The Impact of Online Technical Education on Schooling Outcomes: Evidence from Brazil

Gregory Elacqua
Patricia Navarro-Palau
María Fernanda Prada
Sammara Cavalcanti Soares

Education Division
Social Sector

TECHNICAL NOTE Nº IDB-TN-02292

October 2021

The Impact of Online Technical Education on Schooling Outcomes: Evidence from Brazil

Gregory Elacqua
Patricia Navarro-Palau
María Fernanda Prada
Sammara Cavalcanti Soares

October 2021

Inter-American Development Bank

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

The impact of online technical education on schooling outcomes: evidence from Brazil / Gregory Elacqua, Patricia Navarro-Palau, María Fernanda Prada, Sammara Cavalcanti Soares.
p. cm. — (IDB Technical Note ; 2292)
Includes bibliographic references.
1. Telecommunication in education-Econometric models-Brazil. 2. College preparation programs- Government policy-Brazil. 3. High school dropouts-Prevention-Brazil. I. Elacqua, Gregory M., 1972-. II. Navarro-Palau, Patricia. III. Prada, María Fernanda. IV. Soares, Sammara. V. Inter-American Development Bank. Education Division. VI. Series.
IDB-TN-2292

Keywords: Technical education, online education, elearning, secondary education, Brazil
JEL Classification: I21, I28

http://www.iadb.org

Copyright © 2021 Inter-American Development Bank. This work is licensed under a Creative Commons IGO 3.0 Attribution-NonCommercial-NoDerivatives (CC-IGO BY-NC-ND 3.0 IGO) license (http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode) and may be reproduced with attribution to the IDB and for any non-commercial purpose. No derivative work is allowed.

Any dispute related to the use of the works of the IDB that cannot be settled amicably shall be submitted to arbitration pursuant to the UNCITRAL rules. The use of the IDB's name for any purpose other than for attribution, and the use of IDB's logo shall be subject to a separate written license agreement between the IDB and the user and is not authorized as part of this CC-IGO license.

Note that link provided above includes additional terms and conditions of the license.

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.

BY NC ND

1300 New York Ave, Washington DC, 20577

The impact of online technical education on schooling outcomes: Evidence from Brazil

Gregory Elacqua* Patricia Navarro-Palau†‡
María Fernanda Prada* Sammara Cavalcanti Soares* ‡

October 26, 2021

*Inter-American Development Bank (IADB), 1300 New York Ave NW, Washington, DC 20577. Elacqua: gregorye@iadb.org; Prada: mariafp@iadb.org; Soares: sammarac@iadb.org.

†Organisation for Economic Co-operation and Development (OECD), 2, rue André Pascal 75016 Paris.

‡We thank the Secretary of Education of Pernambuco for providing the data used in this study, Vinícius Schuabb for data processing and Ivan Gonzaga Rabello Brant for data assistance. We are also thankful to the Inter-American Development Bank for funding the research presented in this paper. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent. The authors have no conflicts of interests or financial and material interests in the results. All errors remain our own.

Abstract

This paper studies the impact of online technical education offered to complement regular academic instruction in high school on student schooling outcomes. Using a regression discontinuity design with an oversubscribed large-scale online technical course in Brazil, we find that students who score above the cutoff on the online technical education admission exam are less likely to drop out of high school, while their performance on standardized tests in math and Portuguese is similar to that of students just below the admission exam cutoff. Overall, we provide evidence that complementing high school regular instruction with technical education in an *online* format can be an effective public policy to increase students' work readiness as it reduces the dropout rate from secondary education without negatively affecting students' academic proficiency.

JEL classification: I21, I28

Keywords: Technical education, online education, elearning, secondary education, Brazil

1 Introduction

Advocates have argued that secondary technical education may be a solution to equip youth with the skills needed to improve their employability, smooth the school to work transition, and help address skill shortages. In fact, the ability of this type of education to foster work readiness and close connection to the world of work makes it an attractive option to motivate and engage young people, especially those who are less academically oriented and who come from more disadvantaged backgrounds (Hoeckel & Schwartz (2010)). However, the heterogeneity in the provision of this type of education and the scant literature about it make it difficult to provide definite conclusions about the benefits or disadvantages of technical education during high school.[1]

Proponents have argued that technical education in high school can increase student motivation and engagement, which, in turn, may have an effect on academic outcomes (e.g., Shernoff et al. (2014); Carbonaro (2005); Ryan (2001)). Indeed, recent experimental and quasi-experimental literature has shown the positive effects of technical education combined with academic education during high school on schooling outcomes, particularly on the probability of remaining in and graduating from high school (Kemple & Snipes (2000); Dougherty (2018); Brunner et al. (2021); Hemelt et al. (2019); Elacqua et al. (2019)). In contrast, skeptics maintain that technical education takes time from more relevant core subjects, such as math (World Bank (2017)).

The COVID-19 pandemic has sparked an unprecedented disruption in the provision of education and training, forcing most schools to move to remote classes. Distance virtual/remote education was already emerging as a trend even before the pandemic, which has accelerated investments in technology and education innovation in many school systems (Chun et al. (2021)). Moreover, the educational solutions for the "new normal" have yet to be defined. Post-pandemic education is likely to feature similar or greater levels of online and hybrid education models, with long-lasting and even permanent changes in instruction and learning (Elacqua et al. (2020); Arias et al. (2020)).

[1]The provision of technical education during high school is heterogeneous around the world and even within countries. It ranges from a variety of elective career-oriented courses integrated within a traditional high school curriculum, as is the case in the US, to a fully separate education track offered in technical high schools, such as in Germany and most countries in Latin America and the Caribbean. Technical education, furthermore, encompasses a wide array of fields that affect the type, nature, and intensity of the learning experience. The quality of provision across institutions also varies. This heterogeneity, coupled with the limited literature on the subject, make it difficult to reach conclusions about the benefits or disadvantages of technical education during high school

While still very limited, especially in primary and secondary school years, the evidence shows that online academic education is less effective at improving student learning than hybrid and in-person learning[2] (e.g., Fitzpatrick et al. (2020); Rickles et al. (2018); Ahn & McEachin (2017); Bueno (2020); Alpert et al. (2016); Bettinger et al. (2017)). Given its focus on hands-on learning, technical education in an online or hybrid modality presents additional challenges. To the best of our knowledge, no studies have presented causal evidence on the impact of online or hybrid technical education on schooling outcomes.[3] In this paper, we explore the effects of attending *online* technical education during high school on students' schooling outcomes. Specifically, we ask whether attending an *online* technical course while also taking traditional high school courses has an impact on student performance, the probability of repeating a grade, and on graduating and dropping out of high school.

Using administrative data from the state of Pernambuco in Brazil, we exploit a unique discontinuity in the probability of enrolling in online technical education, namely an admission exam score cutoff.[4] This allows an estimation of the causal effect of enrolling in technical education on academic achievement, as well as on rates of dropping out, repeating a grade, and graduation.

We find that students just above the admission exam cutoff perform as well on standardized tests of math and Portuguese as students below the cutoff, which indicates that participation in this form of technical education does not affect learning in core subjects. In addition, our results suggest that students above the cutoff are less likely to drop out of high school, consistent with the findings in Elacqua et al. (2019) on the impact of in-person technical education on student outcomes. Unlike several programs that provide *online* supplementary and/or remedial content of core subjects and have been shown to be effective, especially for mathematics instruction,[5] the *concurrent*

[2]See Elacqua et al. (2020) for a literature review on the evidence comparing in-person, online, and hybrid learning modalities.

[3]Some countries, such as Argentina and Costa Rica, are currently experimenting with or designing online and hybrid technical education courses, but these experiences are too recent or too small to draw conclusions about their effectiveness.

[4]Students who want to enroll in a type of technical education (state concurrent) must take an admission exam. If a school's technical track is oversubscribed (which happens in 70% of the cases in our data), students are then ranked for admission to their chosen technical track and school according to their admission exam score. Seats in the technical track and school are assigned to students according to their rank. This vacancy-assignment method creates a discontinuity in the probability of being offered a seat in a technical track and, consequently, in the probability of enrolling in this modality of technical education.

[5]https://www.povertyactionlab.org/sites/default/files/publication/education-technology-evidence-review.pdf.

online technical education provides solely technical content. Therefore, any positive potential effects on learning would be the result of indirect mechanisms such as motivation or student engagement. In fact, our results are consistent with the evidence from developed countries, which does not show a positive or negative effect of technical education on academic achievement in high school (Dougherty (2018); Neild et al. (2015)). In addition, the retention rate, as in most massive online courses, remains a major challenge in Pernambuco's online technical education, as discussed in detail in Section 1A of the Appendix.

This paper makes two important contributions to the literature. First, it causally identifies, using quasi-experimental methods, the effects on schooling outcomes of attending online technical classes and secondary academic education concurrently. To the best of our knowledge, no causal studies exist on the effects of online technical education during high school on educational outcomes. Second, our results inform ongoing discussions on the potential effects of the expansion of technical education. This is important because several countries are considering, planning, or implementing expanded online/hybrid offerings, both in technical education and more generally,[6] notably due to the COVID-19 pandemic. In particular, the Brazilian government recently enacted an educational reform that aims to expand technical education by offering a complete technical track option during high school.

The remainder of the paper is organized as follows. The next section describes secondary technical education in Brazil and in the state of Pernambuco. Section 3 introduces the data used in the analysis. Section 4 explains the methodology employed to estimate the effect of being able to enroll in technical high school on schooling outcomes, with Section 5 showing the main estimations and heterogeneous analysis. Section 6 presents some robustness checks. Finally, Section 7 concludes.

[6]Some of the countries that have invested heavily in technical education at the high school level in the last two decades are Finland, Iran, the Dominican Republic and several countries in Central Asia. Others, like Hong Kong, Malaysia, Moldova, South Africa, Costa Rica and the United Kingdom, report positive enrollment trends in technical education (UIS, 2020).

