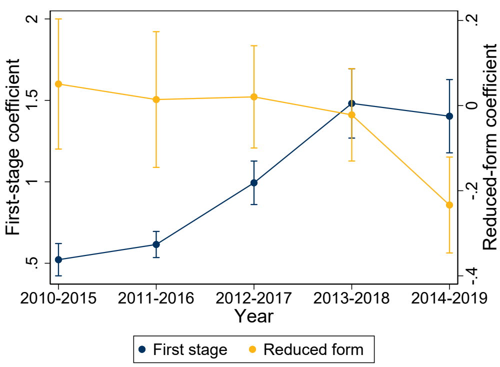
C.2. *Assessing the validity of the instrument*



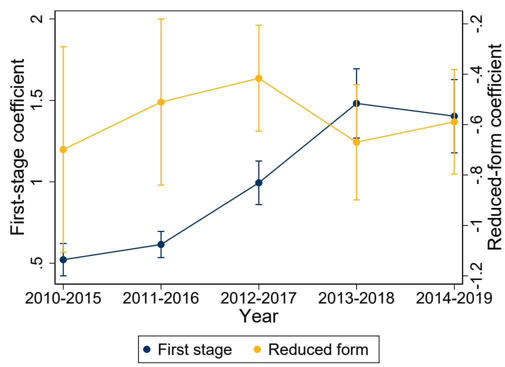
(a) Number of online students



(b) Number of in-person students



(c) Number of in-person degrees



(d) Average log-price of in-person degrees

Figure C.2: IV regression on different subsamples

Notes: