Looking Ahead? Future Educational Opportunities and Pre-College Human Capital

Virna Vidal-Menezes*

November 4, 2026

Job Market Paper

Click here for the latest version

Abstract

How do future schooling opportunities shape early educational decisions? In this paper, I exploit a large expansion of free higher education institutions (HEIs) in Brazil to study whether an increased supply of tertiary education can shift demand for secondary schooling. Combining a novel dataset on HEI rollout with administrative data and a staggered difference-in-differences design, I find that new campuses reduce dropout in every high school grade and increase completion by nearly 1 percentage point, adding over 20,000 high school graduates in the first three exposed cohorts. A simple framework shows that new HEIs can raise the value of high school both as an end, by changing the local economy, and as a gateway to college. Empirically, I find that these short-run effects likely stem from increased college opportunities rather than labor market changes: wages of high school graduates do not change, while local students' enrollment in federal HEIs increases by 140%. Importantly, high school graduation increases only in municipalities with a high pre-policy college premium, highlighting the role of tertiary returns in raising the value of staying in school. My paper shows that expanding access to higher education can improve early educational attainment, especially in low- and middle-income countries where returns to high school alone are not enough to keep students in school.

Keywords: Returns to education, higher education, college premium, educational attainment

JEL Classification: O12, O15, O54, I25, I26

*Department of Economics, University of Notre Dame (vvidalme@nd.edu). I thank the National Academy of Education, the Spencer Foundation, the Kellogg Institute for International Studies, the Program for Interdisciplinary Educational Research (PIER), and the Building Inclusive Growth Lab at the University of Notre Dame for financial support. For comments and suggestions, I thank participants of the Development Economics Workshop at Notre Dame. I extend a special thanks to Taryn Dinkelman, Jeff Denning, Joe Kaboski, and Ethan Lieber for their feedback and advising on this project.

1 Introduction

Barriers to education in low- and middle-income countries (LMICs) have declined substantially in recent decades, yet over 270 million children and youth worldwide remain out of school, the vast majority in LMICs (UNESCO, 2026).[1] A large body of literature has studied both the supply and demand factors that shape schooling attainment. On the supply side, expanding K-12 school construction (e.g., Duflo, 2001 and Burde and Linden, 2013) has substantially increased enrollment. On the demand side, where students decide whether to stay in school by weighing the marginal costs of additional schooling against its (perceived) marginal benefits, interventions have either reduced costs, through cash transfers, scholarships, and tuition-free programs (Schultz, 2004; Barrera-Osorio et al., 2011; Lucas and Mbiti, 2012; Duflo et al., 2021), or increased perceived and immediate returns, by providing information about the returns to schooling (Jensen, 2010; Dinkelman and Martínez, 2014) and through conditional cash transfers.[2] Despite the success of many of these interventions, many countries still face the challenge of how to tackle dropout from high school.

In this paper, I investigate the role of expanding the supply of higher levels of education in raising demand for schooling at lower levels. I ask: What are the effects of increasing the local supply of higher education institutions (HEIs) on persistence through and completion of high school? A new HEI might change the perceived benefits of secondary schooling in a few ways. First, by altering the local economy and labor market, these new institutions might directly change the actual returns to completing high school; a high school degree may become more or less valuable as a result. Second, they might increase the expected returns to staying in school by lowering the costs of accessing higher education and increasing the likelihood that local students enroll in an HEI at all; a high school degree would then become more valuable as a gateway into tertiary education.

To measure the effects of new HEIs on high school outcomes and provide evidence on the mechanisms at play, I study an ambitious federal program in Brazil that more than doubled the supply of free public colleges and trade schools between 2003 and 2014.[3] Brazil is a particularly well-suited setting for this question not only because of the scale of the expansion policy, but also because

[1] By 2023, over half of all countries included free primary and secondary education in their legal frameworks (UNESCO, 2025). In Brazil, the context of this paper, primary and secondary education has been free and a constitutional right for nearly 40 years and became compulsory in 2009. Yet in 2024, over 25% of 25-to-34-year-olds in Brazil did not complete high school (OECD, 2026) and over 1 in 10 high school students dropped out before graduating.

[2] Conditional cash transfers work both by alleviating costs, since they provide cash to families, and by increasing immediate returns to staying in school, since they condition these transfers on school participation.

[3] The policy includes expanding both universities and trade schools. Trade schools traditionally offer technical and professional courses, and majors focused on technical skills, although some do offer traditional majors.

of the country's educational landscape: secondary schooling has been free and a constitutional right for nearly 40 years, yet over 1 in 10 students who start high school drop out. By studying the short-run responses of high schoolers exposed to new HEIs, I assess whether investments at the top of the schooling distribution can raise overall demand for education, even at lower levels.

Empirically, I estimate the impact of a new campus on local public school students' outcomes (likelihood of starting and graduating high school, and dropout rates from 9th to 12th grade) using a staggered difference-in-differences approach proposed by Callaway and Sant'Anna (2021), comparing public school students in municipalities that received a new HEI with those in comparable municipalities that never received one during this period. I build a novel dataset of over 400 HEI openings using information from institutional documents, HEIs' websites, and legal ordinances, combined with rich administrative data from the Ministry of Education covering the universe of K-12 enrollment and college-entrance exam takers, and from the Ministry of Labor covering the universe of formal-sector employees.

My identification strategy faces two key challenges. First, the federal government did not randomly select treated municipalities, and they may differ substantially from control municipalities. To address this selection problem, I control for several municipality-level characteristics observed by the government when selecting receiving municipalities, as stated in institutional documents, and for relevant characteristics documented in the literature, such as political alignment (Arroyo and Petterini, 2020). I also use a difference-in-differences design that requires comparability in trends, not in levels. Event-study estimates show limited evidence of differential pre-trends, even in the absence of controls. Second, the large influx of college students and staff resulting from this policy can alter the composition of primary and secondary schools, potentially driving my results. I investigate whether the shares of public school students who are white, who attended only public schools, and whose mothers completed at least high school are changing as a result of this policy. I find no evidence that changes in composition could explain the observed results. Finally, my results are also robust to alternative definitions of the control group and main outcomes, to different sets of control variables (including none), and to a permutation test, which shows that the estimated effects are unlikely to arise by chance.

I find that public school students in municipalities that received an HEI as part of this program decreased dropout at all three grades of high school by 0.6 to 1.2 percentage points (a 5–6.5% decrease relative to the baseline mean), with no effect on dropout rates at 9th grade, the last grade of primary education. This increase in persistence translates into increased high school completion:

students were nearly 1 percentage point more likely to eventually graduate from high school (a 1.5% increase relative to the pre-policy mean), but no more likely to graduate on time. This amounts to over 20,000 additional high school graduates in the first three cohorts exposed to a new HEI. These gains are concentrated among non-white students, who were the most disadvantaged pre-policy. Students were also 20% more likely to take a national college-entry exam, with no decrease in scores. I find no significant impacts on in-school performance, measured either at 9th grade or at the end of high school.

The effect of new HEIs on high school graduation represents one-fifth of the impact of a conditional cash transfer program in Colombia (Baez and Camacho, 2011), an intervention that directly rewards K-12 school attendance, and one-tenth to one-sixth of the effects found in quasi-experimental studies assessing the impact of being admitted to secondary school or receiving a secondary school voucher on graduation (Ozier, 2018; Angrist et al., 2006). Unlike those interventions, however, this policy did not directly target pre-tertiary students or change the supply of high school opportunities.

This policy provides a unique setting where different municipalities received different types of HEIs. Comparing policy effects by type of HEI received allows me to test whether different types of post-secondary education affect pre-tertiary students' decisions in different ways. I find suggestive evidence that, while both universities and trade schools increase high school completion, only universities appear to change the decision to enroll in high school in the first place. This suggests that the two types of institutions move different marginal students. Trade school opportunities seem more salient to students who would have transitioned to high school but dropped out in the absence of the policy, but now choose to stay the course and graduate.

To understand why higher education expansion might increase persistence at lower levels, I develop a simple framework illustrating how students decide whether to complete high school and how changes in the availability of higher education can alter this decision. In the framework, a new HEI can raise high school completion in two ways. First, it can make high school more valuable on its own, as an *end*: if the local labor market changes in ways that raise the standalone returns to a high school degree, some students who would otherwise drop out can now opt to graduate high school. Second, it can make high school more valuable as a *gateway*. By increasing the likelihood of college attendance and decreasing attendance costs, a new HEI can give students for whom high school alone is not worth completing a reason to stay: the opportunity to attend college. This second path, however, requires a high enough college premium: college opportunities only change

high school decisions if college offers a meaningful increase in expected returns to schooling.

Empirically, I test the framework's predictions to assess which mechanisms and transitions are most likely at play. I find suggestive evidence that the observed high school graduation effects stem from changes in local students' college access rather than in the local labor market. I find no changes in high school graduate wages that could account for the observed effects on school completion. By contrast, local public school students are 0.7 percentage points more likely to enroll in a federal HEI following the policy, a 140% increase relative to the pre-policy mean. Although effects of new HEIs on college enrollment and local wages are similar in high- and low-premium municipalities, only in places where the pre-policy college premium was high do students increase high school completion. This is consistent with students treating high school as a gateway to college rather than as an end.

Given that the college premium seems to play an important role in mediating the effects, I examine how college and high school premiums in the formal labor market change over the six years following the opening of an HEI, to assess whether higher education opportunities might be a sustainable way to increase demand for schooling at lower levels. I find that the college premium remains stable over time, but the high school premium decreases sharply following the opening of an HEI. This drop is driven by an increase in primary-educated wages rather than a decrease in high school graduate wages. This increase may reflect a decrease in the supply of unskilled labor as educational attainment rises, similar to the general equilibrium effects found by Khanna (2023) for primary education expansion in India, or an increase in the local demand for unskilled labor generated by the new campuses. In the framework, a lower high school premium weakens high school's value as an end but expands the pool of students for whom college is the main reason to stay in school. Provided that the college premium remains sizable, which is plausible given the limited supply of college graduates in Brazil (only 7% of adults had completed college in 2000), higher education opportunities should therefore continue to shape pre-tertiary schooling decisions in the long run.

This paper contributes to the broader literature on the effects of college expansion or proximity on educational attainment, which primarily focuses on the direct impact through college enrollment and completion, and local economic impacts (Card, 1993; Currie and Moretti, 2003; Oreopoulos and Petronijevic, 2013; Vu, 2024; Feinmann and Hsu Rocha, 2026; Cox, 2024). I add to this literature by providing evidence on the effects of *expected* higher education attendance on local teenagers' human capital investments, a result especially relevant for low- and middle-income countries where average attainment falls well below high school completion.

I directly build on and contribute to the literature examining the effects of subsequent educational opportunities on earlier educational attainment. While a strand of this literature examines the effects of affirmative action on secondary students' outcomes (Akhtari et al., 2023; Khanna, 2020), my paper more closely aligns with a growing literature on the impact of expanding the physical supply of higher levels of education on students' attainment at lower levels (Bedard, 2001; Jagnani and Khanna, 2020; Katzkowicz et al., 2023; Sandholtz, 2024; Rands and Barsanetti, 2025; Zhang, 2026; Erbabian, 2026). The evidence in this literature is mixed: positive effects of higher education expansion on pre-tertiary schooling outcomes are observed in India (Jagnani and Khanna, 2020) and Kenya (Erbabian, 2026), while negative impacts are documented in the United States (Bedard, 2001) and Chile (Erbabian, 2026).

I build on this literature in a few ways. First, by combining a theoretical framework with empirical evidence, I identify the two possible ways new colleges can increase high school attainment (as an end or as a gateway) and highlight the role of the distribution of returns to schooling in shaping these effects. In doing so, I help reconcile mixed evidence stemming from high- and low- and middle-income countries, contexts that differ substantially in terms of returns to schooling and baseline high school attainment. Second, I provide novel evidence on the types of higher education institutions that matter for teenagers' schooling outcomes. In particular, I show that not only traditional universities but also trade schools can shift local students' high school decisions.

The remainder of this paper is organized as follows. Section 2 details the HEI expansion policy. Section 3 describes the data and provides more detail on the placement of HEIs. Section 4 describes the method used in this paper. Section 5 reports the results for all pre-tertiary outcomes. Section 6 presents a theoretical framework to explain how the expansion of tertiary education can affect students' educational decisions at lower levels. Section 7 provides empirical evidence on the mechanisms through which the policy operates, while Section 8 examines the policy's long-run impact on local wages. Section 9 presents robustness analyses, and Section 10 concludes with a discussion of the results.

2 Background

2.1 Higher education expansion program in Brazil

Tertiary education in Brazil comprises public and private institutions. The former is primarily provided by the federal government through Federal Universities and Trade Schools.[4] These institutions are free to students, and admissions are primarily based on a national exam (ENEM) test score that students completing high school can take once a year.[5] After obtaining their college-entry exam test scores, students apply to federal universities through the Unified Selection System, SISU, a centralized platform that allows students to choose the major and HEI combination they wish to pursue. Private and state universities have their own admission process, although some also use ENEM test scores.

In 2003, the federal government of Brazil launched a large higher education expansion program, with the main goal of reducing the concentration of higher education institutions in higher-income regions and capitals, and using higher education to promote regional economic development, aiming to increase municipalities' income and reduce regional inequalities. The government selected municipalities based on geographical and economic criteria, such as the municipality's Human Development Index (HDI), population size, local economic activities, and pre-existing supply of higher education institutions (Barbosa et al., 2020; Ministério da Educação [MEC], 2011). The policy included the expansion of new public and free universities and trade schools, as well as new campuses of already existing HEIs, and was divided into three phases.

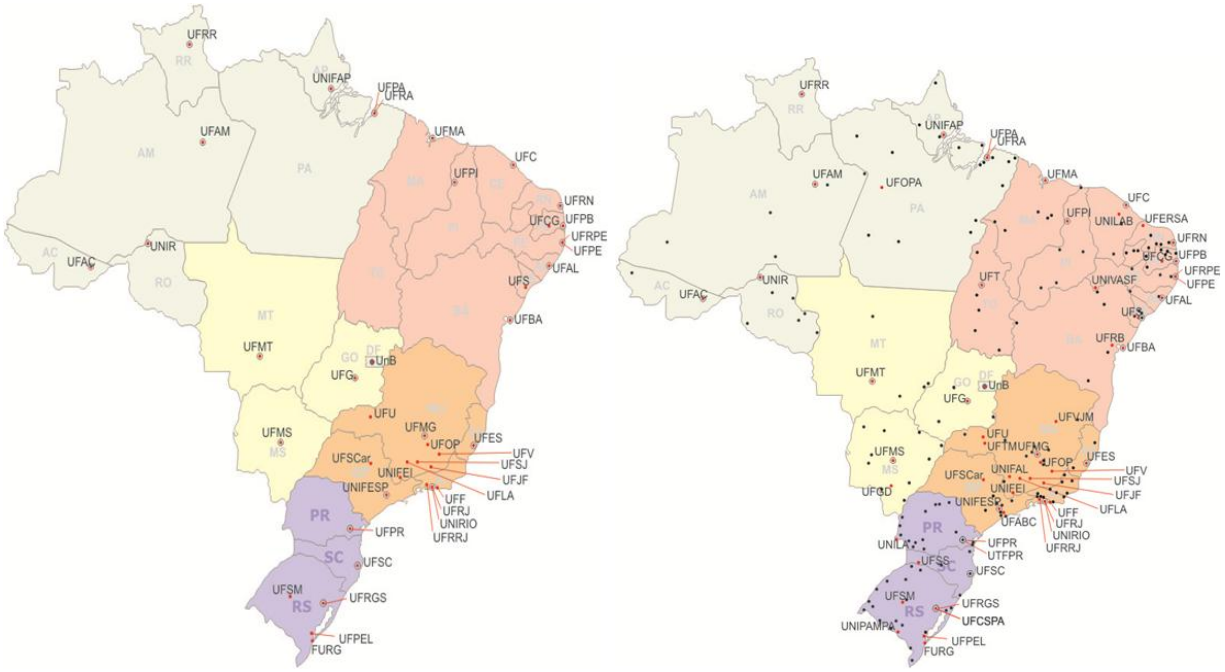
Figure 1 presents the distribution of federal university campuses before and after the expansion program and illustrates the policy's goal of decreasing the concentration of higher education institutions in capitals, metropolitan areas, and more developed regions of the country. In total, 554 municipalities (nearly 10% of all Brazilian municipalities) received at least one new federal HEI between 2003 and 2017.[6] Of those, 252 municipalities received their first higher education institution. Detailed description of the different expansion phases is in the Appendix.

[4] State governments can also create and fund state universities, which are also free to students.

[5] The national unified exam, ENEM, started being used in college admissions in 2010. Beforehand, students took individual exams for each university they would apply to. Trade schools have alternative admission routes, but also use ENEM as part of their admissions process.

[6] Although institutions continued to open until 2017, the policy formally ended in 2014. I therefore restrict my analysis to HEIs that opened through 2014.

Figure 1: Expansion of Federal Universities from 2002 (left) to 2014 (right)



Note: This figure comes from the Brazilian Ministry of Education website. Distribution of federal universities in Brazil in 2002 (left), with 45 institutions in 141 municipalities, with a total of 151 campuses, and in 2014 (right), with 63 institutions in 275 municipalities, with a total of 321 campuses. Colors represent the different regions of Brazil. This map does not include Trade Schools. Before the policy, there were 140 federal trade school campuses in Brazil.

2.2 Federal trade schools

Trade schools account for approximately two-thirds of the HEIs opened under the federal expansion program. These institutions differ from universities in relevant ways that could generate differential effects, depending on the type of HEI a municipality receives. Trade schools are part of the professional and technological education programs in Brazil and were first conceived in the early 1900s as schools to provide lower-income individuals with technical skills, such as carpentry. Between 1930 and 1970, the federal government expanded the supply of technical education to meet the growing demand for middle-skilled labor arising from industrialization, particularly for employment in the automobile and petrochemical industries. At this time, trade schools started providing a high school-equivalent education to attract young individuals (Faveri et al., 2018).

In the 1980s and 1990s, trade schools began offering training in computational skills and information technology (IT) to keep pace with changes in labor demand. Up until 2000, they remained focused on high school-equivalent education and short-term technical courses targeting hard skills. They were then incorporated into the Brazilian Higher Education system, although institutes still

offer technical training that does not require a high school diploma. Since then, trade schools have offered three main types of courses: integrated technical courses, in which students are enrolled in high school and a technical course simultaneously, and which last 3 to 4 years; subsequent technical courses, with a minimum duration of 2 years, in which students need to have a high school diploma to enroll; and undergraduate majors, which are similar to traditional universities[7]. Trade schools can also offer short-term courses, such as elderly caregiving and basic English, as well as graduate education, such as a Master's in Materials Engineering.

3 Data and sample characteristics

To obtain data on college openings, I construct a novel dataset using institutional documents and information provided on college websites. This dataset contains the opening year, the year authorization to operate was granted, municipality name, region and state, and the type of higher education institution (university or trade school) that opened as a part of the expansion program, ranging from 2003 to 2014. I define the treatment year as the year an HEI opened, and, in cases where this information is missing (15%), I replace it with the year authorization to operate was granted. I complement these data with the Higher Education Census data for 2002 and 2018, provided by the Ministry of Education, to obtain information on the number of HEIs pre- and post-policy and to determine whether receiving municipalities had a public or private institution before 2003.

For all of my analyses, I construct a panel dataset of municipalities, covering 2007 to 2019 for 9th-grade test scores, 2007 to 2017 for dropout rates from 9th to 12th grade, high school progression and graduation rates, and labor market outcomes, and 2010 to 2017 for college-going outcomes (share of students taking a college-entry exam, exam scores, and share who applied, were accepted, and enrolled in an HEI). The year restrictions are due to data availability. For all of my pre-tertiary education and college-going outcomes, I restrict my sample to public school students at baseline, but allow them to move to private high schools as a possible response to the policy. This is a key sample restriction, since public school students, on average, are lower-income and less likely to migrate to pursue tertiary education outside their municipality, and would likely be the most affected by this policy.

[7]Examples of integrated technical courses are IT and construction. Examples of subsequent technical courses are business administration, workplace safety, and nursing technician. Examples of undergraduate majors are energy engineering, architecture, and math.

For 9th-grade standardized test scores, I use data from the SAEB/Prova Brasil, a national test administered every two years. Test scores are available biennially from 1995 to 2019, but municipality identifiers are available only from 2007 onward. I thus restrict my test score analysis to municipalities treated between 2008 and 2014. For dropout rates, I use annual municipality-level rates provided by the School Census. To match the sample for test scores and high school graduation, I restrict my dropout rates analysis to municipalities treated between 2008 and 2014. For graduation and high school transition outcomes, I use the universe of K-12 student enrollment in Brazil from the 2007 to 2017 School Census. I build high school graduation outcomes from both 9th- and 10th-grade baselines: conditional on having enrolled in 9th/10th grade, does a student graduate on time or with up to a 2-year delay? High school transition uses the 9th grade baseline (last grade of primary school): conditional on enrolling in 9th grade, does the student enroll in 10th grade within 3 years? I then generate graduation and transition rates at the municipality level. My main outcome, high school completion rates, includes municipalities treated between 2008 and 2013, and high school transition rates include municipalities treated between 2008 and 2014.

My preferred graduation outcome uses the 10th-grade baseline for a few reasons. First, to construct graduation outcomes from the enrollment data, I need to observe students long enough into the future to determine whether they have completed 12th grade. Restricting to a 10th-grade baseline allows me to keep as much of the sample as possible while still restricting my baseline to a period prior to a critical dropout point (between grades 10 and 11) in the Brazilian system. This choice is further supported by the measured effect of the policy on dropout rates: I find no changes in 9th-grade dropout rates, but a decrease in dropout at 10th grade. Second, conditioning on students who have already started their first year of high school (10th grade in Brazil) implies that, for those students, supply-side constraints are unlikely to be binding. Therefore, it follows that dropping out of or graduating from high school is mostly mediated by demand-side factors, which are the focus of this paper. For robustness, I present high school completion estimates using the 9th-grade baseline, which are similar to my preferred estimates.

For both 9th- and 10th-grade baseline graduation outcomes, a limitation of using enrollment data is that I must infer that students who reach 12th grade and then disappear from the K-12 enrollment system have graduated. To bound my estimates, I use 12th-grade yearly municipality dropout rates and present them in Table A1 in the Appendix. The results do not change, and, in some cases, the point estimates are larger. To avoid relying on imputed data, I use the outcome constructed solely from enrollment data as my main outcome.

College-going outcomes stem from the unified selection system, SISU, which has been widely used for admissions to public HEIs since 2010. This contains individual-level information on the applicant's college-entry exam scores, municipality of residence, school municipality, whether they were accepted into college using SISU, and whether they enrolled. I aggregate this information to generate municipality-level mean test scores, share of 17- to 25-year-olds taking the national college entry exam, and application, acceptance, and enrollment rates from 2010 to 2017. It is worth noting that not all HEIs use the national exam and SISU for admission. Thus, the acceptance and enrollment rates calculated from this dataset reflect college-going outcomes only for a specific subset of HEIs. However, nearly all schools opening as part of the expansion policy use SISU as their only form of admission, and these outcomes are therefore informative of whether students attend HEIs created by this policy. Due to data availability, I restrict my analysis of college-going to municipalities that received an HEI between 2011 and 2014.

To analyze possible changes in the local labor market, both in the short- and long-run, I employ administrative data from the Relação Anual de Informações Sociais (RAIS - Annual Report of Social Information), which contains the universe of formal sector employees and provides information on wages and hours worked, along with demographic characteristics such as gender, race, and educational attainment, and spans 2007 to 2017.

Throughout all of my estimates, I use data on municipality GDP per capita, the share of the agricultural sector in GDP, per capita education expenditure, Human Development Index (HDI), whether the local government is aligned with the federal government, the existence of an HEI pre-policy, and population size as pre-policy municipality-level controls. These data come from the IPEA Data database and are discretized into quintiles. In all of my estimations, I restrict my sample to treated municipalities and to controls located at least 45 kilometers from a treated municipality.

3.1 Placement of HEIs

Placement of HEIs in this policy was not random. The federal government followed several criteria to select receiving municipalities, so treated units may differ from control units in many ways. Table 1 presents descriptive statistics on municipality characteristics, students' demographic characteristics, and pre-tertiary outcomes at baseline for treated and control municipalities, as well as for those receiving a trade school or a university.

Overall, treated municipalities have higher GDP per capita, are less agriculture-intensive, have

Table 1: Pretreatment Characteristics by Treatment Status and Type of Treatment Received

	Treated (1)	Control (2)	P-value	Trade Schools (4)	Universities (5)	P-value
Municipality Characteristics						
GDP per capita (R$)	6,428.41	4,682.67	0.00	6,403.90	6,161.99	0.18
Share of Agriculture in GDP	0.13	0.29	0.00	0.12	0.16	0.00
HDI	0.57	0.50	0.00	0.57	0.55	0.00
Mayor aligned with federal government	0.09	0.08	0.09	0.09	0.09	0.86
Education spending per capita (R$)	138.10	238.16	0.00	137.54	155.33	0.00
Had an HEI before the policy	0.42	0.05	0.00	0.43	0.30	0.00
High Secondary Premium	0.58	0.51	0.02	0.58	0.57	0.88
High College Premium	0.62	0.55	0.02	0.62	0.63	0.86
Number of Bolsa Familia Families (per capita)	0.07	0.09	0.00	0.07	0.08	0.22
Students' Characteristics - 9th Grade						
White	0.35	0.34	0.70	0.35	0.34	0.67
Working	0.22	0.27	0.00	0.21	0.23	0.07
Female	0.55	0.55	0.71	0.55	0.56	0.04
Mother completed High School	0.31	0.26	0.00	0.31	0.31	0.73
Mother completed higher education	0.09	0.09	0.20	0.09	0.09	0.94
Outcomes at baseline						
Portuguese Scores - 9th Grade	-0.39	-0.47	0.00	-0.38	-0.40	0.55
Math Scores - 9th Grade	-0.17	-0.21	0.01	-0.16	-0.19	0.52
High School Graduation	0.55	0.59	0.00	0.55	0.53	0.09
Dropout Rate - 9th Grade	7.57	7.81	0.50	7.15	8.63	0.08
Dropout Rate - 10th Grade	18.21	16.35	0.00	17.70	19.01	0.33
Dropout Rate - 11th Grade	14.31	12.15	0.00	13.97	14.54	0.61
Dropout Rate - 12th Grade	10.62	8.28	0.00	10.43	9.99	0.64
Observations	379	2,180		305	46	

Note: Test scores are z-scores standardized with respect to the population mean for each year, which includes private schools in Brazil and, thus, is the reason why public schools' test scores for both treated and control groups are negative. High School Graduation is the share of students entering 10th grade in 2007 who completed high school within five years. Dropout rates are the 2007 dropout rates in urban public schools for each grade. 28 municipalities received both a university and a trade school, and are therefore left out of the trade school vs university breakdown.

a higher HDI, and lower K-12 education spending than control municipalities. Contrary to the federal government's stated criteria, 42% of treated municipalities also had a pre-existing HEI, compared with only 5% in the control group. They are also more likely to be classified as a high secondary and college premium municipality, and have fewer families enrolled in the conditional cash transfer program, Bolsa Familia. Treated municipalities have a lower share of students who reported working outside of the household and a higher share of students whose mothers had completed high school, although a similar share of students had mothers who completed higher education. Public school 9th-graders in treated municipalities have higher Portuguese and Math test scores than those in the control group. However, both groups are below the national average on the standardized tests, a result of the public school student sample restriction. On the other hand,

conditional on starting 10th grade, students in control municipalities were 4 percentage points more likely to graduate from high school. Treated municipalities also had higher dropout rates at all three grades of high school.

Since municipality characteristics may vary by type of treatment received, columns 4 and 5 of Table 1 compare pretreatment characteristics of municipalities that received a university versus those that received a trade school. The latter are less agriculture-intensive and have slightly higher HDI scores than those receiving a university. I do not find large differences in students' characteristics or pre-treatment outcomes between municipalities that received a trade school and those that received a university.

Table 2: Correlation Between Pre-Treatment Outcomes and Treatment Status

	Independent Variable (2007 level)						
				Dropout rate			
	Graduate on time	Eventually graduate	Start HS	9th grade	10th grade	11th grade	12th grade
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable							
Treated	-1.005***	-0.770**	0.552	-0.017**	0.007	0.012**	0.008
	(0.370)	(0.374)	(0.464)	(0.008)	(0.004)	(0.005)	(0.006)
Observations	2,400	2,400	2,400	2,525	2,525	2,525	2,525

Note: Each column reports the probit coefficient from a regression of treatment status on the 2007 level of the outcome in the column header. The sample includes one observation per municipality: municipalities that received their first public HEI between 2008 and 2014 and the never-treated comparison municipalities used in the main results. Graduation outcomes refer to the 2007 10th-grade cohort; Start HS refers to the 2007 9th-grade cohort. All specifications include the same municipality-level control variables as the main results, measured pre-policy and discretized. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

One possible concern is that treated localities were chosen based on pre-treatment outcomes. For instance, municipalities with high secondary graduation rates pre-policy could have received new HEIs at higher rates due to increased demand for tertiary education. To assess if this is the case, I analyze whether pre-treatment high school graduation, transition and dropout rates predict treatment assignment by running a probit regression of treatment status on pre-policy outcomes and municipality-level controls used in the main analysis. Estimates are presented in Table 2. Although a lower 9th-grade dropout rate is correlated with receiving a new HEI, worse outcomes at high school predict treatment: treated municipalities had lower high school completion and higher 11th-grade dropout rates pre-policy. This relationship is expected, given the policy's stated goal of serving underserved and underdeveloped regions. Because my design compares changes rather than levels, these baseline differences do not by themselves bias my estimates. To assess differences in trends, I present event-study estimates in the main results section.

I deal with the observed differences between treated and control municipalities in two ways. First, in a difference-in-differences framework, control and treated units need to be comparable in trends but not in levels. I show in the results section that there is little evidence of pre-trends, which supports the comparisons between these two groups. Second, I control for an array of municipality characteristics that might have played a role in the federal government selection criteria. I expand these to include a political alignment dummy that identifies whether the local government was aligned with the federal government, following Arroyo and Petterini (2020), and per-capita educational expenditure to account for potential differences in pre-treatment school quality. As a robustness check, I include the pre-policy number of families receiving the conditional cash transfer program, Bolsa Familia. This is relevant because the program's conditionality requires students to be enrolled in school. The results do not change when including this additional control.

4 Methods

To estimate the effects of higher education expansion on local students' pre-tertiary education and college-going outcomes, I compare treated versus never-treated municipalities before and after an HEI's opening in a difference-in-differences setting. Equation 1 presents a two-way-fixed-effect estimation for that comparison.

$$y_{m,t} = \alpha + \beta ReceivedHEI_m \cdot Post_{t,m} + \gamma_t + \mu_m + \Pi \mathbf{X}_m + \epsilon_{m,t} \tag{1}$$

Where $y_{m,t}$ is the outcome of interest in municipality m at time t; $ReceivedHEI_m$ is the treatment variable, which equals one if the municipality received a higher education institution as part of this program and zero otherwise; $Post_{t,m}$ equals one if the period t is post-treatment in municipality m and zero otherwise; γ_t are year fixed effects to control for national trends in outcomes; μ_m are municipality fixed effects; and $\mathbf{X}$ is a vector of discretized pre-treatment municipality characteristics, which controls for factors taken into account by the Federal Government to assign treatment[8]. The main coefficient of interest is β, which in this specification does not vary with time.

Since the policy is staggered and the treatment years in my analysis range from 2008 to 2014, I employ the staggered-timing difference-in-differences estimator proposed by Callaway and

[8]Pre-tertiary outcomes are the likelihood of transitioning into high school, high school graduation, 9th grade Portuguese and Math scores, and average college-entry test scores. College-going outcomes are the share of 17- to 25-year-olds who take the college-entry exam, apply to college, are accepted, and enroll. Labor market outcomes are log wages by educational level.

Sant'Anna (2021) and include municipality-level controls. This allows for heterogeneous effects over time that should be relevant to this policy, as well as mitigate concerns about employing a Two-Way Fixed Effects estimator in a staggered treatment setting (Goodman-Bacon, 2021). In this approach, the inverse-probability-weighted (IPW) version of the average treatment effect on the treated is given by Equation 2.

$$ATT(g,t) = E\left[\left(\frac{G_g}{\mathbb{E}[G_g]} - \frac{\frac{p_g(X)(1-G_g)}{1-p_g(X)}}{\mathbb{E}\left[\frac{p_g(X)(1-G_g)}{1-p_g(X)}\right]}\right)\cdot(Y_t - Y_{g-1})\right] \tag{2}$$

where $ATT(g,t)$ is the average treatment effect measured at time t, $t \geq g$, for all municipalities that received an HEI starting in time g. $G_g = 1$ if a municipality is part of the group g, Y_t is the outcome for all units in the post-period of interest, t, and Y_{g-1} is the outcome for all units in the period immediately before when treatment started, $g-1$. $p_g(X) = P(G = g \mid X, G \in \{g, \text{control}\})$ is a propensity score calculated using a set of municipality-level covariates measured pre-policy that captures the likelihood of belonging to cohort g conditional on covariates X and being either in cohort g or the control group. $\mathbb{E}[G_g]$ and $\mathbb{E}\left[\frac{p_g(X)(1-G_g)}{1-p_g(X)}\right]$ are normalizations to ensure that weights sum to one for both treated and control groups. In practice, I use the doubly robust (DR-IPW) version of this estimator, which additionally subtracts from $(Y_t - Y_{g-1})$ the change in the outcome predicted for control municipalities with the same pre-policy characteristics, $m_{g,t}(X)$, and is consistent if either the propensity score or the outcome model is correctly specified (Sant'Anna and Zhao, 2020).

ATTs for each time t and treated cohort g are then grouped following Equation 3, using the weighting function $w(g,t)$. This will allow me to estimate an overall average effect of the policy, θ, which I restrict to up to three periods following the opening of an HEI ($0 \leq t - g \leq 3$).

$$\theta = \sum_{g \in \mathcal{G}} \sum_{t=2}^{\mathcal{T}} w(g,t) \cdot ATT(g,t) \tag{3}$$

My main identification assumption is that, in the absence of treatment, treated and untreated municipalities would have followed the same trends in the outcomes over time. With the inclusion of controls, the parallel trends assumption needs to hold conditional on municipality characteristics. Although this hypothesis cannot be formally tested, I use event-study specifications as described in Equation 4 to provide evidence of no differential pre-trends. This specification allows for the aggregation of ATTs across all treated cohorts, g, for each pre- and post-policy period, t. When

$t < g$, the estimated ATT works as a placebo test that can provide suggestive evidence of whether the parallel trends assumption is met. Control municipalities are at least 45 kilometers away from a treated municipality and are never treated. Municipalities treated before the period for which data are available are excluded from my analysis.

$$\theta(\ell) = \sum_{g \in G} w(g, \ell)\, ATT(g, g + \ell), \quad \text{for each event time } \ell, \tag{4}$$

where $\ell = t - g$ denotes the number of periods relative to the opening of an HEI.

5 The impacts of higher education opportunities on earlier educational decisions

Table 3 shows the effects of higher education expansion on pre-tertiary outcomes, using Callaway and Sant'Anna (2021) and measured within 3 years of an HEI's opening. The opening of a new HEI in a student's municipality does not affect the likelihood that 9th-graders transition to high school in the following three years or 9th-grade dropout rates. However, persistence increases in all high school grades: 10th graders are 1.2 percentage points less likely to drop out, while dropout rates for 11th and 12th graders decrease by 0.8 and 0.6 percentage points, respectively. These effects represent a 5–6% decrease relative to the baseline mean. Figure 2 shows the timing of these effects. While 10th-grade dropout decreases in the period immediately after a new campus opens, the effects in 11th and 12th grade appear only two years after treatment. Tenth graders are thus the most responsive to new college opportunities.

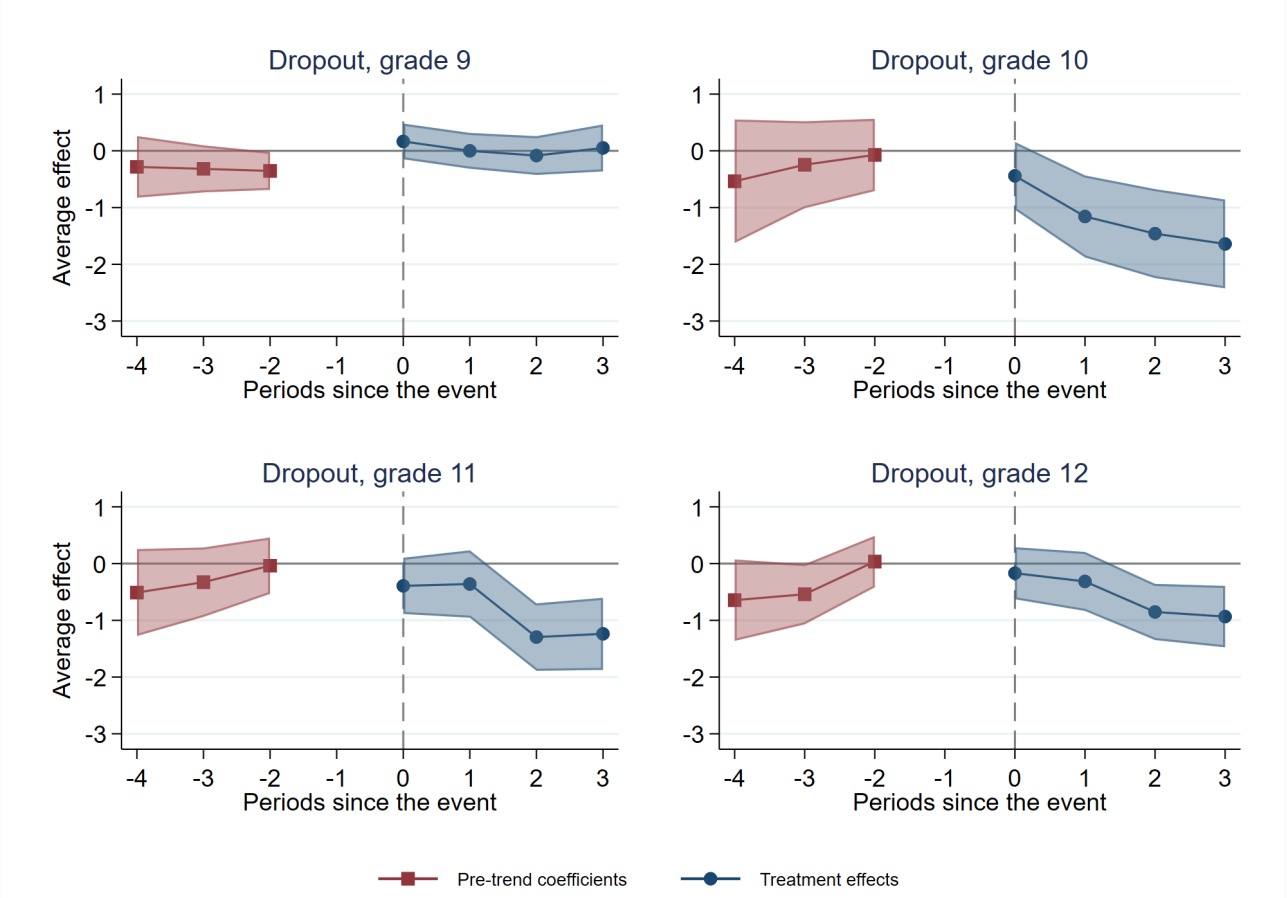
Table 3: Persistence in Secondary School

		Dropout rate					
	Start HS	9th grade	10th grade	11th grade	12th grade	Graduate on time	Eventually graduate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated	0.005*	0.034	-1.174***	-0.822***	-0.568***	-0.003	0.008**
	(0.003)	(0.149)	(0.304)	(0.251)	(0.219)	(0.003)	(0.003)
Observations	19,208	35,241	35,241	35,241	35,241	21,616	16,814
Pre-policy Mean	0.79	7.57	18.21	14.31	10.62	0.46	0.54

Note: Standard errors in parentheses, clustered at the municipality level. All specifications include municipality-level control variables, measured pre-policy and discretized. ATTs are calculated using doubly robust inverse probability weighting (DR-IPW) and aggregated to generate an overall average effect following Callaway and Sant'Anna (2021). Observations vary between outcomes due to years of data availability. Dropout estimates (columns 2–5) use a common sample of municipality-years with non-missing dropout rates for all four grades. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Given that students increase persistence in all grades of high school following the opening of an

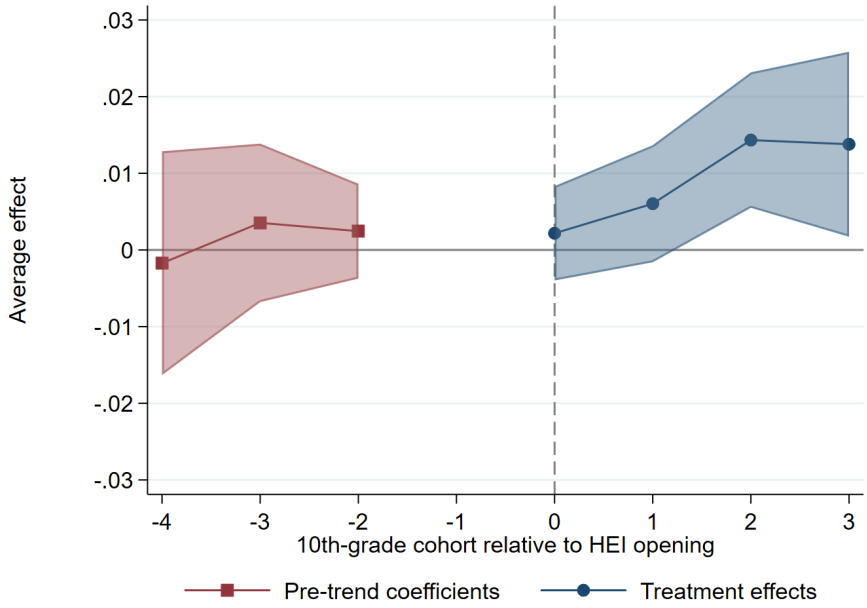
Figure 2: Effect of new HEIs on dropout rates



Note: This figure plots event-study estimates following Callaway and Sant'Anna (2021) (Equation 4), using never-treated units as the comparison group. Event time 0 is the year that the new higher education institution opens in the municipality. Points are estimates and bars are 95% confidence bands, with standard errors clustered at the municipality level. Table 3 reports the aggregated average effect.

HEI, I then assess whether this translates to an increase in the likelihood of completion. Columns 6 and 7 show effects on high school completion for students who were in 10th grade at the time of treatment, and summarize the effects for the first three exposed cohorts. Treated students are not more or less likely to graduate from high school on time. However, they are nearly 1 percentage point more likely to graduate with up to a 2-year delay, which corresponds to a 1.5% increase relative to the pre-policy mean. This corresponds to over 20,000 additional high school graduates in the first three exposed cohorts following the policy. Figure 3 shows that the effects on high school graduation increase over time, similarly to the observed pattern for dropout rates at 10th grade. In both sets of event studies, there is limited evidence of differential trends pre-policy.

Figure 3: Likelihood of graduating High School with up to a 2-year delay



Note: This figure plots event-study estimates following Callaway and Sant'Anna (2021) (Equation 4), using never-treated units as the comparison group. Event time 0 is the cohort that was in 10th grade when a new higher education institution opened in the municipality. Points are estimates and bars are 95% confidence bands, with standard errors clustered at the municipality level. Table 3 reports the aggregated average effect.

Table 4 presents performance outcomes for 9th-graders and for students exiting high school who choose to take the college-entry exam. I find no changes in either 9th-graders' Portuguese and Math performance or college-entry exam scores. I do, however, find meaningful increases in the share of local young adults taking an optional college exam. The share of students writing the exam following the opening of an HEI increases by 2 percentage points, a 20% increase from the pre-policy mean. This increase in take-up does not seem to come at the expense of performance, as measured by college-entry exam scores.

Table 4: Performance Outcomes

	9th Grade Math (1)	9th Grade Portuguese (2)	College Exam Scores (3)	Share Taking Exam (4)
Treated	0.01	0.01	-0.62	0.02***
	(0.01)	(0.01)	(1.38)	(0.00)
Observations	17,548	17,548	10,576	10,576
Pre-policy Mean	-0.17	-0.39	525.48	0.10

Note: Standard errors in parentheses, clustered at the municipality level. All specifications include municipality-level control variables, measured pre-policy and discretized. ATTs are calculated using doubly robust inverse probability weighting (DR-IPW) and aggregated to generate an overall average effect following Callaway and Sant'Anna (2021). Observations vary between outcomes due to years of data availability. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.1 Does the type of Higher Education matter?

The positive effects above are comparable in magnitude with those in the Indian setting, as measured by Jagnani and Khanna (2020). Unlike that setting, however, the Brazilian policy did not only expand existing 4-year universities but also established new 4-year higher education institutions and 2-year trade schools. This policy's unique feature allows for testing whether other types of higher education opportunities can shift pre-tertiary students' schooling decisions. In this section, I examine whether the average effects measured above are driven by a specific type of HEI.

Table 5 shows the differential impacts of new HEIs on secondary schooling persistence by type of HEI received. Although not statistically different, 9th graders in municipalities that received a university are nearly 2 percentage points more likely to start high school in subsequent years, with no effect for students in localities that received a trade school. Although trade schools do not seem to drive students to start high school in the first place, they meaningfully change persistence in high school, decreasing dropout rates at all grades and increasing high school graduation.

This suggests that the opportunity to access a trade school is more salient to students later in their schooling and who would have started high school in the absence of the policy. University openings, on the other hand, might change the composition of who starts high school in the first place. This difference in the composition of high school is further evidenced when comparing the 9th- and 10th-grade baseline estimates of the likelihood of eventually graduating, Table A3 in the Appendix. When conditioning on 9th-grade enrollment, I find a large 4% increase in the likelihood of eventually completing high school for students exposed to a university, although only marginally significant and not statistically different from the effect for trade schools. This effect completely dissipates if conditioning on enrollment in 10th grade, likely due to changes in high school composition driven by the newly opened university.

Table 5: Persistence in Secondary School, by Type of HEI Received

		Dropout rate					
	Start HS	9th grade	10th grade	11th grade	12th grade	Graduate on time	Eventually graduate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Received a Trade School	0.003	0.038	-1.239***	-0.760***	-0.584**	-0.004	0.010**
	(0.003)	(0.159)	(0.333)	(0.277)	(0.238)	(0.004)	(0.004)
Observations	18,616	34,149	34,149	34,149	34,149	20,950	16,296
Pre-policy Mean	0.79	7.15	17.70	13.97	10.43	0.47	0.55
Received a University	0.016*	0.225	-0.803	-0.949	-0.321	-0.007	-0.003
	(0.009)	(0.436)	(0.739)	(0.593)	(0.563)	(0.008)	(0.010)
Observations	16,544	30,355	30,355	30,355	30,355	18,619	14,483
Pre-policy Mean	0.78	8.63	19.01	14.54	9.99	0.44	0.53
P-value (diff)	0.18	0.69	0.59	0.77	0.67	0.74	0.21
Bootstrap p-value (diff)	0.18	0.70	0.60	0.77	0.66	0.77	0.23

Note: Standard errors in parentheses, clustered at the municipality level. Each type of HEI is compared with the same never-treated municipalities. All specifications include municipality-level control variables, measured pre-policy and discretized. ATTs are calculated using doubly robust inverse probability weighting (DR-IPW) and aggregated to generate an overall average effect following Callaway and Sant'Anna (2021). P-value (diff) tests the equality of the trade school and university estimates using $\sqrt{se_{TS}^2 + se_U^2}$, which ignores that both groups share the same comparison municipalities. Bootstrap p-value (diff) uses the standard error of the difference from a municipality-level bootstrap (200 replications, stratified by opening cohort), re-estimating both ATTs in each draw. Observations vary between outcomes due to years of data availability. Dropout estimates (columns 2–5) use a common sample of municipality-years with non-missing dropout rates for all four grades. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.2 Who benefits from higher education expansion?

Since all students in my sample are public school students, the overall effects are informative of how the policy affects lower-income, highly constrained students. However, within public schools, some students are more disadvantaged than others. In this section, I analyze whether the policy had heterogeneous effects across racial groups. Pre-policy, non-white students are 6 percentage points less likely to ever enroll in high school (70% versus 76% for white students) and 6 percentage points less likely to ever graduate (53% compared to 59% for white students).

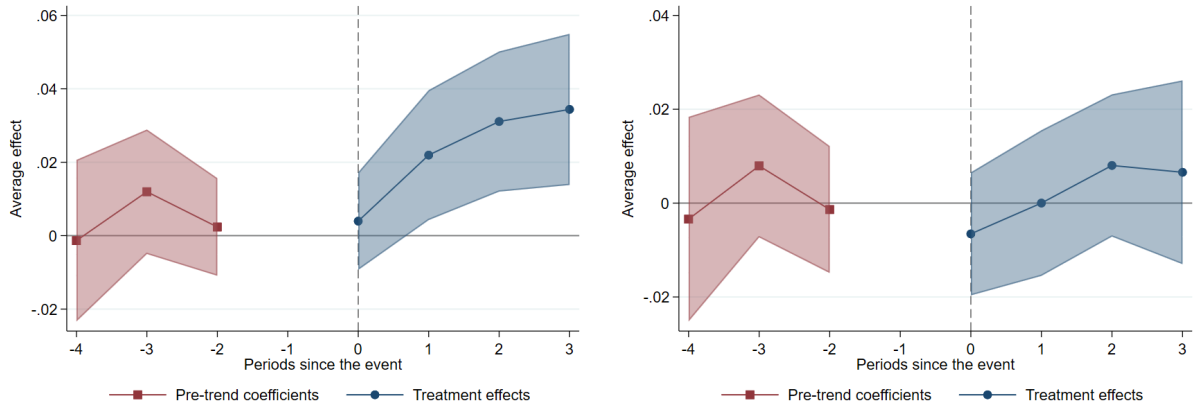
Table 6 shows that the policy disproportionately benefits non-white students. They are 2 percentage points more likely to ever enroll in high school (a 3% increase from the pre-policy mean), closing 1/3 of the pre-policy gap between white and non-white students. Although only marginally statistically different, estimates on the likelihood of ever graduating high school for non-white students are four times as large as those for white students: conditional on being enrolled in 10th grade, non-white students are 2 percentage points more likely to eventually graduate high school, a 4% increase from the pre-policy mean and 1/3 of the pre-policy gap relative to white students. Figure 4 shows the event studies for the share of non-white and white students starting high school following the policy. The effect is increasing over time for non-white students, and flat for white students.

Table 6: Persistence in Secondary School, by Race

	Start HS (1)	Graduate on time (2)	Eventually graduate (3)
White	0.001	-0.006	0.005
	(0.007)	(0.007)	(0.008)
Observations	17,529	20,413	15,807
Pre-policy Mean	0.76	0.50	0.59
Non-white	0.021***	0.004	0.020***
	(0.007)	(0.006)	(0.007)
Observations	17,529	20,413	15,807
Pre-policy Mean	0.70	0.42	0.53
P-value (diff)	0.05	0.27	0.18
Bootstrap p-value (diff)	0.02	0.13	0.08

Note: Standard errors in parentheses, clustered at the municipality level. Estimates for white and non-white students are obtained in separate regressions that use the same treated and never-treated municipalities, and each column keeps the same municipality-years for both groups. All specifications include municipality-level control variables, measured pre-policy and discretized. ATTs are calculated using doubly robust inverse probability weighting (DR-IPW) and aggregated to generate an overall average effect following Callaway and Sant'Anna (2021). P-value (diff) tests the equality of the two estimates using $\sqrt{se_1^2 + se_2^2}$, which ignores that both groups come from the same municipalities. Bootstrap p-value (diff) uses the standard error of the difference from a municipality-level bootstrap (200 replications, stratified by opening cohort), re-estimating both ATTs in each draw. Observations vary between outcomes due to years of data availability. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 4: Share of non-white (left) and white (right) 9th Graders who Start High School



Note: This figure plots event-study estimates following Callaway and Sant'Anna (2021) (Equation 4), using never-treated units as the comparison group. Event time 0 is the cohort that was in 9th grade when a new higher education institution opened in the municipality. Points are estimates and bars are 95% confidence bands, with standard errors clustered at the municipality level. Table 6 reports the aggregated average effect.

6 Why and when do higher education opportunities matter? A theoretical framework

The results in Section 5 show that local students exposed to a new HEI are more likely to persist through and graduate from high school. Yet an increased provision of higher education in a student's municipality does not directly affect the availability or costs of secondary schooling. Why, then, would a new college keep students in high school? In this section, I develop a simple model to illustrate how students decide whether to complete high school and how the availability of higher education factors into this decision. The model has two goals. First, it identifies the mechanisms through which the availability of higher education influences the decision to graduate from high school, and the specific parameters that may change with an increased supply of local HEIs. Second, by identifying the conditions under which each mechanism operates, it helps reconcile mixed findings stemming from different contexts: effects similar to mine have been documented in India and Kenya (Jagnani and Khanna, 2020; Erbabian, 2026), but opposite effects have been found in the United States and Chile (Bedard, 2001; Erbabian, 2026).

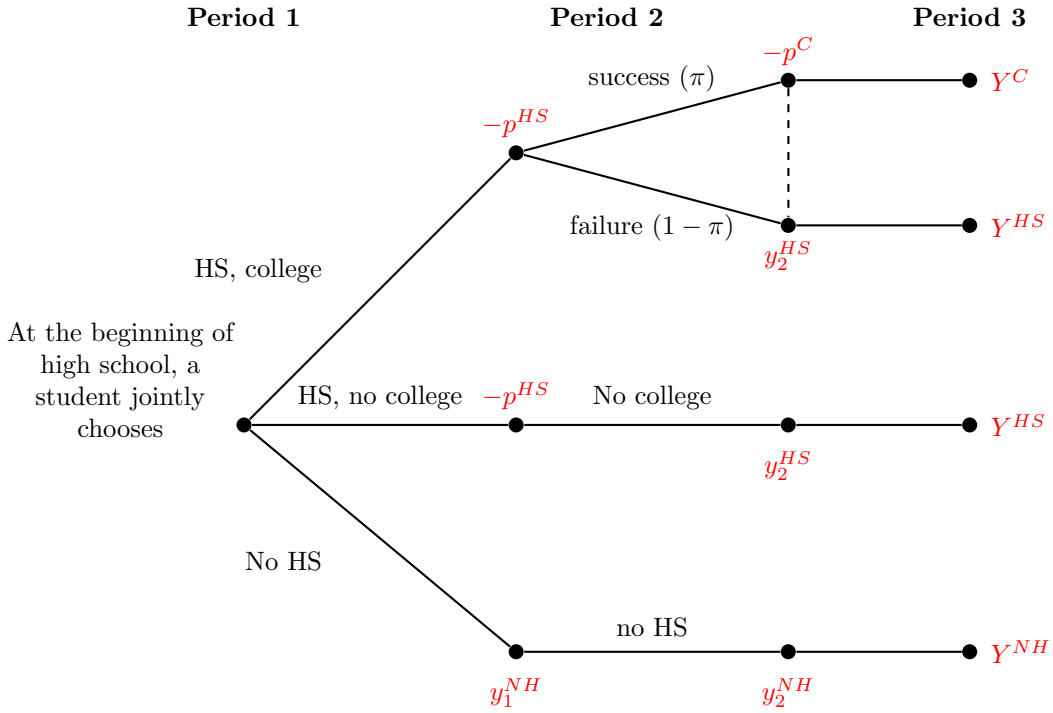
6.1 A student's decision: baseline model

Figure 5 illustrates the decision of a forward-looking student just starting 10th grade, the first year of high school in the Brazilian system. At that point, a student must simultaneously decide whether to continue in and graduate from high school and whether to *attempt to* graduate from higher education. Three possible paths are available to a student. If they decide to drop out of high school, they automatically choose not to pursue higher education, follow the (no high school, no college) path, and immediately join the labor force in period 1. On the other hand, if a student graduates from high school, they remain in high school in period 1, and then follow one of two paths depending on their higher education decision. If choosing not to attend college, they follow the (high school, no college) path and join the labor force in period 2, immediately after completing secondary education. If choosing the (high school, college) path, they will successfully enroll in and graduate from higher education in period 2 with probability π_c, and fail and join the labor force with a high school degree $(1-\pi_c)$ – similarly to those who chose the (high school, no college) path[9]. This uncertainty is key to modeling the Brazilian setting, where public higher education spots are

[9]In this framework, I abstract away selection issues that would lead high school-level earnings to be different for a student who chose high school and no college, and a student who chose high school, college but failed to progress to higher education.

limited, acceptance rates are below 10%. Lower-income public school students, the focus of this paper, therefore face significant constraints in attending higher education and may not attend even after choosing to try.

Figure 5: Decision tree with joint choice of high school and college attendance



Note: This figure illustrates the schooling decision of a student at the beginning of high school (10th grade). The student chooses among three paths: dropping out of high school (No HS), completing high school without attempting college (HS, no college), or completing high school and attempting college (HS, college). Period 1 corresponds to the high school years, Period 2 to the typical duration of college, and Period 3 to the remainder of working life. Red terms denote costs and earnings: p^{HS} and p^{C} are the direct costs of attending high school and college; y^{NH} and y^{HS} are per-period earnings of workers without and with a high school degree in Periods 1 and 2; and Y^{NH}, Y^{HS}, and Y^{C} are earnings in Period 3 by final educational attainment. A student who attempts college is admitted with probability π and otherwise earns high school wages; the dashed line connects the two possible outcomes of this admission lottery, which is realized after the schooling decision is made.

In this setting, both high school and college attendance incur transportation/relocation and material costs, and a student must pay p^{HS} and p^{C} if continuing in high school in period 1 and higher education in period 2, respectively. For simplicity, I abstract from tuition-related costs. Since public high school is readily available and HEIs opening as part of the Brazilian policy are free, this is a reasonable abstraction[10]. When a student enters the labor force, they start earning a

[10]Private higher education is also a relevant share of tertiary institutions in Brazil. For simplicity and due to data availability, I do not directly model private HEI enrollment as a student outcome. For my sample, which consists only of public school students coming from lower-income households, and for most of this time period, when loans to attend private colleges were limited, this is an acceptable simplification.

constant wage specific to their final educational level: y^{NH} for high school dropouts, y^{HS} for high school graduates, and y^{C} for college graduates.

In a setting where higher education is not available or where distance and moving costs are so high for the median student that college is not in their choice set, students will choose only between (high school, no college) and (no high school, no college), or simply whether to graduate high school. In this case, if the high school premium is high enough to justify the forgone earnings and attendance costs in period 1, students will choose to stay in school rather than drop out and join the labor force. Now consider a world where higher education is available. In addition to the high school decision, a high school freshman will also choose whether to pursue higher education after graduation. In this scenario, choosing to graduate high school opens an additional path, (high school, college), which might lead to greater lifetime earnings.

The decision to attend high school varies greatly depending on which world a student is in. Assuming students have certainty about wages by education level, when higher education is not an option, the expected earnings from graduating high school are simply the high school wage, and the schooling decision becomes a simple comparison of lifetime earnings with and without graduating high school. On the other hand, with the added possibility of pursuing higher education (and thus accessing higher-education wages), students now face uncertainty about their future earnings when choosing the path corresponding to (high school, college). Or: $\mathbb{E}[y^{HS}_{hs,nocol}] = y^{HS}$, while $\mathbb{E}[y^{HS}_{hs,col}] = (1-\pi_c)y^{HS} + \pi_c y^{C}$. Simply put, this additional educational opportunity represents a potential increase in the expected returns to graduating high school, provided that wage levels vary considerably by educational level in the local labor market.

I formalize the decision-making process illustrated in Figure 5 as a simple three-period problem described below. The student chooses their educational path from $(s,d) \in \{(0,0),(1,0),(1,1)\}$, where $s=1$ if they choose to graduate high school and 0 otherwise, and $d=1$ if they choose to pursue college. Because graduating from college requires first graduating from high school, the pair $(0,1)$ is not feasible. The student maximizes the expected utility:

$$\max_{(s,d)\in\{(0,0),(1,0),(1,1)\}} U(C_1) + \beta_2\,\mathbb{E}\big[U(C_2)\big] + \beta_3\,\mathbb{E}\big[U(C_3)\big], \tag{5}$$

where $u = ln(c)$ is a standard strictly increasing and strictly concave utility function and $\beta_2 = \beta^3$ and $\beta_3 \equiv \sum_{t=8}^{N}\beta^{t-1}$. β is the student's discount factor. The consumption streams for each period are:

Period 1. Period 1 lasts 3 years, the duration of high school in Brazil, grades 10 to 12:

$$C_1 = \bar{k}_1 + (1-s)\, y_1^{NH} - s\, p^H,$$

where $\bar{k}$ is the level of household assets, which includes the earnings of other family members, $\bar{k}_1 = 3\bar{k}$ and $y_1^{NH} = 3y^{NH}$.

Period 2. Period 2 lasts 4 years, the typical duration of tertiary education in Brazil:

$$C_2 = \begin{cases} \bar{k}_2 + y_2^{NH} & \text{if } (s,d) = (0,0), \\ \bar{k}_2 + y_2^{HS} & \text{if } (s,d) = (1,0), \\ \bar{k}_2 + y_2^{HS} & \text{if } (s,d) = (1,1) \text{ and not admitted with probability } (1-\pi_c), \\ \bar{k}_2 - p^c & \text{if } (s,d) = (1,1) \text{ and admitted with probability } \pi_c. \end{cases}$$

where $\bar{k}_2 = 4\bar{k}$, $y_2^{NH} = 4y^{NH}$ and $y_2^{HS} = 4y^{HS}$.

Period 3. Period 3 summarizes the remainder of an individual's working life as the present value of the per-year consumption C_3. After period 2, an individual's education level is fixed, and consumption will be constant for the following N periods during which the individual remains employed.

$$C_3 = \begin{cases} \bar{k} + y^{NH} & \text{if } (s,d) = (0,0), \\ \bar{k} + y^{HS} & \text{if } (s,d) = (1,0), \\ \bar{k} + y^{HS} & \text{if } (s,d) = (1,1) \text{ and not admitted}, \\ \bar{k} + y^{C} & \text{if } (s,d) = (1,1) \text{ and admitted}, \end{cases}$$

For simplicity, I assume $\bar{k}$ to be constant over time. Similarly, earnings by level of education are constant and known to students at the time of the schooling decision, and students' final educational attainment cannot change after this initial decision and the realization of the college draw. For paths $(0,0)$ and $(1,0)$, there is no uncertainty given that the student will not apply to college, and the only source of uncertainty is the college admission probability, $\pi_c \in (0,1)$.

Costs and earnings incurred by the students, as well as the timing for when they will join the labor force, are as described above. For simplicity, I will assume that costs are drawn independently and uniformly: $p^H \sim U[0, \bar{P}^H]$ and $p^c \sim U[0, \bar{P}^c]$. I further assume that $\bar{P}^H < \bar{k}_1$ and $\bar{P}^c < \bar{k}_2$

so that period-1 and period-2 consumption are strictly positive on all paths. A student will decide whether to follow $(0,0)$, $(1,0)$, or $(1,1)$ by comparing the expected lifetime utility they will receive from each path. Using $U(c) = \ln(c)$, the lifetime utilities of the three paths are:

$$V(0,0) = \ln(\bar{k}_1 + y_1^{NH}) + \beta_2 \ln(\bar{k}_2 + y_2^{NH}) + \beta_3 \ln(\bar{k} + y^{NH}), \tag{6}$$

$$V(1,0) = \ln(\bar{k}_1 - p^H) + \beta_2 \ln(\bar{k}_2 + y_2^{HS}) + \beta_3 \ln(\bar{k} + y^{HS}), \tag{7}$$

$$\begin{aligned} V(1,1) = \ln(\bar{k}_1 - p^H) &+ \beta_2 \big[(1-\pi_c)\ln(\bar{k}_2 + y_2^{HS}) + \pi_c \ln(\bar{k}_2 - p^c)\big] \\ &+ \beta_3 \big[(1-\pi_c)\ln(\bar{k} + y^{HS}) + \pi_c \ln(\bar{k} + y^C)\big]. \end{aligned} \tag{8}$$

By setting $V(1,0) = V(0,0)$, $V(1,1) = V(0,0)$ and $V(1,1) = V(1,0)$, I derive thresholds for when a student will prefer $(1,0)$ over $(0,0)$ (threshold $\bar{p}_{10}^H$), $(1,1)$ over $(0,0)$ (threshold $\bar{p}_{11|00}^c$) and $(1,1)$ over $(1,0)$ (threshold $\bar{p}_{11|10}^c$), respectively. The detailed derivation of these thresholds is in the Appendix. Combining the three thresholds, the student's decision rule is:

$$(s,d)^* = \begin{cases} (1,1) & \text{if } \big[p^H \le \bar{p}_{10}^H \text{ and } p^c \le \bar{p}_{11|10}^c\big] \text{ or } \big[p^H > \bar{p}_{10}^H \text{ and } p^c \le \bar{p}_{11|00}^c\big], \\ (1,0) & \text{if } p^H \le \bar{p}_{10}^H \text{ and } p^c > \bar{p}_{11|10}^c, \\ (0,0) & \text{if } p^H > \bar{p}_{10}^H \text{ and } p^c > \bar{p}_{11|00}^c. \end{cases} \tag{9}$$

where:

$$\bar{p}_{10}^H = \bar{k}_1 - (\bar{k}_1 + y_1^{NH}) \left(\frac{\bar{k}_2 + y_2^{NH}}{\bar{k}_2 + y_2^{HS}}\right)^{\beta_2} \left(\frac{\bar{k} + y^{NH}}{\bar{k} + y^{HS}}\right)^{\beta_3}. \tag{10}$$

$$\bar{p}_{11|10}^c = \bar{k}_2 - (\bar{k}_2 + y_2^{HS}) \left(\frac{\bar{k} + y^{HS}}{\bar{k} + y^C}\right)^{\beta_3/\beta_2}. \tag{11}$$

$$\bar{p}_{11|00}^c = \bar{k}_2 - \left[\left(\frac{\bar{k}_1 + y_1^{NH}}{\bar{k}_1 - p^H}\right)^{1/\beta_2} \frac{\bar{k}_2 + y_2^{NH}}{(\bar{k}_2 + y_2^{HS})^{1-\pi_c}} \left(\frac{\bar{k} + y^{NH}}{(\bar{k} + y^{HS})^{1-\pi_c}(\bar{k} + y^C)^{\pi_c}}\right)^{\beta_3/\beta_2}\right]^{1/\pi_c}. \tag{12}$$

Each threshold marks the point where a student is indifferent between two options: at $p^H = \bar{p}_{10}^H$, a student is indifferent between following the (no high school, no college) and the (high school, no college) paths; at $p^c = \bar{p}_{11|10}^c$, a student is indifferent between the (high school, college) and (high

school, no college) paths; and at $p^c = \bar{p}^c_{11|00}$, a student is indifferent between (high school, college) and (no high school, no college). In the absence of higher education, the relevant threshold would only be $\bar{p}^H_{10}$, which mediates the choice between (no high school, no college) and (high school, no college) and simply compares the value of staying in high school versus dropping out. With the added higher education opportunity, it is now possible for students who would drop out of high school via $\bar{p}^H_{10}$ to choose the (high school, college) path through the $\bar{p}^c_{11|00}$ threshold: for those with prohibitive p^H to choose high school alone, high school now becomes attainable when $p^c \leq \bar{p}^c_{11|00}$.

Let S^* denote the aggregate share of students who graduate high school. A student graduates high school either by choosing $(1,0)$ (so $p^H \leq \bar{p}^H_{10}$) *or* $(1,1)$ over $(0,0)$ (so $p^H > \bar{p}^H_{10}$ and $p^c \leq \bar{p}^c_{11|00}$). Hence:

$$S^* = \Pr(p^H \leq \bar{p}^H_{10}) + \Pr\big(p^H > \bar{p}^H_{10} \text{ and } p^c \leq \bar{p}^c_{11|00}\big). \tag{13}$$

Applying the law of total probability and independence of p^H and p^c:

$$S^* = \Pr(p^H \leq \bar{p}^H_{10}) + \int_{\bar{p}^H_{10}}^{\bar{P}^H} \Pr\big(p^c \leq \bar{p}^c_{11|00}(p^H)\big) \cdot f(p^H)\, dp^H,$$

where $f(p^H) = 1/\bar{P}^H$. Under $p^H \sim U[0, \bar{P}^H]$, $\Pr(p^H \leq \bar{p}^H_{10}) = \bar{p}^H_{10}/\bar{P}^H \equiv \theta_H$. Similarly, under $p^c \sim U[0, \bar{P}^c]$, $\Pr(p^c \leq \bar{p}^c_{11|00}(p^H)) = \bar{p}^c_{11|00}(p^H)/\bar{P}^c$. Therefore:

$$S^* = \theta_H + \frac{1}{\bar{P}^H \bar{P}^c} \int_{\bar{p}^H_{10}}^{\bar{P}^H} \bar{p}^c_{11|00}(p^H)\, dp^H. \tag{14}$$

The first term represents the share of students who find high school worthwhile on its own and will graduate high school to follow the (high school, no college) path, while the second is the share that would not graduate high school alone but is induced to complete the full college path when college costs are low enough. Note that S^* does *not* depend on $\bar{p}^c_{11|10}$, the threshold pertaining to students who move from $(1,0)$ to $(1,1)$, as these transitions do not impact the share of students graduating high school.

6.2 How can college supply impact pre-tertiary educational attainment?

Given the composition of S^* described above, changes in the share of students completing high school can stem from two distinct transitions. The first transition moves students from $(0,0)$

(no high school, no college) to $(1,0)$ (high school, no college). The threshold that governs this transition, $\bar{p}_{10}^{H}$, depends only on household assets and income, $\bar{k}$, and the local wage levels for primary-educated and high school graduate workers, y^{NH} and y^{HS}. Importantly, it does not depend on parameters related to higher education opportunities that might be directly affected by HEI expansion, such as college attendance costs, p^c, and the probability of attending higher education, π_c. Thus, movements from $(0,0)$ to $(1,0)$ do not reflect expanded access to higher education, but result solely from changes in the local labor market that alter the standalone returns to high school. The marginal student in this case completes high school because high school itself has become more valuable, not because of the possibility of attending college.

The second transition moves students from $(0,0)$ to $(1,1)$ and operates through the $\bar{p}_{11|00}^{c}$ threshold. Equation 12 shows that $\bar{p}_{11|00}^{c}$ depends not only on earnings by education level and household assets and income, $\bar{k}$, but also on college attendance costs (via the support $\bar{P}^c$) and the probability of attending higher education, π_c. Unlike the first transition, this transition can occur even if the returns to high school itself remain unchanged: an HEI expansion can shift $\bar{p}_{11|00}^{c}$ simply by making the returns to college more accessible, through lower attendance costs or a higher probability of attending. The marginal student in this case is one for whom high school alone is not worth completing (i.e., who prefers $(0,0)$ to $(1,0)$), but who is now willing to complete high school *because* of the possibility of attending college. For each transition, I next describe how the HEI expansion can move marginal students onto a high school completion path, and in which contexts such shifts are more likely to occur.

6.2.1 High school as an end: $(0,0) \rightarrow (1,0)$

Any change that directly affects the returns to completing high school could move students who would otherwise drop out to stay and finish. In the framework, these appear as changes in local wage levels for high school dropouts and high school graduates. As shown in the Appendix, either an increase in high school graduate wages or a decrease in high school dropout wages would predict an increase in high school completion.

But why would an increase in the local supply of higher education affect the local labor market? As the literature widely documents, new colleges can improve local infrastructure, change local labor supply, and attract new firms, expanding employment opportunities (Jagnani and Khanna, 2020; Vu, 2024; Cox, 2024). For this channel to explain my results, the labor market response would have to be immediate: the effects on dropout appear within three years of an HEI opening, and

the effects on completion appear among the first three exposed 10th-grade cohorts. Many of the channels documented in the literature, such as firm entry and infrastructure improvements, are likely to take longer to materialize. However, if high school graduates already in the labor market respond immediately by enrolling in college, this could meaningfully reduce the supply of high school graduates in these municipalities and expand employment opportunities that require a high school degree, effectively raising the returns to completing high school. New HEIs could have a similar effect by directly creating jobs that require a high school degree.

For this transition to be possible, the framework implies two necessary conditions. First, in the absence of the policy, dropping out must be an attractive option for at least some students. This implies that the high school premium cannot be too large; otherwise, students would already be completing high school, and the policy would have no one to move. Second, for students to choose to graduate high school following the policy, the high school premium must be positive and large enough to compensate for the earnings forgone while in school. Together, these conditions require the high school premium to be positive but modest: large enough to make graduating worthwhile for some students, but not so large that all students already graduate pre-policy.

Empirically, I test whether the arrival of an HEI changes the local labor market in ways that could explain my results. In particular, I examine whether local wages by educational level change within three years of the arrival of a new campus, which directly maps to the predictions of the framework.

6.2.2 High school as a gateway: $(0,0) \to (1,1)$

In addition to directly changing the returns to high school, new HEIs can also change the expected returns to completing secondary education by expanding college opportunities. This role of high school as a gateway is captured by transitions from (no high school, no college) directly to (high school, college) in the framework: students who would not otherwise graduate high school now do so *because* of the opportunity to attend higher education.

Two parameters specific to college access govern these transitions in the framework. First, new HEIs can raise local students' likelihood of college attendance, π_c, by increasing the number of available seats in their municipality. Second, by reducing the distance to the nearest free HEI, new campuses can substantially lower college attendance costs, shifting the cost distribution down (a lower $\bar{P}^c$) so that more students fall below the threshold $\bar{p}^c_{11|00}$. Both changes would increase the expected returns to the (high school, college) path, and could in turn drive students to stay in high

school in order to attend higher education. Beyond these college-specific parameters, increases in high school graduate and college graduate wages would also increase high school completion via this transition, although the former would also move some students into the (high school, no college) path.

As with the first transition, the framework implies two necessary conditions for this transition to occur. The first condition is the same as before: some students must choose (no high school, no college) in the absence of the policy. The second condition is specific to this move. For students to stay in high school only *because* of college, the college premium must be sufficiently high. That is, the returns to college must meaningfully increase the expected returns to high school for students who find high school alone not worth completing. Taken together, these conditions imply that this transition is most likely in contexts where standalone returns to high school are modest, but the college premium is high: that is, where returns to schooling are highly convex.

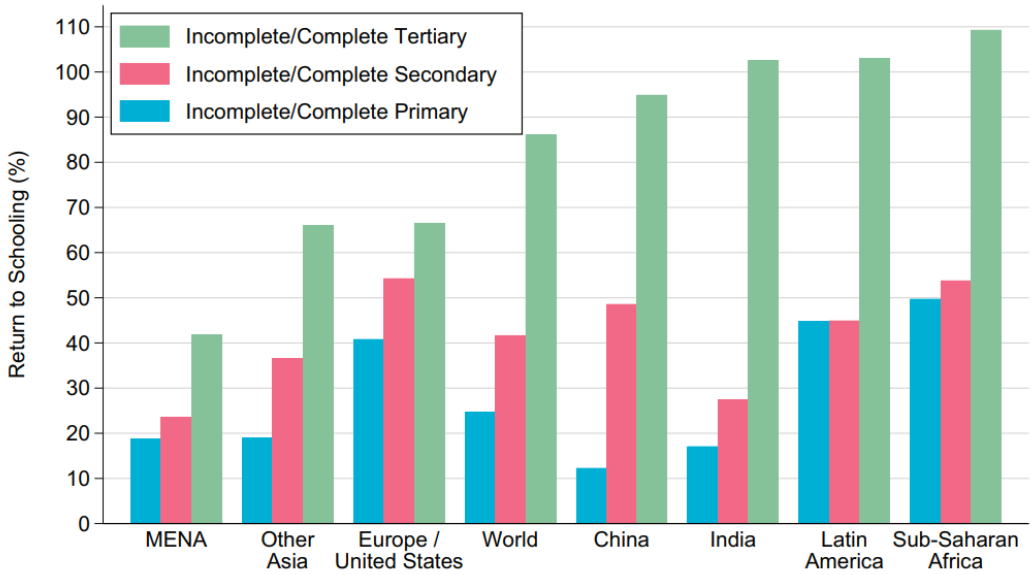
Empirically, I test whether college enrollment for local public school students changes following the opening of an HEI. This measure does not directly capture the college enrollment of the marginal students who stayed in high school in response to the new HEI. Instead, it measures an overall increase in the likelihood that local public school students enroll in a federal HEI, which forward-looking 10th-graders deciding whether to stay in school might take as a proxy for their own likelihood of attending college. However, an increase in college enrollment could also decrease the local supply of high school graduates (and open up new employment opportunities that require a high school degree), operating through the first transition. Enrollment alone therefore cannot distinguish between the two. I thus examine the role of the college premium in mediating my results. In the framework, the college premium matters for the high school decision only if students treat high school as a gateway to college earnings, and not the end goal. If high school completion responds similarly in low- and high-college-premium municipalities, this would be more consistent with transitions into (high school, no college) driving my results. If, instead, completion increases only where the college premium is high, this provides suggestive evidence that at least some of the response is driven by students who stay in high school because of the added opportunity to attend college.

6.2.3 Where should higher education opportunities matter?

The framework highlights the role of the underlying distribution of returns to schooling in determining whether higher education opportunities should matter for schooling demand at earlier

levels. In particular, in contexts where standalone returns to high school are modest, more students drop out at baseline, leaving more room for college opportunities to change high school decisions. Additionally, college opportunities should matter more for high school decisions where tertiary education brings a large jump in earnings relative to primary and secondary education.

Figure 6: Returns by region - Gethin (2025)



Note: This figure comes from Gethin (2025), published in the Quarterly Journal of Economics. These are OLS estimates of returns to each level of education, controlling for an age quartic, gender, and interactions between gender and the age quartic.

Figure 6 shows substantial variation in returns to schooling by education level across regions, as measured in Gethin (2025). In places like India, Latin America, and Sub-Saharan Africa, tertiary returns greatly exceed those for secondary schooling, whereas primary and high school returns are much closer in magnitude. In high-income countries in Europe and the United States, high school returns are significantly larger than those for primary education, and tertiary education gains are much smaller than those observed in LMICs. According to this simple model and the necessary conditions described above, settings like India, Latin America, and Sub-Saharan Africa are the most suitable for an expansion in college opportunities to increase high school completion. This is in line with the observed impacts in the literature for Kenya and India, and the opposite evidence in higher-income settings such as the United States and Chile.

7 High school as an end or a gateway? Evidence on mechanisms

Although I cannot empirically observe which path each student takes, the framework yields distinct testable predictions for the two transitions. For students to graduate high school as an end, the value of a high school degree must increase through higher wages or more employment opportunities. By contrast, for students to graduate high school as a gateway to college, new HEIs must change the parameters associated with college-going: the likelihood that a student attends college and the costs of attending. This transition also requires returns to college to be high enough to make high school worth completing for students who would otherwise drop out. Therefore, it should only occur in places where the college premium is high.

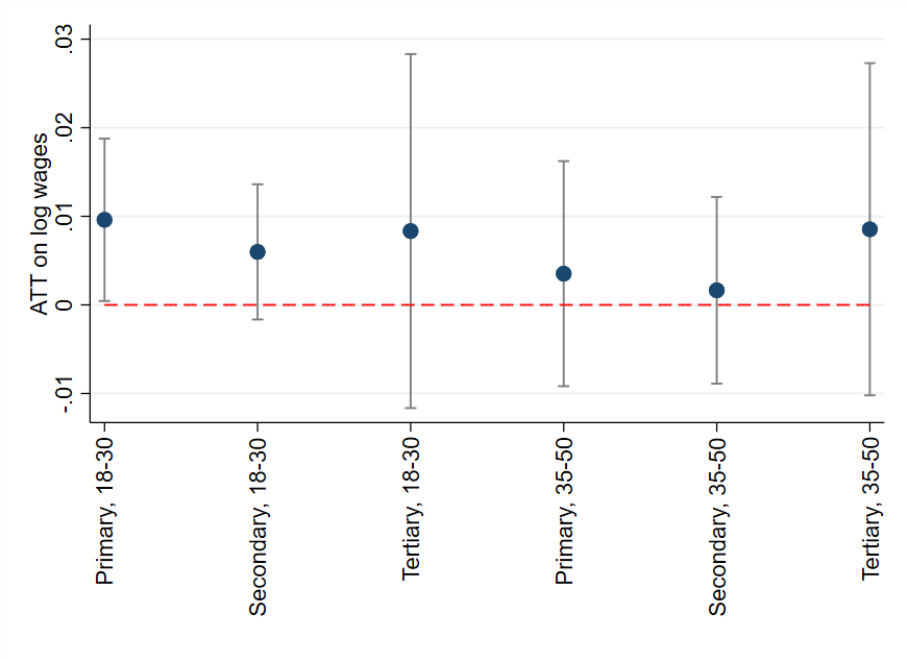
In this section, I first assess whether local wages change following the policy. I then show that new HEIs increase college enrollment among local public school students in the short run. This increase can be read in two ways. It may serve as a proxy for π_c, the likelihood of attending college, as perceived by students deciding whether to stay in school. Alternatively, it may open up jobs for new high school graduates, as those already in the labor market leave to enroll in college. To help distinguish between these two interpretations, I measure the differential effects of HEI openings by pre-policy municipality college premium.

7.1 Local wages

In this section, I assess whether the policy affected formal-labor-market wages in the short run (within 3 years of an HEI's opening) across education levels and ages. If new HEIs directly change the local labor market in ways that make staying in high school more worthwhile, students will be led to graduate high school at higher rates. Increases in college-graduate and high school graduate wages, as well as decreases in primary-educated wages, would all be consistent with an increase in high school graduation following the opening of an HEI.

Changes in local wages in the formal labor market do not seem to be driving the observed increases in high school completion. Point estimates are largest and significant only for primary-educated earnings among young workers, although all estimates shown in Figure 7 are indistinguishable from one another. These increases likely result from young individuals choosing to stay in school longer because of the policy, thus decreasing the supply of low-skilled labor in the short run. If primary wages are increasing as a result of the policy, the framework predicts that dropouts should increase, and high school graduation rates would move in the opposite direction of the

Figure 7: Effect on formal-sector wages by level of education and age



Note: This figure plots the average effect of a new higher education institution on municipalities' average log wage for all individuals employed in the formal labor market, by education level and age group. This effect is estimated following Callaway and Sant'Anna (2021), as described in Equation 2, and aggregated as described in Equation 3. Standard errors are clustered at the municipality level. Points are the estimated coefficient and bars are 95% confidence intervals.

observed effects. Therefore, wage changes in this direction would dampen the policy's effects on pre-tertiary educational attainment rather than driving them.

Changes in wages might, however, not fully depict immediate changes in the local labor market caused by new HEIs. If new job opportunities that require a high school degree become available, students might find high school more attractive because it leads to better employment prospects. Because the high school response is immediate, any new job opportunities driving these effects would most likely come from two sources: positions at the new HEIs themselves, or vacancies left by high school graduates who leave the labor market to enroll in college.

7.2 College enrollment

Although new HEIs directly lower the cost of attendance by reducing distance to the nearest HEI, changes in the likelihood of accessing higher education are less straightforward. The national college entrance exam and unified higher education application system allow higher-income students from other municipalities to migrate and take advantage of newly created spots. In addition, spots

at federal tuition-free HEIs are limited, and acceptance rates are below 10%, since the application is nearly costless. To assess whether the policy increases higher education attendance for local public school students, I estimate the effect of receiving an HEI on the likelihood that local students apply to, are accepted to, and enroll in federal HEIs up to three years following the policy. Table 7 shows that local students are 2.5 percentage points more likely to apply (14% increase relative to the pre-policy mean), 1.8 percentage points more likely to be accepted (nearly a 130% increase relative to the pre-policy mean), and 0.7 percentage points more likely to enroll (a 140% increase) in a federal HEI. Enrollment effects are stable over time, as shown in Figure 8, confirming that local students experience a continued increased probability of attending higher education.

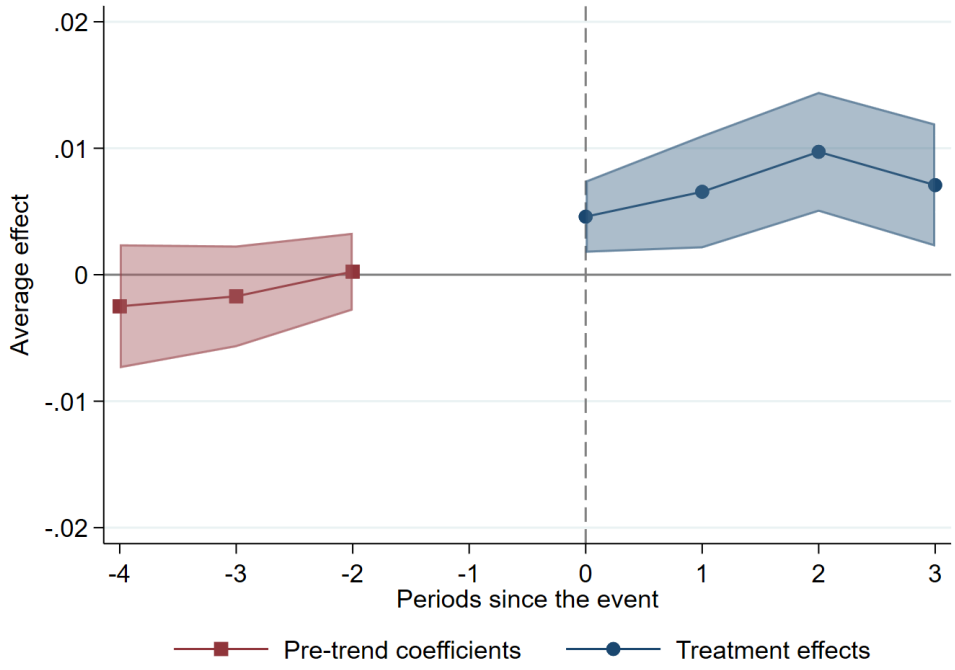
Table 7: College-Going Outcomes

	Applied (1)	Accepted (2)	Enrolled (3)
Treated	0.025***	0.018***	0.007***
	(0.007)	(0.005)	(0.002)
Observations	17,704	17,704	17,704
Pre-policy Mean	0.179	0.014	0.005

Note: Standard errors in parentheses, clustered at the municipality level. All specifications include municipality-level control variables, measured pre-policy and discretized. ATTs are calculated using doubly robust inverse probability weighting (DR-IPW) and aggregated to generate an overall average effect following Callaway and Sant'Anna (2021). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Because this increase happens immediately after HEIs open, it is not driven by the same marginal students behind the observed increases in high school completion. Instead, this response likely comes from students who had already graduated from high school or were in 11th or 12th grade when the policy took effect. In the context of the framework, this result can be interpreted in two ways. First, if 10th-graders deciding whether to stay in school take the college-going of older local public school students as a proxy for their own likelihood of attending an HEI, this result corresponds to an increase in π_c in the framework. Alternatively, if this increase in enrollment frees up jobs that require a high school diploma, it could raise the value of a high school degree. The first would keep students in school by making high school a gateway to college; the second by making high school more valuable as an end. To provide suggestive evidence on which interpretation is more likely, I next examine whether the effects on high school completion vary with the pre-policy college premium.

Figure 8: Likelihood of Enrolling in a Federal HEI



Note: This figure plots event-study estimates following Callaway and Sant'Anna (2021) (Equation 4), using never-treated units as the comparison group. Event time 0 is the year when a new higher education institution opened in the municipality. Points are estimates and bars are 95% confidence bands, with standard errors clustered at the municipality level. Table 7 reports the aggregated average effect.

7.3 The role of the college premium

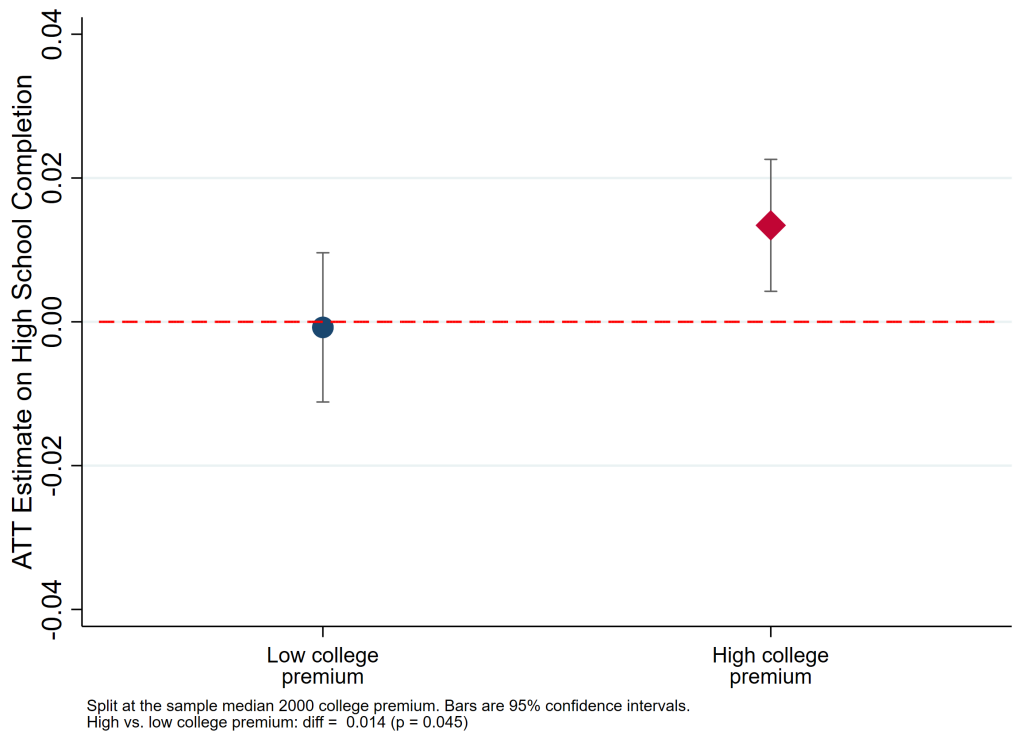
As derived in the framework, increases in high school completion stemming from direct increases in college opportunities (high school as a gateway) should only occur in places where the college premium is sufficiently high. That is, the possibility of completing college must meaningfully increase the expected returns to high school for students for whom high school alone is not worth completing.

This prediction around the college premium provides a direct empirical test. If high school completion responds mainly to changes in the value of high school itself, driven by changes in the local economy, effects should not differ by municipalities' baseline college premium. By contrast, if effects are disproportionately concentrated in places with a high college premium pre-policy, this provides suggestive evidence that they are at least partially driven by students who see high school as a gateway to college.

Using data from the 2000 Brazilian Census, I construct municipality-level pre-policy college

and high school premiums, defined as $\log(\frac{y_m^C}{y_m^{HS}})$ and $\log(\frac{y_m^{HS}}{y_m^{NH}})$, respectively, where y_m^C, y_m^{HS} and y_m^{NH} are municipality-level average earnings of workers with a college degree, a high school degree, and less than a high school degree, respectively. I then classify municipalities as having a high (low) college and high school premium if they are above (below) the median pre-policy premium across municipalities in my sample. Figure 9 plots the average treatment effects of receiving an HEI on high school graduation rates for municipalities with low and high pre-policy college premiums. High school completion increases only in municipalities with a high college premium. This effect is positive, significant, and statistically different from the effect in municipalities with below-median college premiums (p-value = 0.045).

Figure 9: High School Completion effect, by pre-policy college premium



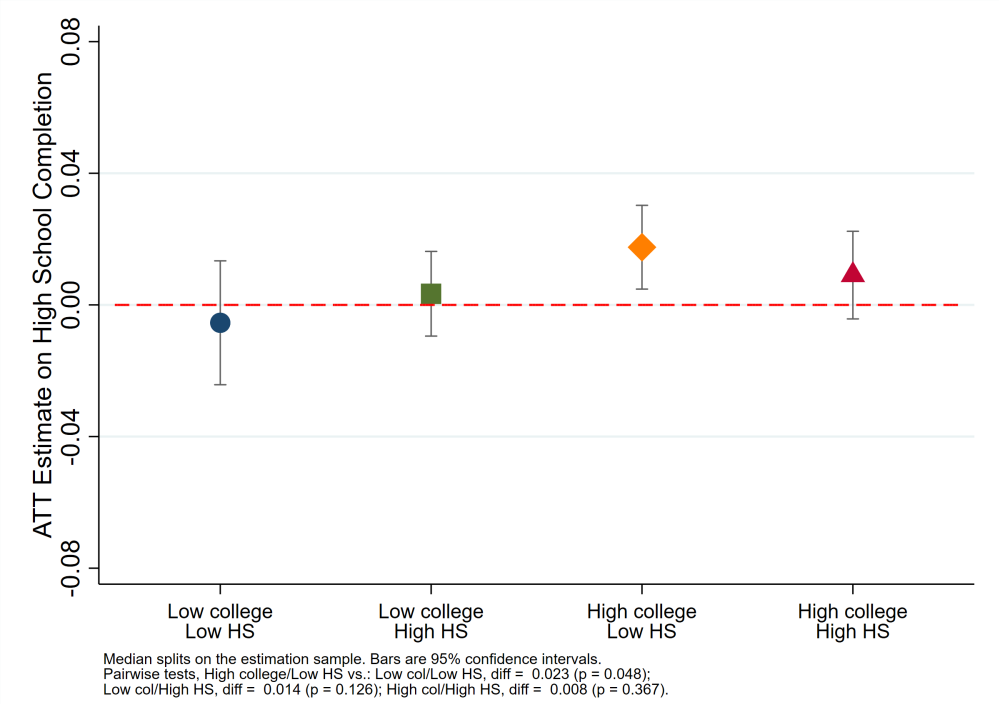
Note: This figure plots the effect of new higher education institutions on high school completion (Table 3, column 7) separately by the municipality's pre-policy college premium. Municipalities with a low (high) college premium are those below (above) the median pre-policy college premium across municipalities in my sample. Points are estimates and bars are 95% confidence intervals, with standard errors clustered at the municipality level. The difference between the two estimates is statistically significant at the 5% level (p-value = 0.045).

Figure A2 in the Appendix plots the differential effects by pre-policy college premium for all of my main outcomes. Consistent with the pattern observed for high school completion, decreases in dropout at 11th and 12th grade are larger in municipalities with a high college premium, although

the differences are not statistically significant. Importantly, effects on local wages and on college enrollment (which could free up jobs for high school graduates) are similar across these two settings. The greater response in high school completion in places with a high college premium therefore provides suggestive evidence that students are, at least partially, responding to expanded college opportunities and treating high school as a gateway to college.

Because a low high school premium implies that more students would choose to drop out pre-policy, I further divide municipalities into four groups, based on the baseline college and high school premiums. Figure 10 plots the average treatment effects of receiving an HEI on high school graduation rates for each of these groups. Within municipalities with a high college premium, point estimates are largest for those with a low high school premium, as predicted by the framework, although the differences are not statistically significant.

Figure 10: High School Completion effect, by pre-policy distribution of college and high school premium



Note: This figure plots the effect of new higher education institutions on high school completion (Table 3, column 7) separately for four groups of municipalities, defined by their pre-policy college and high school premiums. Municipalities with a low (high) college or high school premium are those below (above) the median of the corresponding pre-policy premium across municipalities in my sample. Points are estimates and bars are 95% confidence intervals, with standard errors clustered at the municipality level. Differences between groups are not statistically significant at conventional levels.

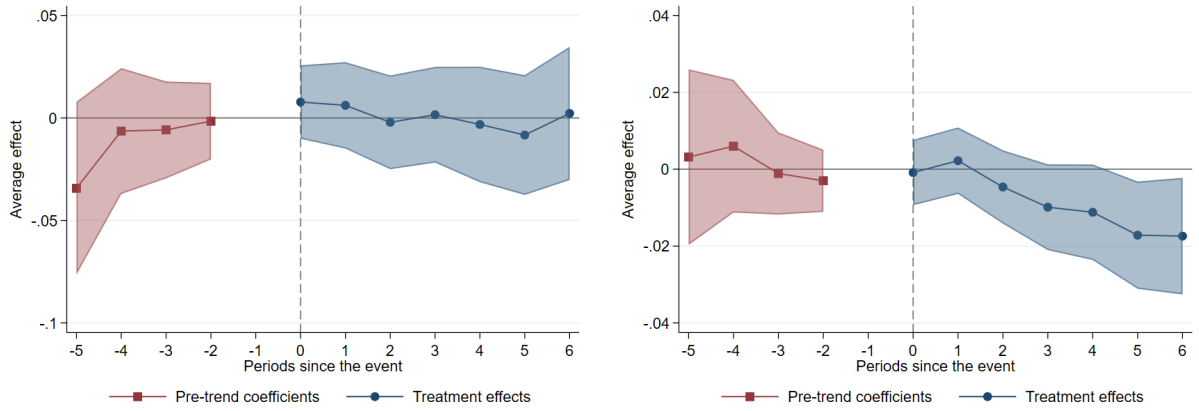
These results suggest that the overall effects shown in Table 3 mask important heterogeneity

in who responds to an increased supply of HEIs by changing pre-tertiary schooling decisions, and provide suggestive evidence for the role of expanded college access in raising demand for secondary schooling.

8 How sustainable is the policy effect in the long run?

Since responses are concentrated where the pre-policy college premium is high, the long-run effects of HEI expansion on high school attainment may depend on whether that premium persists. As students respond to this policy by increasing high school and college attainment, returns to both levels of education might decrease, diminishing the impact of higher education opportunities on high school completion over time. In this section, I provide suggestive evidence on how municipality-level college and high school premiums (in the formal labor market) behave following the opening of an HEI. This will inform on whether we can expect higher education expansion to permanently increase the returns to high school, or whether these short-run effects should dissipate over time.

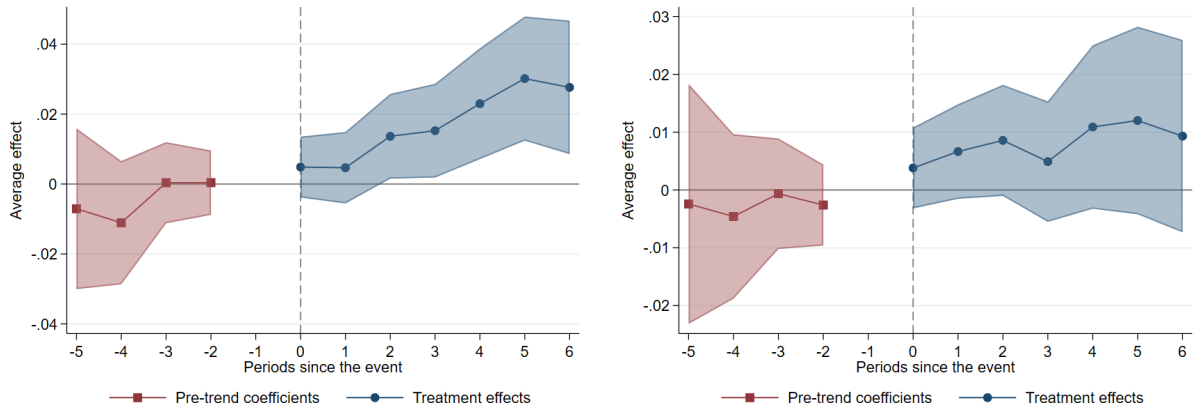
Figure 11: Municipality college-premium (left) and high-school premium (right) in the formal labor market



Note: This figure plots event-study estimates following Callaway and Sant'Anna (2021) (Equation 4), using never-treated units as the comparison group. College and high school premiums are calculated for the formal labor market, and defined as $\log(\frac{y_m^C}{y_m^{HS}})$ and $\log(\frac{y_m^{HS}}{y_m^{NH}})$, respectively, where y_m^C, y_m^{HS} and y_m^{NH} are municipality-level average earnings of workers with a college degree, a high school degree, and less than a high school degree, respectively. Event time 0 is the year when a new higher education institution opened in the municipality. Points are estimates and bars are 95% confidence bands, with standard errors clustered at the municipality level.

Figure 11 shows that local college premiums remain stable for up to 6 years after an HEI's opening. On the other hand, high-school premiums begin to decrease sharply two years after the policy. This decrease in high school returns comes from an increase in primary-education wages,

Figure 12: Primary-educated (left) and high school graduate wages (right) in the formal labor market



Note: This figure plots event-study estimates following Callaway and Sant'Anna (2021) (Equation 4), using never-treated units as the comparison group. Wages for primary-educated and high school-educated workers are calculated using the universe of formal labor market employees. Event time 0 is the year when a new higher education institution opened in the municipality. Points are estimates and bars are 95% confidence bands, with standard errors clustered at the municipality level.

rather than a decrease in earnings for high school graduates, as shown in Figure 12. This likely reflects both a reduced supply of primary-educated young workers due to increased educational attainment following the opening of an HEI and the potential creation of new jobs that might increase demand for low-skilled labor. Overall, college premiums seem to remain stable 6 years after the policy, suggesting that the higher-education-induced increase in returns to high school can be sustained over time, provided that college returns remain high relative to the outside option (in this case, primary education).

9 Robustness analysis

Changes in composition. One possible threat to identification is that, due to a potential influx of people into the municipality caused by the arrival of an HEI, such as professors, staff, and new students, the composition of public K-12 schools could change in ways that affect high school graduation rates and test scores. The observed results would then not be attributable to the expansion of higher education itself, but to compositional changes generated by the policy. To assess whether this is the case, I replace the outcome variable in my estimation with the shares of white students, students who attended only public schools throughout their lives, and students whose

mothers completed high school. This measures whether the policy affects student composition[11].

As shown in Table A7 in the Appendix, student composition does not appear to change as a result of the policy. In fact, all significant estimates would bias the results downward, given that non-white and public school students, on average, have lower high school completion rates and lower test scores than white and private school students.

Control variables. Results are robust to alternative sets and combinations of controls, including the number of families receiving Bolsa Família per capita. When controls are omitted entirely, estimates for dropout rates remain similar and are even larger, but the effect on high school completion is smaller (0.005 compared with 0.008 in the main specification) and no longer statistically significant. Because treated municipalities were selected based on observable characteristics that may also be associated with trends in educational outcomes, the identifying assumption in my main specification is parallel trends conditional on these pre-policy characteristics. The smaller unconditional estimate is consistent with this: without controls, treated municipalities are compared with control municipalities that differ systematically in characteristics such as income, population, and education spending. Estimates without controls are shown in Appendix Tables A4 and A5.

Permutation test. To test whether my results could arise by chance, I conduct a randomization-inference test for the effect on eventually graduating from high school (10th grade baseline). I randomly reassign the observed treatment cohorts, including never-treated status, across municipalities 2,000 times, holding fixed the number of municipalities in each cohort. In each draw, I re-estimate the ATT with the same specification as in the main results. Because the reassigned cohorts are unrelated to where HEIs actually opened, these placebo estimates show the distribution of effects expected in the absence of any true policy effect. As expected, the placebo estimates are centered at zero (mean -0.0002). The observed t-statistic exceeds 98.7% of the placebo t-statistics in absolute value. It is therefore unlikely that my main results are driven by random noise. Figure A1 in the Appendix shows the distribution of placebo estimates.

Technical high schools. Although trade schools are part of the Brazilian higher education system, they sometimes offer technical training to students enrolled in high school. To assess whether these high school programs drive the increase in high school completion, I estimate my

[11]Public school attendance in Brazil is highly correlated with income. The questionnaire asks students if, starting in first grade, they studied in public only, private only, or both private and public schools.

main results separately for municipalities with and without technical training in high school. As shown in Table A6 in the Appendix, the effects are not statistically different across groups. This suggests that the increase in high school completion is not driven by changes in the supply of technical high school programs.

10 Discussion

Although future educational opportunities can meaningfully shape early educational attainment, this channel has only recently been explored in the literature and remains largely absent from the policy debate as a tool for sustainably increasing overall demand for schooling. By bringing higher education within reach of low-income students, expansion policies can increase pre-tertiary educational attainment and reduce educational disparities. Exploring a large-scale expansion of higher education institutions in Brazil, I find that local students exposed to a new HEI increase persistence in high school by 1 percentage point (5-6% relative to the baseline mean) and are 1.5% more likely to graduate high school – increasing the stock of high school graduates by over 20,000 in the first three exposed cohorts. More disadvantaged, non-white public students benefit the most: they are 3% more likely to start high school after an HEI opens and, among those who reach 10th grade, 4% more likely to graduate, closing one-third of the completion gap relative to white students.

Trade schools and universities both increase high school completion, but appear to affect students at different stages of their educational careers. Ninth graders in municipalities that received a university were nearly 2 percentage points more likely to transition to high school, while those in municipalities that received a trade school were not. Instead, trade schools shift schooling decisions among older pre-tertiary students who would have started high school in the absence of the policy, but are now more likely to stay in school and graduate. This result is particularly relevant for LMICs, since trade schools are smaller and less costly than four-year universities, and potentially better aligned with the needs of local labor markets.

Combining the framework's predictions and empirical evidence, I find that my effects stem at least partially from increased educational opportunities rather than from changes in the local labor market. High school graduate wages do not change following the opening of an HEI. By contrast, students are 14% more likely to apply to an HEI and 0.7 percentage points more likely to enroll, a 140% increase from the pre-policy mean. Most importantly, high school graduation increases only in municipalities with a high pre-policy college premium, even though college enrollment rises similarly across high- and low-premium municipalities. This pattern suggests that students are treating high school as a gateway to college rather than as an end goal.

Given that the college premium seems to play an important role in mediating these effects, I examine whether these results can be sustained over time by analyzing how premiums in the

formal labor market evolve up to six years after an HEI opens. While the college premium remains stable over time, the high school premium declines sharply as wages for primary-educated workers rise. This is consistent with the general equilibrium effects of expanding education documented by Khanna (2023), who finds that a large expansion of primary schooling in India raised wages for unskilled workers. As long as the returns to college remain high relative to dropping out, higher education opportunities are likely to continue shaping high school decisions in the long run.

My empirical results contrast with Bedard (2001). While she finds that expanding higher education in the United States increased high school dropout rates, I find that students exposed to new HEIs in Brazil responded by increasing their high school persistence and completion rates. This difference likely reflects differences in context, and both results are consistent with my framework. First, average years of schooling in the United States in the 1970s were approximately 12 years, equivalent to high school completion, whereas average attainment in my setting is below 9 years, equivalent to primary school completion in Brazil. For an expansion in higher education to affect high school completion, enough students must be dropping out of high school in the absence of college opportunities. This is true in my setting, but much less so in the United States. Second, in the signaling model of Bedard (2001), in the absence of HEIs, both low-type and high-type students graduate from high school in a pooling equilibrium. However, after an HEI opens, only high-type students attend college, while low-type students can, at most, complete high school. This leads to a separating equilibrium, allowing employers to correctly identify students' types and pay low wages to low types (high school dropouts or graduates) and high wages to high types (college graduates). In my framework, this is equivalent to the policy lowering wages for high school graduates, which in turn reduces high school completion and increases dropout, as found by Bedard (2001).

In a context more closely related to this paper, I also build on previous work by Jagnani and Khanna (2020) and provide further external validity for their findings. Importantly, I find significant gains in persistence at lower levels of education following higher-education expansion in a setting where most new HEIs are newly created, not prestigious, and predominantly technically focused (trade schools). I show that not only four-year colleges, but also smaller two-year trade schools can have important trickle-down effects on high school completion. By developing a theoretical framework and conducting an empirical analysis of the mechanisms, I explore the channels through which these effects may operate and highlight the key role of skill premium and increased college access opportunities in shaping this relationship.

Overall, expanding higher education in Brazil shifted earlier educational decisions among pre-

tertiary students in receiving municipalities, raising students' high school persistence and graduation rates. Local students were also more likely to attend federal higher education institutions. This policy provides a relevant example of how the availability of future educational opportunities, coupled with an effective increase in access for low-income students, can raise the expected returns to basic education and generate meaningful improvements in human capital accumulation, particularly among disadvantaged students and in places where returns to high school alone are not enough to keep students in school.

References

Akhtari, M., Bau, N., and Laliberte, J. (2023). Affirmative action and pre-college human capital. *American Economic Journal: Applied Economics*.

Angrist, J., Bettinger, E., and Kremer, M. (2006). Long-term educational consequences of secondary school vouchers: Evidence from administrative records in colombia. *American Economic Review*, 96(3):847–862.

Arroyo, A. F. and Petterini, F. C. (2020). A influência da orientação política dos prefeitos na probabilidade de implementação de institutos federais e campi universitários em tempos de reuni. *Textos De Economia*, 23(2):1–31.

Baez, J. E. and Camacho, A. (2011). Assessing the long-term effects of conditional cash transfers on human capital: Evidence from Colombia. Policy Research Working Paper 5681, World Bank, Washington, DC.

Barbosa, M. P., Petterini, F. C., and Ferreira, R. T. (2020). Política de expansão das universidades federais: É possível potencializar os impactos econômicos? *Revista De Administração Contemporânea*, 24(1):3–24.

Barrera-Osorio, F., Bertrand, M., Linden, L. L., and Perez-Calle, F. (2011). Improving the design of conditional transfer programs: Evidence from a randomized education experiment in Colombia. *American Economic Journal: Applied Economics*, 3(2):167–195.

Bedard, K. (2001). Human capital versus signaling models: University access and high school dropouts. *Journal of Political Economy*, 109(4):749–775.

Burde, D. and Linden, L. L. (2013). Bringing education to Afghan girls: A randomized controlled trial of village-based schools. *American Economic Journal: Applied Economics*, 5(3):27–40.

Callaway, B. and Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Card, D. (1993). Using geographic variation in college proximity to estimate the return to schooling. *NBER Working Paper*.

Cox, A. (2024). From classroom to prosperity: Fostering development through higher education. Working paper, Yale University.

Currie, J. and Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*, 118(4):1495–1532.

Dinkelman, T. and Martínez, C. (2014). Investing in schooling in Chile: The role of information about financial aid for higher education. *Review of Economics and Statistics*, 96(2):244–257.

Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91(4):795–813.

Duflo, E., Dupas, P., and Kremer, M. (2021). The impact of free secondary education: Experimental evidence from Ghana. Working Paper 28937, National Bureau of Economic Research.

Erbabian, M. (2026). Degrees of structural transformation: How college expansion reshapes local skill composition across the schooling ladder. Working paper, University of California, San Diego.

Faveri, D. B., Petterini, F. C., and Barbosa, M. P. (2018). Uma avaliação do impacto da política de expansão dos institutos federais nas economias dos municípios brasileiros. *Planejamento e Políticas Públicas*, 50:125–147. Available at https://www.ipea.gov.br/ppp/index.php/PPP/article/view/742.

Feinmann, J. and Hsu Rocha, R. (2026). Inequality and higher education: The role of elite public colleges. Working paper, Paris School of Economics and Nova SBE.

Gethin, A. (2025). Distributional growth accounting: Education and the reduction of global poverty, 1980–2019. *Quarterly Journal of Economics*.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Jagnani, M. and Khanna, G. (2020). The effects of elite public colleges on primary and secondary schooling markets in india. *Journal of Development Economics*, 146:102512.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *Quarterly Journal of Economics*, 125(2):515–548.

Katzkowicz, N., Lavy, V., Querejeta, M., and Rosá, T. (2023). Schooling and intergenerational mobility: Consequences of expanding higher education institutions. Working Paper 31906, National Bureau of Economic Research.

Khanna, G. (2020). Does affirmative action incentivize schooling? evidence from india. *The Review of Economics and Statistics*, 102(2):219–233.

Khanna, G. (2023). Large-scale education reform in general equilibrium: Regression discontinuity evidence from india. *Journal of Political Economy*, 131(2):549–591.

Lucas, A. M. and Mbiti, I. M. (2012). Access, sorting, and achievement: The short-run effects of free primary education in Kenya. *American Economic Journal: Applied Economics*, 4(4):226–253.

Ministério da Educação [MEC] (2011). Expansão da educação superior e profissional e tecnológica: Mais formação e oportunidades para os brasileiros. Government publication. Brazilian Ministry of Education.

OECD (2026). *Education at a Glance 2026: OECD Indicators*. OECD Publishing, Paris.

Oreopoulos, P. and Petronijevic, U. (2013). Making college worth it: A review of research on the returns to higher education. *NBER Working Paper*, (19053).

Ozier, O. (2018). The impact of secondary schooling in kenya: A regression discontinuity analysis. *Journal of Human Resources*, 53(1):157–188.

Rands, T. and Barsanetti, B. (2025). College access spillover on high school performance: Evidence from college openings. *Economics of Education Review*, 109.

Sandholtz, W. A. (2024). Secondary school access raises primary school achievement. Working paper.

Sant'Anna, P. H. C. and Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1):101–122.

Schultz, T. P. (2004). School subsidies for the poor: Evaluating the Mexican Progresa poverty program. *Journal of Development Economics*, 74(1):199–250.

UNESCO (2025). The right to education: Past, present and future directions. Technical report, United Nations Educational, Scientific and Cultural Organization. Accessed: May 23, 2026.

UNESCO (2026). Global education monitoring report 2026: Access and equity, countdown to 2030. Technical report, UNESCO, Paris. Available at https://www.unesco.org/gem-report/en/publication/equity-and-access.

Vu, K. (2024). Higher education expansion and the rise of the skill-intensive service sector. Job Market Paper, University of Minnesota.

Zhang, R. (2026). Higher education expansion and the rural–urban gap in secondary schooling and labor outcomes: Evidence from China's 1999 reform. *Journal of Development Economics*, 181:103721.

A Additional Figures and Tables

Table A1: Graduation Outcomes Adjusted by 12th-grade Dropout Rates

	9th grade baseline		10th grade baseline	
	Graduate on time	Eventually graduate	Graduate on time	Eventually graduate
	(1)	(2)	(3)	(4)
Treated	0.001	0.011**	0.000	0.010***
	(0.004)	(0.004)	(0.004)	(0.004)
Observations	19,208	14,410	21,616	16,814
Pre-policy Mean	0.40	0.49	0.46	0.54

Note: Standard errors in parentheses, clustered at the municipality level. All specifications include municipality-level control variables, measured pre-policy and discretized. Outcomes are the adjusted versions of the persistence measures. ATTs are calculated using doubly robust inverse probability weighting (DR-IPW) and aggregated to generate an overall average effect following Callaway and Sant'Anna (2021). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A2: Graduation Outcomes by Baseline Grade

	9th grade baseline		10th grade baseline	
	Graduate on time	Eventually graduate	Graduate on time	Eventually graduate
	(1)	(2)	(3)	(4)
Treated	0.000	0.011**	-0.003	0.008**
	(0.004)	(0.004)	(0.003)	(0.003)
Observations	19,208	14,410	21,616	16,814
Pre-policy Mean	0.44	0.54	0.46	0.54

Note: Standard errors in parentheses, clustered at the municipality level. All specifications include municipality-level control variables, measured pre-policy and discretized. ATTs are calculated using doubly robust inverse probability weighting (DR-IPW) and aggregated to generate an overall average effect following Callaway and Sant'Anna (2021). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A3: High School Completion by Baseline Grade, by Type of HEI Received

	9th grade baseline		10th grade baseline	
	Graduate on time	Eventually graduate	Graduate on time	Eventually graduate
	(1)	(2)	(3)	(4)
Received a Trade School	-0.002	0.010**	-0.004	0.010**
	(0.005)	(0.005)	(0.004)	(0.004)
Observations	18,616	13,966	20,950	16,296
Pre-policy Mean	0.45	0.54	0.47	0.55
Received a University	0.011	0.021*	-0.007	-0.003
	(0.010)	(0.012)	(0.008)	(0.010)
Observations	16,544	12,412	18,619	14,483
Pre-policy Mean	0.44	0.54	0.44	0.53
P-value (diff)	0.27	0.42	0.74	0.21
Bootstrap p-value (diff)	0.26	0.40	0.77	0.23

Note: Standard errors in parentheses, clustered at the municipality level. Each type of HEI is compared with the same never-treated municipalities. Graduate on time is the share of students in the baseline-grade cohort who complete high school on time; Eventually graduate allows for up to a two-year delay. All specifications include municipality-level control variables, measured pre-policy and discretized. ATTs are calculated using doubly robust inverse probability weighting (DR-IPW) and aggregated to generate an overall average effect following Callaway and Sant'Anna (2021). P-value (diff) tests the equality of the trade school and university estimates using $\sqrt{se_{TS}^2 + se_U^2}$, which ignores that both groups share the same comparison municipalities. Bootstrap p-value (diff) uses the standard error of the difference from a municipality-level bootstrap (200 replications, stratified by opening cohort), re-estimating both ATTs in each draw. Observations vary between outcomes due to years of data availability. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A4: Persistence in Secondary School (No Controls)

		Dropout rate					
	Start HS	9th grade	10th grade	11th grade	12th grade	Graduate on time	Eventually graduate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated	0.003	-0.110	-1.340***	-1.068***	-0.925***	-0.003	0.005
	(0.003)	(0.137)	(0.286)	(0.236)	(0.202)	(0.003)	(0.003)
Observations	19,208	35,241	35,241	35,241	35,241	21,616	16,814
Pre-policy Mean	0.79	7.57	18.21	14.31	10.62	0.46	0.54

Note: Standard errors in parentheses, clustered at the municipality level. No municipality-level control variables are included: ATTs are estimated under unconditional parallel trends and aggregated to generate an overall average effect following Callaway and Sant'Anna (2021). Observations vary between outcomes due to years of data availability. Dropout estimates (columns 2–5) use a common sample of municipality-years with non-missing dropout rates for all four grades. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Performance Outcomes (No Controls)

	9th Grade Math	9th Grade Portuguese	College Exam Scores	Share Taking Exam
	(1)	(2)	(3)	(4)
Treated	0.01	0.01	-0.51	0.03***
	(0.01)	(0.01)	(1.26)	(0.00)
Observations	17,548	17,548	16,485	16,485
Pre-policy Mean	-0.17	-0.39	525.48	0.10

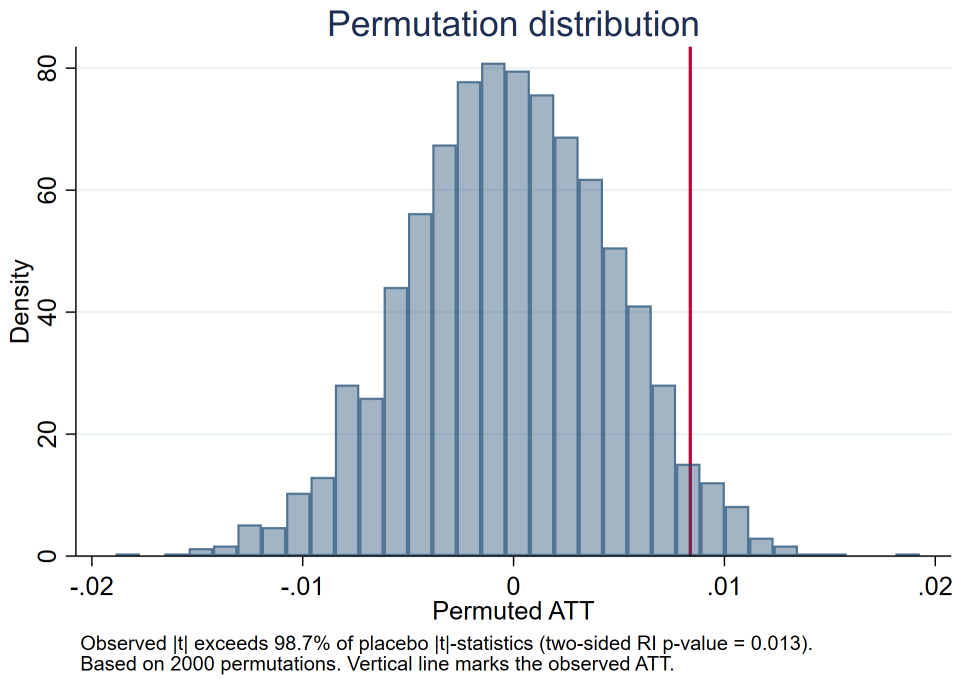
Note: Standard errors in parentheses, clustered at the municipality level. No municipality-level control variables are included: ATTs are estimated under unconditional parallel trends and aggregated to generate an overall average effect following Callaway and Sant'Anna (2021). Observations vary between outcomes due to years of data availability. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A6: Persistence in Secondary School in Trade School Municipalities, by Federal Technical High School Supply

	Dropout rate				
	9th grade	10th grade	11th grade	12th grade	Eventually graduate
	(1)	(2)	(3)	(4)	(5)
All trade school municipalities	0.038	-1.239***	-0.760***	-0.584**	0.010**
	(0.159)	(0.333)	(0.277)	(0.238)	(0.004)
Observations	34,149	34,149	34,149	34,149	16,296
Pre-policy Mean	7.15	17.70	13.97	10.43	0.55
With federal technical HS	0.162	-1.169***	-0.821***	-0.491*	0.011***
	(0.185)	(0.375)	(0.313)	(0.265)	(0.004)
Observations	33,053	33,053	33,053	33,053	15,771
Pre-policy Mean	6.96	17.34	13.63	10.35	0.55
Without federal technical HS	-0.334	-1.377**	-0.518	-0.863*	-0.001
	(0.270)	(0.699)	(0.559)	(0.511)	(0.014)
Observations	30,774	30,774	30,774	30,774	14,686
Pre-policy Mean	7.74	18.79	14.99	10.67	0.54
P-value (diff)	0.13	0.79	0.64	0.52	0.42
Bootstrap p-value (diff)	0.13	0.81	0.66	0.52	0.48

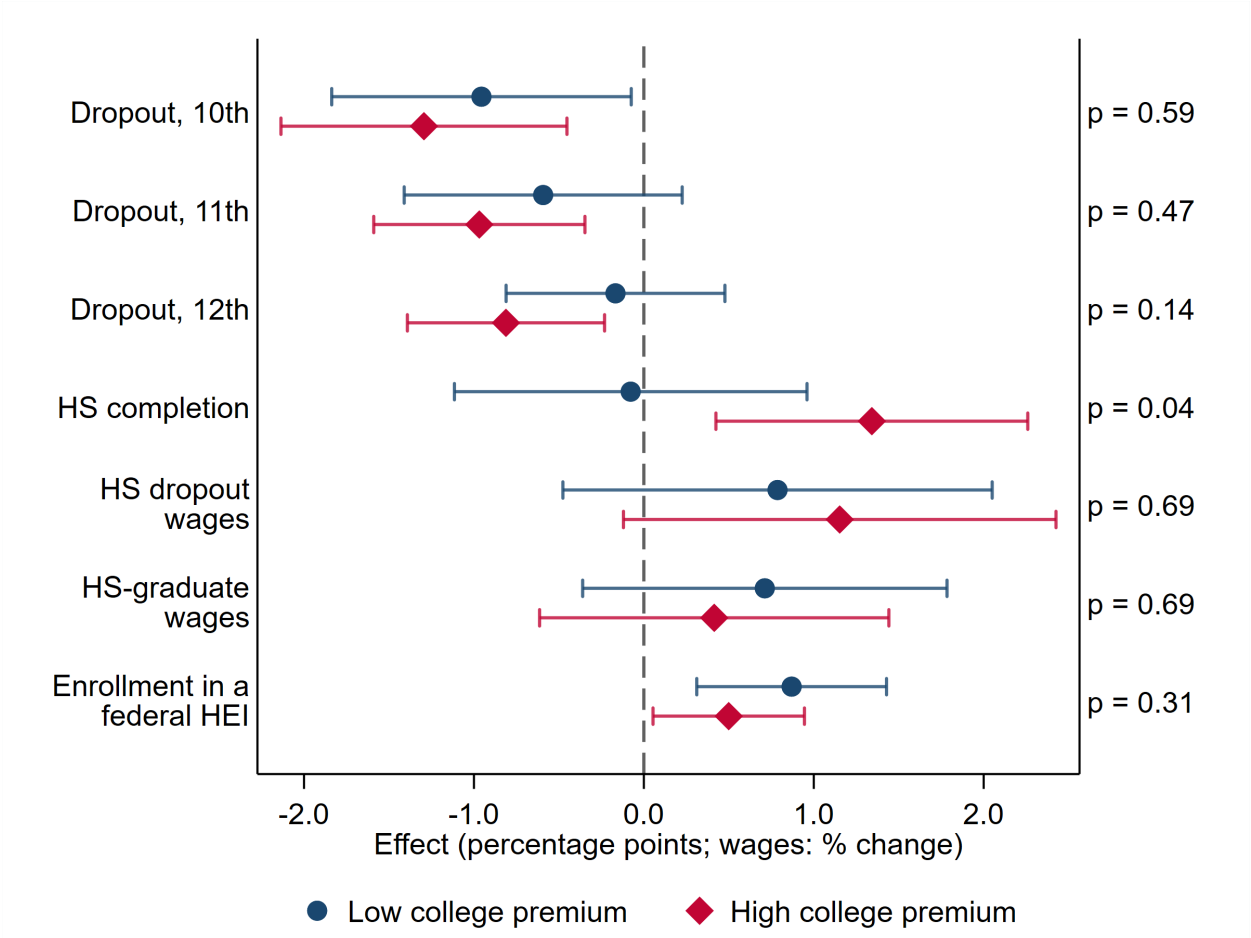
Note: Standard errors in parentheses, clustered at the municipality level. The sample includes municipalities that received only a trade school and never-treated municipalities. The first row pools all trade school municipalities. The next rows split them by whether, in 2014, the municipality had a federal technical education campus (Rede Federal de EPCT) offering integrated technical high school or concurrent technical training; both groups are compared with the same never-treated municipalities. All specifications include municipality-level control variables, measured pre-policy and discretized. ATTs are calculated using doubly robust inverse probability weighting (DR-IPW) and aggregated to generate an overall average effect following Callaway and Sant'Anna (2021). Dropout estimates (columns 1–4) use a common sample of municipality-years with non-missing dropout rates for all four grades. P-value (diff) tests the equality of the two estimates using $\sqrt{se_1^2 + se_2^2}$, which ignores that both groups share the same control municipalities. Bootstrap p-value (diff) uses the standard error of the difference from a municipality-level bootstrap (200 replications, stratified by opening cohort and group), re-estimating both ATTs in each draw. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A1: Permutation Test: Distribution of Placebo Estimates



Note: Distribution of ATT estimates from 2,000 random reassignments of treatment cohorts (including never-treated status) across municipalities, holding fixed the number of municipalities in each cohort. The vertical line marks the observed estimate. Outcome: eventually graduating from high school, 10th grade baseline.

Figure A2: Heterogeneity by pre-policy college premium



Point estimates and 95% CIs. HS completion and enrollment x 100 (pp);
wages x 100 (approx. % change). Split at the sample median 2000 college premium
Right: p-value of the high - low difference.

Table A7: Effects of Higher Education Expansion on Student Composition

Panel A: *5th Grade*	White	Studied in Public Schools Only	Mother Completed High School
	(1)	(2)	(3)
Treated	-0.008*	-0.012	0.000
	(0.004)	(0.007)	(0.004)
Observations	14,470	14,471	14,471
Pre-policy Mean	0.35	0.57	0.29
Panel B: *9th Grade*	White	Studied in Public Schools Only	Mother Completed High School
	(1)	(2)	(3)
Treated	0.002	0.020***	-0.001
	(0.003)	(0.005)	(0.004)
Observations	15,093	15,093	15,092
Pre-policy Mean	0.35	0.90	0.31

Note: Standard errors in parentheses, clustered at the municipality level. All specifications include municipality-level control variables, measured pre-policy and discretized. ATTs are calculated using doubly robust inverse probability weighting (DR-IPW) and aggregated to generate an overall average effect following Callaway and Sant'Anna (2021). Observations vary between outcomes due to years of data availability. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

B Complete overview of the model

B.1 Setup

At the beginning of 10th grade, the first grade of high school in Brazil, a student enrolled in high school chooses a schooling path $(s, d) \in \{(0,0), (1,0), (1,1)\}$, where $s = 1$ denotes completing high school, $s = 0$ denotes dropping out of high school before completion, and $d = 1$ denotes attempting college. Because college requires first graduating high school, the pair $(0, 1)$ is not feasible. The student maximizes expected utility:

$$\max_{(s,d)\in\{(0,0),(1,0),(1,1)\}} U(C_1) + \beta_2 \,\mathbb{E}\big[U(C_2)\big] + \beta_3 \,\mathbb{E}\big[U(C_3)\big], \tag{B1}$$

where $U(c) = \ln c$ and the expectation is over the college admission lottery. For paths $(0, 0)$ and $(1, 0)$ there is no uncertainty, so the expectation is trivial. For path $(1, 1)$, the student is admitted with probability $\pi_c \in (0, 1)$; conditional on admission they enroll and pay costs p^c, and conditional on non-admission they earn high-school wages. When choosing s and d, a student knows their cost

of high school attendance, p^H, college cost, p^c, and the likelihood of being accepted to an HEI, π_c, but does not know the college admission outcome.

Period 1. Period 1 lasts 3 years (the duration of high school in Brazil, grades 10–12):

$$C_1 = \bar{k}_1 + (1-s)\, y_1^{NH} - s\, p^H,$$

where $\bar{k}_1 = 3\bar{k}$ and $y_1^{NH} = 3y^{NH}$.

Period 2. Period 2 lasts 4 years (the typical higher education duration), with $\bar{k}_2 = 4\bar{k}$:

$$C_2 = \begin{cases} \bar{k}_2 + y_2^{NH} & \text{if } (s,d) = (0,0), \\ \bar{k}_2 + y_2^{HS} & \text{if } (s,d) = (1,0), \\ \bar{k}_2 + y_2^{HS} & \text{if } (s,d) = (1,1) \text{ and not admitted (probability } 1-\pi_c), \\ \bar{k}_2 - p^c & \text{if } (s,d) = (1,1) \text{ and admitted (probability } \pi_c), \end{cases}$$

where $y_2^{NH} = 4y^{NH}$ and $y_2^{HS} = 4y^{HS}$.

Period 3. Period 3 represents the remainder of working life. C_3 denotes *per-year* consumption in that period, which is assumed to be constant:

$$C_3 = \begin{cases} \bar{k} + y^{NH} & \text{if } (s,d) = (0,0), \\ \bar{k} + y^{HS} & \text{if } (s,d) = (1,0), \\ \bar{k} + y^{HS} & \text{if } (s,d) = (1,1) \text{ and not admitted}, \\ \bar{k} + y^{C} & \text{if } (s,d) = (1,1) \text{ and admitted}. \end{cases}$$

Timing and discount factors. Period 1 lasts 3 years, period 2 lasts 4 years, and period 3 runs from year 8 to the end of working life, year N. Accordingly,

$$\beta_2 = \beta^3, \qquad \beta_3 \;\equiv\; \sum_{t=8}^{N} \beta^{t-1} \;=\; \beta^7\, \frac{1-\beta^{N-7}}{1-\beta} \;>\; 0. \tag{B2}$$

Because C_3 is measured per year, β_3 is the exact present value, as of period 1, of one unit of per-year consumption received in every year $t = 8, \ldots, N$.

Wages. y^{NH} is the wage for a worker who did not complete high school, y^{HS} for a high-school graduate, and y^C for a college graduate. I assume all are constant over time. An individual's education level is fixed after the schooling decision is made and the college admission lottery is realized.

Endowment and costs. $\bar{k}$ is the household endowment per year. $p^H > 0$ is the direct cost of remaining enrolled in and completing high school and $p^c > 0$ is the direct cost of attending college. Both p^H and p^c do not include tuition costs, and are mostly comprised of transportation or relocation costs. Throughout, I assume $p^H < \bar{k}_1$ and $p^c < \bar{k}_2$ so that period-1 and period-2 consumption are strictly positive on all paths.

College access. $\pi_c \in (0,1)$ is the probability of admission to a higher-education institution. The student knows π_c when choosing (s,d), but the admission outcome is only realized after the choice.

Cost distributions. Costs are drawn independently and uniformly: $p^H \sim U[0, \bar{P}^H]$ and $p^c \sim U[0, \bar{P}^c]$.

B.2 Value Functions

Substituting $U(c) = \ln c$ and taking the expectation over the admission lottery for path $(1,1)$:

$$V(0,0) = \ln(\bar{k}_1 + y_1^{NH}) + \beta_2 \ln(\bar{k}_2 + y_2^{NH}) + \beta_3 \ln(\bar{k} + y^{NH}), \tag{B3}$$

$$V(1,0) = \ln(\bar{k}_1 - p^H) + \beta_2 \ln(\bar{k}_2 + y_2^{HS}) + \beta_3 \ln(\bar{k} + y^{HS}), \tag{B4}$$

$$\begin{aligned} V(1,1) = \ln(\bar{k}_1 - p^H) &+ \beta_2 \big[(1-\pi_c)\ln(\bar{k}_2 + y_2^{HS}) + \pi_c \ln(\bar{k}_2 - p^c)\big] \\ &+ \beta_3 \big[(1-\pi_c)\ln(\bar{k} + y^{HS}) + \pi_c \ln(\bar{k} + y^C)\big]. \end{aligned} \tag{B5}$$

B.3 Decision Thresholds and Decision Rule

B.3.1 Threshold 1: $(1,0)$ vs. $(0,0)$

Derivation. A student chooses $(1,0)$ over $(0,0)$ when $V(1,0) \geq V(0,0)$. Setting $V(1,0) = V(0,0)$ using (B4) and (B3):

$$\ln(\bar{k}_1 - p^H) + \beta_2 \ln(\bar{k}_2 + y_2^{HS}) + \beta_3 \ln(\bar{k} + y^{HS}) = \ln(\bar{k}_1 + y_1^{NH}) + \beta_2 \ln(\bar{k}_2 + y_2^{NH}) + \beta_3 \ln(\bar{k} + y^{NH}).$$

Isolating p^H on one side, I can find a threshold for which a student will be indifferent between $(1,0)$ and $(0,0)$:

$$\begin{aligned}
\ln(\bar{k}_1 - p^H) - \ln(\bar{k}_1 + y_1^{NH}) &= \beta_2\big[\ln(\bar{k}_2 + y_2^{NH}) - \ln(\bar{k}_2 + y_2^{HS})\big] + \beta_3\big[\ln(\bar{k} + y^{NH}) - \ln(\bar{k} + y^{HS})\big] \\
\implies \quad \ln\left(\frac{\bar{k}_1 - p^H}{\bar{k}_1 + y_1^{NH}}\right) &= \beta_2 \ln\left(\frac{\bar{k}_2 + y_2^{NH}}{\bar{k}_2 + y_2^{HS}}\right) + \beta_3 \ln\left(\frac{\bar{k} + y^{NH}}{\bar{k} + y^{HS}}\right) \\
\implies \quad \ln\left(\frac{\bar{k}_1 - p^H}{\bar{k}_1 + y_1^{NH}}\right) &= \ln\left[\left(\frac{\bar{k}_2 + y_2^{NH}}{\bar{k}_2 + y_2^{HS}}\right)^{\beta_2}\left(\frac{\bar{k} + y^{NH}}{\bar{k} + y^{HS}}\right)^{\beta_3}\right] \\
\implies \quad \bar{p}_{10}^H &= \bar{k}_1 - (\bar{k}_1 + y_1^{NH})\left(\frac{\bar{k}_2 + y_2^{NH}}{\bar{k}_2 + y_2^{HS}}\right)^{\beta_2}\left(\frac{\bar{k} + y^{NH}}{\bar{k} + y^{HS}}\right)^{\beta_3}.
\end{aligned}$$

Since $V(1,0) - V(0,0)$ is strictly decreasing in p^H, a student chooses $(1,0)$ over $(0,0)$ if and only if $p^H \leq \bar{p}_{10}^H$.

B.3.2 Threshold 2: $(1,1)$ vs. $(1,0)$

Derivation. A student chooses $(1,1)$ over $(1,0)$ when $V(1,1) \geq V(1,0)$. Setting $V(1,1) = V(1,0)$. The terms $\ln(\bar{k}_1 - p^H)$ are identical in (B5) and (B4) and cancel, leaving

$$\begin{aligned}
&\beta_2\big[(1-\pi_c)\ln(\bar{k}_2 + y_2^{HS}) + \pi_c \ln(\bar{k}_2 - p^c)\big] + \beta_3\big[(1-\pi_c)\ln(\bar{k} + y^{HS}) + \pi_c \ln(\bar{k} + y^C)\big] \\
&\quad = \beta_2 \ln(\bar{k}_2 + y_2^{HS}) + \beta_3 \ln(\bar{k} + y^{HS}).
\end{aligned}$$

Subtract the right side from the left. In the β_2 block, $(1-\pi_c)\ln(\bar{k}_2 + y_2^{HS}) - \ln(\bar{k}_2 + y_2^{HS}) = -\pi_c \ln(\bar{k}_2 + y_2^{HS})$, and similarly in the β_3 block:

$$\pi_c\beta_2\big[\ln(\bar{k}_2 - p^c) - \ln(\bar{k}_2 + y_2^{HS})\big] + \pi_c\beta_3\big[\ln(\bar{k} + y^C) - \ln(\bar{k} + y^{HS})\big] = 0.$$

Dividing through by $\pi_c > 0$ and $\beta_2 > 0$, and exponentiating:

$$\frac{\bar{k}_2 - p^c}{\bar{k}_2 + y_2^{HS}} = \left(\frac{\bar{k} + y^{HS}}{\bar{k} + y^C}\right)^{\beta_3/\beta_2}.$$

Multiplying by $(\bar{k}_2 + y_2^{HS}) > 0$ and subtracting from $\bar{k}_2$:

$$\bar{p}_{11|10}^c = \bar{k}_2 - (\bar{k}_2 + y_2^{HS})\left(\frac{\bar{k} + y^{HS}}{\bar{k} + y^C}\right)^{\beta_3/\beta_2}. \tag{B6}$$

Since $V(1,1) - V(1,0)$ is strictly decreasing in p^c, a student chooses $(1,1)$ over $(1,0)$ if and only if $p^c \leq \bar{p}^c_{11|10}$. Additionally, since $(1,0)$ is available for that student, it must be true that $p^H \leq \bar{p}^H_{10}$.

B.3.3 Threshold 3: $(1,1)$ vs. $(0,0)$

Derivation. A student chooses $(1,1)$ over $(0,0)$ when $V(1,1) \geq V(0,0)$. Setting $V(1,1) = V(0,0)$:

$$\begin{aligned}\ln(\bar{k}_1 - p^H) + \beta_2\big[(1-\pi_c)\ln(\bar{k}_2 + y_2^{HS}) + \pi_c \ln(\bar{k}_2 - p^c)\big] + \beta_3\big[(1-\pi_c)\ln(\bar{k} + y^{HS}) + \pi_c \ln(\bar{k} + y^C)\big] \\ = \ln(\bar{k}_1 + y_1^{NH}) + \beta_2 \ln(\bar{k}_2 + y_2^{NH}) + \beta_3 \ln(\bar{k} + y^{NH}).\end{aligned}$$

Isolating p^c:

$$\begin{aligned}\pi_c\beta_2 \ln(\bar{k}_2 - p^c) = \ln\left(\frac{\bar{k}_1 + y_1^{NH}}{\bar{k}_1 - p^H}\right) + \beta_2 \ln(\bar{k}_2 + y_2^{NH}) - \beta_2(1-\pi_c)\ln(\bar{k}_2 + y_2^{HS}) \\ + \beta_3 \ln(\bar{k} + y^{NH}) - \beta_3(1-\pi_c)\ln(\bar{k} + y^{HS}) - \beta_3\pi_c \ln(\bar{k} + y^C).\end{aligned}$$

Divide every term by $\pi_c\beta_2 > 0$ and rearranging:

$$\ln(\bar{k}_2 - p^c) = \frac{1}{\pi_c}\left\{\frac{1}{\beta_2}\ln\left(\frac{\bar{k}_1 + y_1^{NH}}{\bar{k}_1 - p^H}\right) + \ln\left(\frac{\bar{k}_2 + y_2^{NH}}{(\bar{k}_2 + y_2^{HS})^{1-\pi_c}}\right) + \frac{\beta_3}{\beta_2}\ln\left(\frac{\bar{k} + y^{NH}}{(\bar{k} + y^{HS})^{1-\pi_c}(\bar{k} + y^C)^{\pi_c}}\right)\right\}.$$

$$\bar{p}^c_{11|00}(p^H) = \bar{k}_2 - \big[E(p^H)\big]^{1/\pi_c}, \tag{B7}$$

where, written out in full,

$$E(p^H) \;\equiv\; \left(\frac{\bar{k}_1 + y_1^{NH}}{\bar{k}_1 - p^H}\right)^{1/\beta_2} \cdot \frac{\bar{k}_2 + y_2^{NH}}{(\bar{k}_2 + y_2^{HS})^{1-\pi_c}} \cdot \left(\frac{\bar{k} + y^{NH}}{(\bar{k} + y^{HS})^{1-\pi_c}(\bar{k} + y^C)^{\pi_c}}\right)^{\beta_3/\beta_2} \;>\; 0. \tag{B8}$$

By the same monotonicity argument as in Section B.3.2, a student chooses $(1,1)$ over $(0,0)$ if and only if $p^c \leq \bar{p}^c_{11|00}(p^H)$. Since a student is comparing $(1,1)$ to $(0,0)$ and not $(1,0)$, it must be true that $p^H > \bar{p}^H_{10}$ and $(1,0)$ is not available for that student. Note that $\bar{p}^c_{11|00}$ is a *function* of p^H, since p^H appears in $E(p^H)$.

Boundary condition $\bar{p}^c_{11|00}(\bar{p}^H_{10}) = \bar{p}^c_{11|10}$. Evaluating $\bar{p}^c_{11|00}$ at $p^H = \bar{p}^H_{10}$ and using

$$\Pi \;\equiv\; \left(\frac{\bar{k}_2 + y_2^{NH}}{\bar{k}_2 + y_2^{HS}}\right)^{\beta_2}\left(\frac{\bar{k} + y^{NH}}{\bar{k} + y^{HS}}\right)^{\beta_3} \;>\; 0, \qquad \text{so} \qquad \bar{p}^H_{10} = \bar{k}_1 - (\bar{k}_1 + y_1^{NH})\,\Pi.$$

we have that $\bar{k}_1 - \bar{p}_{10}^H = (\bar{k}_1 + y_1^{NH})\Pi$, so:

$$\left(\frac{\bar{k}_1 + y_1^{NH}}{\bar{k}_1 - \bar{p}_{10}^H}\right)^{1/\beta_2} = \Pi^{-1/\beta_2} = \frac{\bar{k}_2 + y_2^{HS}}{\bar{k}_2 + y_2^{NH}} \left(\frac{\bar{k} + y^{HS}}{\bar{k} + y^{NH}}\right)^{\beta_3/\beta_2}.$$

Substituting into (B8), the period-2 factors give

$$\frac{\bar{k}_2 + y_2^{HS}}{\bar{k}_2 + y_2^{NH}} \cdot \frac{\bar{k}_2 + y_2^{NH}}{(\bar{k}_2 + y_2^{HS})^{1-\pi_c}} = (\bar{k}_2 + y_2^{HS})^{\pi_c},$$

since $(\bar{k}_2 + y_2^{NH})$ cancels and $x/x^{1-\pi_c} = x^{\pi_c}$; and the period-3 factors give

$$\left(\frac{\bar{k} + y^{HS}}{\bar{k} + y^{NH}} \cdot \frac{\bar{k} + y^{NH}}{(\bar{k} + y^{HS})^{1-\pi_c}(\bar{k} + y^C)^{\pi_c}}\right)^{\beta_3/\beta_2} = \left(\frac{\bar{k} + y^{HS}}{\bar{k} + y^C}\right)^{\pi_c\beta_3/\beta_2}.$$

Hence $E(\bar{p}_{10}^H) = (\bar{k}_2 + y_2^{HS})^{\pi_c}\big((\bar{k} + y^{HS})/(\bar{k} + y^C)\big)^{\pi_c\beta_3/\beta_2}$, and raising to $1/\pi_c$ divides every exponent by π_c:

$$\bar{k}_2 - \bar{p}_{11|00}^c(\bar{p}_{10}^H) = (\bar{k}_2 + y_2^{HS}) \left(\frac{\bar{k} + y^{HS}}{\bar{k} + y^C}\right)^{\beta_3/\beta_2} = \bar{k}_2 - \bar{p}_{11|10}^c.$$

Therefore

$$\bar{p}_{11|00}^c(\bar{p}_{10}^H) = \bar{p}_{11|10}^c. \tag{B9}$$

This is the internal consistency condition of the model: at $p^H = \bar{p}_{10}^H$ the student is indifferent between $(1, 0)$ and $(0, 0)$, so comparing $(1, 1)$ against either must yield the same critical college cost.

Strict monotonicity of $\bar{p}_{11|00}^c(\cdot)$. Writing $\big((\bar{k}_1 + y_1^{NH})/(\bar{k}_1 - p^H)\big)^{1/\beta_2} = (\bar{k}_1 + y_1^{NH})^{1/\beta_2}(\bar{k}_1 - p^H)^{-1/\beta_2}$ and raising $E(p^H)$ to $1/\pi_c$, the p^H-dependent factor becomes $(\bar{k}_1 - p^H)^{-1/(\pi_c\beta_2)}$. Setting $\alpha \equiv 1/(\pi_c\beta_2) > 0$,

$$\bar{p}_{11|00}^c(p^H) = \bar{k}_2 - \Lambda\,(\bar{k}_1 - p^H)^{-\alpha}, \Lambda \equiv \left[\frac{(\bar{k}_1 + y_1^{NH})^{1/\beta_2}(\bar{k}_2 + y_2^{NH})}{(\bar{k}_2 + y_2^{HS})^{1-\pi_c}} \left(\frac{\bar{k} + y^{NH}}{(\bar{k} + y^{HS})^{1-\pi_c}(\bar{k} + y^C)^{\pi_c}}\right)^{\beta_3/\beta_2}\right]^{1/\pi_c} > 0, \tag{B10}$$

where Λ contains no p^H. Differentiating:

$$\frac{\partial \bar{p}_{11|00}^c(p^H)}{\partial p^H} = -\Lambda\,\alpha\,(\bar{k}_1 - p^H)^{-\alpha-1} < 0.$$

So $\bar{p}^c_{11|00}(\cdot)$ is strictly decreasing, and with (B9),

$$\bar{p}^c_{11|00}(p^H) < \bar{p}^c_{11|00}(\bar{p}^H_{10}) = \bar{p}^c_{11|10} \qquad \text{for every } p^H > \bar{p}^H_{10}. \tag{B11}$$

A student who has paid a larger p^H has lower period-1 consumption, so the critical college cost at which she is willing to take the college path is lower. A useful identity, immediate from (B10), is

$$\Lambda\,(\bar{k}_1 - p^H)^{-\alpha} = \bar{k}_2 - \bar{p}^c_{11|00}(p^H), \qquad \text{and in particular} \qquad \Lambda\,(\bar{k}_1 - \bar{p}^H_{10})^{-\alpha} = \bar{k}_2 - \bar{p}^c_{11|10}. \tag{B12}$$

B.3.4 Decision Rule

Combining the three thresholds:

$$(s,d)^* = \begin{cases} (1,1) & \text{if } \left[p^H \le \bar{p}^H_{10} \text{ and } p^c \le \bar{p}^c_{11|10}\right] \text{ or } \left[p^H > \bar{p}^H_{10} \text{ and } p^c \le \bar{p}^c_{11|00}(p^H)\right], \\ (1,0) & \text{if } p^H \le \bar{p}^H_{10} \text{ and } p^c > \bar{p}^c_{11|10}, \\ (0,0) & \text{if } p^H > \bar{p}^H_{10} \text{ and } p^c > \bar{p}^c_{11|00}(p^H). \end{cases} \tag{B13}$$

The rule is exhaustive and the four cases are internally consistent:

- $p^H \le \bar{p}^H_{10}$, $p^c \le \bar{p}^c_{11|10}$: Threshold 1 gives $(1,0) \succ (0,0)$ and Threshold 2 gives $(1,1) \succ (1,0)$, so by transitivity $(1,1)$ is chosen.
- $p^H \le \bar{p}^H_{10}$, $p^c > \bar{p}^c_{11|10}$: $(1,0) \succ (0,0)$ and $(1,0) \succ (1,1)$, so $(1,0)$ is chosen.
- $p^H > \bar{p}^H_{10}$, $p^c \le \bar{p}^c_{11|00}(p^H)$: $(0,0) \succ (1,0)$, so $(1,0)$ is dominated; Threshold 3 gives $(1,1) \succ (0,0)$ and transitivity gives $(1,1) \succ (1,0)$, so $(1,1)$ is chosen.
- $p^H > \bar{p}^H_{10}$, $p^c > \bar{p}^c_{11|00}(p^H)$: $(0,0)$ dominates both alternatives.

B.4 Aggregate High School Graduation

Definition. S^* is the aggregate share of students who graduate high school, i.e. who choose $s = 1$ either through $(1,0)$ or $(1,1)$.

Derivation. From (B13), a student graduates in exactly one of two mutually exclusive ways: either $p^H \le \bar{p}^H_{10}$, in which case she chooses $(1,0)$ or $(1,1)$ depending on p^c but graduates either way,

or $p^H > \bar{p}^H_{10}$ and $p^c \le \bar{p}^c_{11|00}(p^H)$, in which case she moves directly from $(0,0)$ to $(1,1)$. Hence

$$S^* = \Pr(p^H \le \bar{p}^H_{10}) + \Pr\big(p^H > \bar{p}^H_{10} \text{ and } p^c \le \bar{p}^c_{11|00}(p^H)\big).$$

Since the threshold $\bar{p}^c_{11|00}(p^H)$ depends on the realization of p^H, I apply the law of total probability and conditioning on p^H to get:

$$\Pr\big(p^H > \bar{p}^H_{10} \text{ and } p^c \le \bar{p}^c_{11|00}(p^H)\big) = \int_{\bar{p}^H_{10}}^{\bar{P}^H} \Pr\big(p^c \le \bar{p}^c_{11|00}(p^H) \,\big|\, p^H\big)\, f(p^H)\, dp^H,$$

with $f(p^H) = 1/\bar{P}^H$. Because the two cost draws are *independent random variables*, the conditional probability equals the unconditional distribution function of p^c evaluated at the realized threshold: $\Pr(p^c \le \bar{p}^c_{11|00}(p^H) \mid p^H) = \bar{p}^c_{11|00}(p^H)/\bar{P}^c$, valid whenever $\bar{p}^c_{11|00}(p^H) \in [0, \bar{P}^c]$, which I assume throughout (Assumption 2 below), together with $\bar{p}^H_{10} \in (0, \bar{P}^H)$. Writing $\theta_H \equiv \bar{p}^H_{10}/\bar{P}^H$,

$$S^* = \theta_H + \frac{1}{\bar{P}^H \bar{P}^c} \int_{\bar{p}^H_{10}}^{\bar{P}^H} \bar{p}^c_{11|00}(p^H)\, dp^H. \tag{B14}$$

[Regularity] $\bar{p}^H_{10} \in (0, \bar{P}^H)$.

[Interior solution] $\bar{p}^c_{11|00}(p^H) \in (0, \bar{P}^c)$ for all $p^H \in [\bar{p}^H_{10}, \bar{P}^H]$.

Closed form. Substituting $\bar{p}^c_{11|00}(p^H)$ below into (B14)

$$\bar{p}^c_{11|00}(p^H) = \bar{k}_2 - \Lambda\,(\bar{k}_1 - p^H)^{-\alpha}, \Lambda \equiv \left[\frac{(\bar{k}_1 + y_1^{NH})^{1/\beta_2}(\bar{k}_2 + y_2^{NH})}{(\bar{k}_2 + y_2^{HS})^{1-\pi_c}} \left(\frac{\bar{k} + y^{NH}}{(\bar{k} + y^{HS})^{1-\pi_c}(\bar{k} + y^C)^{\pi_c}}\right)^{\beta_3/\beta_2}\right]^{1/\pi_c} > 0,$$

We have:

$$\int_{\bar{p}^H_{10}}^{\bar{P}^H} \bar{p}^c_{11|00}(p^H)\, dp^H = \bar{k}_2(\bar{P}^H - \bar{p}^H_{10}) - \Lambda \int_{\bar{p}^H_{10}}^{\bar{P}^H} (\bar{k}_1 - p^H)^{-\alpha}\, dp^H,$$

and substituting $u = \bar{k}_1 - p^H$, $du = -dp^H$,

$$\int_{\bar{p}^H_{10}}^{\bar{P}^H} (\bar{k}_1 - p^H)^{-\alpha}\, dp^H = \int_{\bar{k}_1 - \bar{P}^H}^{\bar{k}_1 - \bar{p}^H_{10}} u^{-\alpha}\, du = \frac{(\bar{k}_1 - \bar{p}^H_{10})^{1-\alpha} - (\bar{k}_1 - \bar{P}^H)^{1-\alpha}}{1-\alpha},$$

valid for $\alpha \neq 1$, i.e. $\pi_c \beta_2 \neq 1$. The expression is strictly positive in both cases: for $\alpha < 1$ the numerator is positive and $1-\alpha > 0$; for $\alpha > 1$ both are negative. Hence

$$S^* = \theta_H + \frac{1}{\bar{P}^H \bar{P}^c} \underbrace{\left\{ \bar{k}_2(\bar{P}^H - \bar{p}^H_{10}) - \Lambda \, \frac{(\bar{k}_1 - \bar{p}^H_{10})^{1-\alpha} - (\bar{k}_1 - \bar{P}^H)^{1-\alpha}}{1-\alpha} \right\}}_{\equiv \; \mathcal{S} \; = \; \int_{\bar{p}^H_{10}}^{\bar{P}^H} \bar{p}^c_{11|00}(p^H)\, dp^H} \tag{B15}$$

B.4.1 Monotonicity of S^*

If the share of individuals completing high school as represented by S^* is a monotonic function of the thresholds, I can analyze changes in the thresholds to understand how policy-induced changes in parameters can affect S^*.

Independence from $\bar{p}^c_{11|10}$. $\bar{p}^c_{11|10}$ governs only the split between $(1,0)$ and $(1,1)$ among students with $p^H \leq \bar{p}^H_{10}$. Both choices involve graduating, so moving a student between them leaves S^* unchanged. Therefore, $\bar{p}^c_{11|10}$ appears in neither (B14) nor (B15).

S^* is strictly increasing in $\bar{p}^H_{10}$. Differentiating (B15) directly: $\partial \theta_H / \partial \bar{p}^H_{10} = 1/\bar{P}^H$ and $\partial \big[\bar{k}_2(\bar{P}^H - \bar{p}^H_{10})\big] / \partial \bar{p}^H_{10} = -\bar{k}_2$. Since only $(\bar{k}_1 - \bar{p}^H_{10})^{1-\alpha}$ depends on $\bar{p}^H_{10}$ in the remaining term, and by the chain rule, we have that $\partial (\bar{k}_1 - \bar{p}^H_{10})^{1-\alpha} / \partial \bar{p}^H_{10} = -(1-\alpha)(\bar{k}_1 - \bar{p}^H_{10})^{-\alpha}$, then:

$$\frac{\partial}{\partial \bar{p}^H_{10}} \left[-\Lambda \, \frac{(\bar{k}_1 - \bar{p}^H_{10})^{1-\alpha} - (\bar{k}_1 - \bar{P}^H)^{1-\alpha}}{1-\alpha} \right] = +\Lambda \, (\bar{k}_1 - \bar{p}^H_{10})^{-\alpha},$$

By (B12), we have that $\Lambda \, (\bar{k}_1 - \bar{p}^H_{10})^{-\alpha} = \bar{k}_2 - \bar{p}^c_{11|10}$. Replacing it and combining both derivatives:

$$\frac{\partial S^*}{\partial \bar{p}^H_{10}} = \frac{1}{\bar{P}^H} - \frac{\bar{p}^c_{11|10}}{\bar{P}^H \bar{P}^c} = \frac{1}{\bar{P}^H} \left(1 - \frac{\bar{p}^c_{11|10}}{\bar{P}^c} \right) = \frac{1 - \theta_{c,10}}{\bar{P}^H} > 0, \tag{B16}$$

with $\theta_{c,10} \equiv \bar{p}^c_{11|10} / \bar{P}^c$, strictly positive because $\theta_{c,10} < 1$ under Assumption B.4.

S^* is strictly increasing in the college threshold function. Using the integral form of (B14), in which $\bar{p}^c_{11|00}(\cdot)$ appears explicitly as the integrand. Since S^* is an affine functional of $\bar{p}^c_{11|00}(\cdot)$ with strictly positive coefficient $1/(\bar{P}^H \bar{P}^c)$, the functional derivative at any $p^H_0 \in (\bar{p}^H_{10}, \bar{P}^H)$ is

$$\frac{\delta S^*}{\delta \, \bar{p}^c_{11|00}(p^H_0)} = \frac{1}{\bar{P}^H \bar{P}^c} > 0.$$

So any parameter change that raises $\bar{p}^c_{11|00}(p^H)$ within $[\bar{p}^H_{10}, \bar{P}^H]$, and lowers it nowhere, strictly raises S^*.

B.5 Necessary Conditions

The share of individuals completing high school might move as a result of changes in two channels: the indirect labor market channels, which moves students from $(0,0)$ to $(1,0)$ and increases S^* by increasing θ_H , and the direct education effect via transitions from $(0,0)$ to $(1,1)$, which operates through the second term. In this section, I derive the necessary conditions for these transitions to take place.

B.5.1 Condition 1: a non-empty pool at $(0,0)$, $\bar{p}^H_{10} < \bar{P}^H$

For either channel to be active, the path $(0,0)$ need to attract at least some students pre-policy. A student is at $(0,0)$ if and only if $p^H > \bar{p}^H_{10}$ *and* $p^c > \bar{p}^c_{11|00}(p^H)$. Conditioning on p^H and using independence of the two draws,

$$\begin{aligned}\Pr\big((s,d)^* = (0,0)\big) &= \int_{\bar{p}^H_{10}}^{\bar{P}^H} \Pr\big(p^c > \bar{p}^c_{11|00}(p^H) \,\big|\, p^H\big)\, f(p^H)\, dp^H = \int_{\bar{p}^H_{10}}^{\bar{P}^H} \frac{\bar{P}^c - \bar{p}^c_{11|00}(p^H)}{\bar{P}^c} \cdot \frac{1}{\bar{P}^H}\, dp^H \\ &= \frac{1}{\bar{P}^H \bar{P}^c} \int_{\bar{p}^H_{10}}^{\bar{P}^H} \big[\bar{P}^c - \bar{p}^c_{11|00}(p^H)\big]\, dp^H,\end{aligned}$$

where the middle step uses $\Pr(p^c > \bar{p}^c_{11|00}(p^H)) = 1 - \bar{p}^c_{11|00}(p^H)/\bar{P}^c$, which is valid because Assumption B.4 places $\bar{p}^c_{11|00}(p^H)$ inside $(0, \bar{P}^c)$.

This condition implies that $\bar{p}^H_{10} < \bar{P}^H$. Assume $\bar{p}^H_{10} \geq \bar{P}^H$. Then every student draws $p^H \leq \bar{P}^H \leq \bar{p}^H_{10}$, so every student graduates and $S^* = 1$. Hence no parameter change can raise S^*, and $\bar{p}^H_{10} < \bar{P}^H$ is necessary. Recall the formula for $\bar{p}^H_{10}$:

$$\bar{p}^H_{10} = \bar{k}_1 - (\bar{k}_1 + y_1^{NH}) \left(\frac{\bar{k}_2 + y_2^{NH}}{\bar{k}_2 + y_2^{HS}}\right)^{\beta_2} \left(\frac{\bar{k} + y^{NH}}{\bar{k} + y^{HS}}\right)^{\beta_3}.$$

As $y^{HS} \to \infty$ both ratios tend to 0 and $\bar{p}^H_{10} \to \bar{k}_1$. If $\bar{k}_1 \geq \bar{P}^H$, then $\theta_H \to 1$ and the pool at $(0,0)$ disappears. Condition 1 therefore places an *upper* bound on the high-school wage premium: it must not be so large that all students graduate even in the complete absence of any college option.

B.5.2 Condition 2: a reachable pool

Condition 1 guarantees that students sit at $(0,0)$. However, it does not guarantee that either transitions can take place. That is, that students sitting at $(0,0)$ can move to either $(1,0)$ or $(1,1)$. Below, I derive the necessary conditions for each transition to take place.

Necessary conditions for the indirect labor market channel. For $(0,0)$ to $(1,0)$ transitions to take place, it needs to be true that $p^H < \bar{p}^H_{10}$ is attainable for at least some students. That implies that $\bar{p}^H_{10}$ must be positive, since $p^H > 0$.

From:

$$\bar{p}^H_{10} = \bar{k}_1 - (\bar{k}_1 + y_1^{NH}) \left(\frac{\bar{k}_2 + y_2^{NH}}{\bar{k}_2 + y_2^{HS}} \right)^{\beta_2} \left(\frac{\bar{k} + y^{NH}}{\bar{k} + y^{HS}} \right)^{\beta_3}.$$

$\bar{p}^H_{10} > 0$ if and only if $(\bar{k}_1 + y_1^{NH}) \left(\frac{\bar{k}_2 + y_2^{NH}}{\bar{k}_2 + y_2^{HS}} \right)^{\beta_2} \left(\frac{\bar{k} + y^{NH}}{\bar{k} + y^{HS}} \right)^{\beta_3} < \bar{k}_1$. Dividing by $(\bar{k}_1 + y_1^{NH}) > 0$ and substituting $\bar{k}_1 = 3\bar{k}$, $y_1^{NH} = 3y^{NH}$ on the right,

$$\frac{\bar{k}_1}{\bar{k}_1 + y_1^{NH}} = \frac{3\bar{k}}{3(\bar{k} + y^{NH})} = \frac{\bar{k}}{\bar{k} + y^{NH}},$$

so the exact necessary and sufficient condition is

$$\bar{p}^H_{10} > 0 \iff \left(\frac{\bar{k}_2 + y_2^{NH}}{\bar{k}_2 + y_2^{HS}} \right)^{\beta_2} \left(\frac{\bar{k} + y^{NH}}{\bar{k} + y^{HS}} \right)^{\beta_3} < \frac{\bar{k}}{\bar{k} + y^{NH}}. \tag{B17}$$

If $y^{HS} \leq y^{NH}$ then $y_2^{HS} \leq y_2^{NH}$, so both ratios are ≥ 1. Since $y^{NH} > 0$, $\bar{k}/(\bar{k} + y^{NH}) < 1$. Thus, $y^{HS} \leq y^{NH}$ implies that $\bar{p}^H_{10} \leq 0$. Contrapositively, $\bar{p}^H_{10} > 0 \Rightarrow y^{HS} > y^{NH}$.

Necessary conditions for the direct education channel. For $(0,0)$ to $(1,1)$ transitions to take place, it needs to be true that $p^c < \bar{p}^c_{11|00}$ is attainable for at least some students. That implies that $\bar{p}^c_{11|00}$ is positive, since $p^c > 0$. By (B11), if $\bar{p}^c_{11|10} \leq 0$, then $\bar{p}^c_{11|00} < 0$. Thus, $\bar{p}^c_{11|10} > 0$ implies $\bar{p}^c_{11|00} > 0$. I can then study when $\bar{p}^c_{11|10} > 0$ to determine the necessary conditions for $\bar{p}^c_{11|00} > 0$.

Using $\bar{k}_2 = 4\bar{k}$ and $\bar{k}_2 + y_2^{HS} = 4(\bar{k} + y^{HS})$:

$$\bar{p}^c_{11|10} = 4 \left[\bar{k} - (\bar{k} + y^{HS}) \left(\frac{\bar{k} + y^{HS}}{\bar{k} + y^C} \right)^{\beta_3/\beta_2} \right].$$

Note that at $y^C = y^{HS}$ the ratio equals 1 and $\bar{p}^c_{11|10} = 4[\bar{k} - (\bar{k} + y^{HS})] = -4y^{HS} < 0$. As $y^C \to \infty$ the ratio tends to 0 and $\bar{p}^c_{11|10} \to 4\bar{k} > 0$. Since $\bar{p}^c_{11|10}$ is increasing in y^C, there is a

unique threshold $\bar{y}^C$ where $\bar{p}^c_{11|10} = 0$, and for which any $y^C > \bar{y}^C$ implies that $\bar{p}^c_{11|10} > 0$. Then $\bar{p}^c_{11|10} > 0 \iff y^C > \bar{y}^C$. Setting $\bar{p}^c_{11|10} = 0$ and dividing by $4(\bar{k} + y^{HS})$:

$$\frac{\bar{k}}{\bar{k} + y^{HS}} = \left(\frac{\bar{k} + y^{HS}}{\bar{k} + \bar{y}^C}\right)^{\beta_3/\beta_2}.$$

Raising both sides to the power $\beta_2/\beta_3 > 0$ and multiplying by $(\bar{k} + y^{HS})$:

$$\bar{y}^C = (\bar{k} + y^{HS})\left(\frac{\bar{k} + y^{HS}}{\bar{k}}\right)^{\beta_2/\beta_3} - \bar{k}. \tag{B18}$$

Since $y^{HS} > 0$, it follows that $\left((\bar{k} + y^{HS})/\bar{k}\right)^{\beta_2/\beta_3} > 1$, and thus $\bar{y}^C > y^{HS}$ strictly. Therefore, a necessary condition for $\bar{p}^c_{11|10} > 0$, and consequently, $\bar{p}^c_{11|00} > 0$, is that $y^C > \bar{y}^C > y^{HS}$.

B.6 Comparative Statics

B.6.1 The chain rule

To analyze how the share of students completing high school changes with policy-driven changes in parameters, I differentiate the closed form expression (B15) with respect to key parameters. Recall $\mathcal{S}$ denotes the braced expression, and by construction $\mathcal{S} = \int_{\bar{p}^H_{10}}^{\bar{P}^H} \bar{p}^c_{11|00}(p^H)\, dp^H$. For a given parameter x,

$$\frac{\partial S^*}{\partial x} = \frac{1}{\bar{P}^H}\frac{\partial \bar{p}^H_{10}}{\partial x} + \frac{1}{\bar{P}^H \bar{P}^c}\frac{d\mathcal{S}}{dx}, \qquad \frac{d\mathcal{S}}{dx} = \frac{\partial \mathcal{S}}{\partial \bar{p}^H_{10}}\frac{\partial \bar{p}^H_{10}}{\partial x} + \left.\frac{\partial \mathcal{S}}{\partial x}\right|_{\bar{p}^H_{10} \text{ fixed}},$$

where $\partial \mathcal{S}/\partial \bar{p}^H_{10} = -\bar{p}^c_{11|10}$. Collecting the terms in $\partial \bar{p}^H_{10}/\partial x$ and using $1/\bar{P}^H - \bar{p}^c_{11|10}/(\bar{P}^H \bar{P}^c) = (1 - \theta_{c,10})/\bar{P}^H$:

$$\frac{\partial S^*}{\partial x} = \underbrace{\frac{1 - \theta_{c,10}}{\bar{P}^H} \cdot \frac{\partial \bar{p}^H_{10}}{\partial x}}_{\text{(0,0) to (1,0)}} + \underbrace{\frac{1}{\bar{P}^H \bar{P}^c}\int_{\bar{p}^H_{10}}^{\bar{P}^H} \frac{\partial \bar{p}^c_{11|00}(p^H)}{\partial x}\, dp^H}_{\text{(0,0) to (1,1)}}. \tag{B19}$$

B.6.2 Signing the college-threshold term

Define $\Delta(p^c, x; p^H) \equiv V(1,1) - V(0,0)$:

$$V(1,1) - V(0,0) = \ln\left(\frac{\bar{k}_1 - p^H}{\bar{k}_1 + y_1^{NH}}\right) + \beta_2 \ln\left(\frac{(\bar{k}_2 + y_2^{HS})^{1-\pi_c}(\bar{k}_2 - p^c)^{\pi_c}}{\bar{k}_2 + y_2^{NH}}\right) + \beta_3 \ln\left(\frac{(\bar{k} + y^{HS})^{1-\pi_c}(\bar{k} + y^C)^{\pi_c}}{\bar{k} + y^{NH}}\right), \tag{B20}$$

So that $\Delta = 0$ at $p^c = \bar{p}^c_{11|00}(p^H)$. The only p^c-dependent term is $\pi_c\beta_2 \ln(\bar{k}_2 - p^c)$, so

$$\frac{\partial \Delta}{\partial p^c} = -\frac{\pi_c \beta_2}{\bar{k}_2 - p^c} < 0,$$

which is strictly negative because $p^c < \bar{k}_2$. The implicit function theorem therefore gives

$$\frac{\partial \bar{p}^c_{11|00}(p^H)}{\partial x} = -\frac{\partial \Delta / \partial x}{\partial \Delta / \partial p^c} = \frac{\bar{k}_2 - \bar{p}^c_{11|00}(p^H)}{\pi_c \beta_2} \cdot \frac{\partial \Delta}{\partial x}, \tag{B21}$$

Since $\frac{\bar{k}_2 - \bar{p}^c_{11|00}(p^H)}{\pi_c \beta_2} > 0$, then $\operatorname{sign}\big(\partial \bar{p}^c_{11|00}/\partial x\big) = \operatorname{sign}\big(\partial \Delta / \partial x\big)$, and it suffices to sign $\partial \Delta / \partial x$.

B.6.3 Effect of p^c via the support $\bar{P}^c$

$\bar{P}^c$ enters S^* only through the explicit factor $1/\bar{P}^c$ in (B14), and

$$\frac{\partial S^*}{\partial \bar{P}^c} = -\frac{1}{\bar{P}^H \, \bar{P}^{c2}} \int_{\bar{p}^H_{10}}^{\bar{P}^H} \bar{p}^c_{11|00}(p^H)\, dp^H. \tag{B22}$$

Under the necessary condition that $\bar{p}^c_{11|10} > 0$, $\bar{p}^c_{11|00}(\bar{p}^H_{10}) = \bar{p}^c_{11|10} > 0$ by (B9) and, by continuity, $\bar{p}^c_{11|00} > 0$ on a set of positive measure, so the integral is strictly positive and (B22) is strictly negative. On the other hand, if the necessary condition is not met and If $\bar{p}^c_{11|10} \leq 0$, then by (B11) $\bar{p}^c_{11|00}(p^H) \leq 0$ throughout the range, and $\Pr(p^c \leq \bar{p}^c_{11|00}(p^H)) = 0$ for every student. The integral is zero, and $\partial S^*/\partial \bar{P}^c = 0$. Hence

$$\frac{\partial S^*}{\partial \bar{P}^c} < 0 \quad \text{if and only if} \quad \bar{p}^c_{11|10} > 0, \tag{B23}$$

so lowering college costs unambiguously raises S^* exactly when the college channel is non-empty.

B.6.4 Effect of π_c

Changes in S^* brought about by changes in π_c run only through the educational opportunities channel, so $\partial \bar{p}^H_{10}/\partial \pi_c = 0$. Then, through the education opportunities channel, we have that:

$$\frac{\partial \Delta}{\partial \pi_c} = \beta_2 \ln\left(\frac{\bar{k}_2 - p^c}{\bar{k}_2 + y^{HS}_2}\right) + \beta_3 \ln\left(\frac{\bar{k} + y^C}{\bar{k} + y^{HS}}\right), \tag{B24}$$

The first term of (B24) is strictly negative for *every* student and every admissible p^c, because $y^{HS}_2 > 0$ and $p^c > 0$ give $\bar{k}_2 - p^c < \bar{k}_2 < \bar{k}_2 + y^{HS}_2$. This is the period-2 cost of raising the admission

probability, since a larger π_c makes the student more likely to be paying p^c instead of earning y_2^{HS}. The second term is the period-3 benefit, which is positive when $y^C > y^{HS}$. At an *arbitrary* p^c, the sign of (B24) is sign is ambiguous even when $y^C > y^{HS}$, since which term will dominate depends on the size of p^c.

By (B21), $\partial \bar{p}^c_{11|00}(p^H)/\partial \pi_c$ is $\partial \Delta / \partial \pi_c$ evaluated *at* $p^c = \bar{p}^c_{11|00}(p^H)$, which is the point of indifference (threshold) and the only point that moves when π_c changes. By (B11), $\bar{p}^c_{11|00}(p^H) < \bar{p}^c_{11|10}$ for $p^H > \bar{p}^H_{10}$, so $\bar{k}_2 - \bar{p}^c_{11|00}(p^H) > \bar{k}_2 - \bar{p}^c_{11|10}$, and by (B6) the latter equals $(\bar{k}_2 + y_2^{HS})\big((\bar{k} + y^{HS})/(\bar{k} + y^C)\big)^{\beta_3/\beta_2}$. Dividing by $(\bar{k}_2 + y_2^{HS}) > 0$, taking logarithms and multiplying by $\beta_2 > 0$,

$$\beta_2 \ln\left(\frac{\bar{k}_2 - \bar{p}^c_{11|00}(p^H)}{\bar{k}_2 + y_2^{HS}}\right) > -\beta_3 \ln\left(\frac{\bar{k} + y^C}{\bar{k} + y^{HS}}\right),$$

i.e. $\partial \Delta / \partial \pi_c > 0$ at the threshold.

Under the necessary condition that $\bar{p}^c_{11|10} > 0$, $\partial \bar{p}^c_{11|00}(p^H)/\partial \pi_c > 0$ for every $p^H \in (\bar{p}^H_{10}, \bar{P}^H)$. If $\bar{p}^c_{11|10} \leq 0$ then $\bar{p}^c_{11|00}(p^H) \leq 0$ throughout, and the college channel is identically zero for every π_c. Hence:

$$\frac{\partial S^*}{\partial \pi_c} = \frac{1}{\bar{P}^H \bar{P}^c} \int_{\bar{p}^H_{10}}^{\bar{P}^H} \frac{\partial \bar{p}^c_{11|00}(p^H)}{\partial \pi_c} \, dp^H > 0 \quad \text{if and only if} \quad \bar{p}^c_{11|10} > 0. \tag{B25}$$

B.6.5 Effect of y^{HS}

Changes in high-school level wages operate through both the indirect labor market and the direct educational channel. Let:

$\bar{p}^H_{10} = \bar{k}_1 - (\bar{k}_1 + y_1^{NH})\Pi$ with

$$\ln \Pi = \beta_2 \ln\left(\frac{\bar{k}_2 + y_2^{NH}}{\bar{k}_2 + y_2^{HS}}\right) + \beta_3 \ln\left(\frac{\bar{k} + y^{NH}}{\bar{k} + y^{HS}}\right),$$

so $\partial \ln \Pi / \partial y^{HS} = -(\beta_2 + \beta_3)/(\bar{k} + y^{HS})$

Using $\bar{k}_1 + y_1^{NH} = 3(\bar{k} + y^{NH})$ throughout,

$$\frac{\partial \bar{p}^H_{10}}{\partial y^{HS}} = \frac{3(\bar{k} + y^{NH})(\beta_2 + \beta_3)\,\Pi}{\bar{k} + y^{HS}} \; > \; 0, \tag{B26}$$

On the direct educational channel, we have that:

$$\frac{\partial \Delta}{\partial y^{HS}} = \frac{(1-\pi_c)(\beta_2+\beta_3)}{\bar{k}+y^{HS}} \;>\; 0, \tag{B27}$$

Thus, higher high-school wages raise the high school threshold, pulling students from $(0,0)$ into $(1,0)$, and the college threshold rises too, because y^{HS} is earned in the non-admission state on the college path. Both terms of (B19) are strictly positive, so

$$\frac{\partial S^*}{\partial y^{HS}} > 0 \tag{B28}$$

B.6.6 Effect of y^{NH}

Similarly, changes in high-school dropout wages operate through both channels. Let: $\bar{p}_{10}^{H} = \bar{k}_1 - (\bar{k}_1 + y_1^{NH})\Pi$ with

$$\ln \Pi = \beta_2 \ln\left(\frac{\bar{k}_2 + y_2^{NH}}{\bar{k}_2 + y_2^{HS}}\right) + \beta_3 \ln\left(\frac{\bar{k}+y^{NH}}{\bar{k}+y^{HS}}\right),$$

so $\partial \ln \Pi / \partial y^{NH} = +(\beta_2+\beta_3)/(\bar{k}+y^{NH})$. Using $\bar{k}_1 + y_1^{NH} = 3(\bar{k}+y^{NH})$ throughout,

$$\frac{\partial \bar{p}_{10}^{H}}{\partial y^{NH}} = -3\,\Pi\,(1+\beta_2+\beta_3) \;<\; 0, \tag{B29}$$

On the direct educational channel, we have:

$$\frac{\partial \Delta}{\partial y^{NH}} = -\frac{1+\beta_2+\beta_3}{\bar{k}+y^{NH}} \;<\; 0, \tag{B30}$$

Both terms of (B19) are strictly negative, so

$$\frac{\partial S^*}{\partial y^{NH}} < 0 \tag{B31}$$

B.6.7 Effect of y^{C}

College earnings operate solely through the direct education channel. We have:

$$\frac{\partial \Delta}{\partial y^{C}} = \frac{\beta_3 \pi_c}{\bar{k}+y^{C}} \;>\; 0, \tag{B32}$$

And thus:

$$\frac{\partial \bar{p}^c_{11|00}(p^H)}{\partial y^C} = \frac{\bar{k}_2 - \bar{p}^c_{11|00}(p^H)}{\pi_c \beta_2} \cdot \frac{\beta_3 \pi_c}{\bar{k} + y^C} = \frac{\beta_3 \big(\bar{k}_2 - \bar{p}^c_{11|00}(p^H)\big)}{\beta_2 (\bar{k} + y^C)} \; > \; 0,$$

the factor π_c cancelling. Hence

$$\frac{\partial S^*}{\partial y^C} = \frac{\beta_3}{\beta_2 \, \bar{P}^H \, \bar{P}^c \, (\bar{k} + y^C)} \int_{\bar{p}^H_{10}}^{\bar{P}^H} \big(\bar{k}_2 - \bar{p}^c_{11|00}(p^H)\big) \, dp^H > 0 \quad \text{if and only if} \quad \bar{p}^c_{11|10} > 0. \tag{B33}$$

Under the necessary condition that $\bar{p}^c_{11|10} > 0$, an increase in college-level wages should increase the share of students graduating high school.

C Background: the three phases of the HEI expansion

The HEI expansion led by the federal government was divided in three phases. The first expansion phase, called Expansion I, occurred between 2003 and 2007, and its main goal was to expand higher education institutions into municipalities in the inner part of the country, thus earning the name "interiorization phase". During this period, 98 institutions were created. The second expansion phase, called REUNI, occurred between 2008 and 2012 and combined the continuation of the expansion started in 2003 with the restructuring of existing universities, along with expanded enrollment capacity and infrastructure improvements, including new buildings and classrooms. In this second phase, 247 new institutions were created.

The third phase of the expansion occurred between 2012 and 2014, although institutions continued to open up until 2017. It focused on coupling higher education expansion with initiatives to develop underserved regions and on launching an array of public programs, such as policies to expand medical school enrollment and to include people with disabilities in higher education. All these additional policies targeted college students enrolled in these newly opened institutions. During this phase, 239 HEIs were created.