2 Background: Secondary Technical Education in Brazil and Pernambuco[7]

Brazil has three public school networks: municipal, state, and federal, in addition to private schools. The states run most secondary schools in the country, with 84% of high school students enrolled in state public schools according to the 2019 School Census.

Public high schools in Brazil provide two modalities of academic education, with a regular or full-time school day. Technical education is also offered in two modalities. The first, *technical integrated high school* (TIHS), provides a mix of in-person academic and technical curricular content in a three-year technical high school. The second modality of public technical education, *concurrent*, offers technical courses to students enrolled in their final two years of secondary school in a separate public technical institution, generally provided by a state school.

There is, furthermore, a third type of secondary-level technical education called *subsequent*. In *subsequent* technical education, the candidate must have finished high school to be eligible. Although this modality serves a large proportion of students,[8] we do not include them in our analysis as we focus exclusively on high school students still attending school and the options available to them.

The structure of high school provision in Pernambuco is the same as that in the rest of Brazil in terms of the academic and technical modalities offered and the predominance of state public schools. Table 1 summarizes the differences between academic high schools and technical education modalities.

Table 5 provides a breakdown of high school enrollment in 2013 and 2018 for Pernambuco and Brazil. Regular academic high school accounts for more than 80% of high school enrollment in the country.[9] TIHS enrollment represents less than 7% of total enrollment (4.1% in 2013 and 6.5% in 2018).[10] Full-time academic high school has in-

[7]Brazil has 27 federal states and one federal district. Pernambuco is the 7th largest state in Brazil in terms of population, with the 10th largest GDP and the 3rd highest education quality index for its secondary public schools. According to the basic education development index of 2017 (*Índice de desenvolvimento da educação básica - IDEB*), Pernambuco ranks third in education quality among all states in Brazil.

[8]In 2018, 894,862 students, accounting for 12% of secondary school enrollment in the country.

[9]In 2013, this figure was 90% of high school students, which decreased to 85% in 2018 (according to Brazil's school census). Academic schools that offer 7 or more hours per day are considered "full time." Less than 3% of high school students in Brazil are enrolled in such schools.

[10]The Ministry of Education aims to triple the number of students enrolled in technical education by

creased over the last decade but remains a small fraction of total secondary education enrollment, increasing from 0.5% in 2013 to 2.5% in 2018. Lastly, *concurrent* technical education went from 3.7% of high school enrollment in 2013 to 4.6% in 2018 in Brazil and from 1.8% to 3.1% in Pernambuco.

2.1 Online Technical Education in Pernambuco

The state school network of Pernambuco is the only one in Brazil in which the *concurrent* technical courses are offered exclusively online. This modality is available to students attending their 2^{nd} and 3^{rd} years of high school (*"concurrent students"*) and to high school graduates (*"subsequent students"*).[11]

The online technical network in Pernambuco gradually expanded from 7,900 seats in 2012 to approximately 20,475 in 2017, distributed into ten different technical tracks within three main career areas: STEM (science, technology, engineering, and math), humanities, and general services.[12]

The online technical education curriculum requires 12 hours of online activities/classes along with 3 hours of in-person activities *per week* in accredited supportive poles, for a total time commitment of 15 hours per week (see Table 6). Depending on the chosen track, the total course load ranges from 800 to 1200 hours. An average course load is normally composed of three modules of approximately 300 hours, which take 60 weeks or 15 months to complete, with a total of around 720 hours of online classes/activities and 180 hours of in-person classes/activities. Teachers conduct "formative" evaluations every week and an assessment at the end of each module. Students are required to attend a minimum 75% of in-person and 75% of online activities to achieve a minimum score at the end of each module and to graduate from the course.

Supportive poles are spaces equipped with tutors and laboratories for the in-person activities and reinforce the online technical classes. They provide essential infrastruc-

2024 (Law No. 13.005/2014)

[11] *Concurrent* technical courses provided by the state network compete directly with in-person and online technical courses offered by *Sistema-S*, a nonprofit network composed of private entities that offers a range of public interest services in Brazil. A high school student attending their 2^{nd} or 3^{rd} year may apply to a competitive federal scholarship (*Bolsa Formação*) in order to attend a free technical course in *Sistema-S*.

[12] The technical tracks are Business Administration, Librarian, Interior Design, System Developer, Logistics, Multimedia Teaching, Human Resources, School Secretary, Workplace Safety and Restaurant and Bar Attendant. Dutra P. (2017) does not report the criteria used by the Secretary of Education of Pernambuco to define the number of seats each year.

ture, internet access, and technical and pedagogical assistance to all students, helping to strengthen a sense of belonging and aiding those who cannot connect from their homes.

Since 2012, the State Secretary of Education of Pernambuco has invested substantially in the online technical education network, increasing the number of supportive poles from 18 in 2012 to 91 in 2017. Figure 3 displays this enlargement from a spatial perspective, given that a key goal of the state government was to distribute the poles throughout Pernambuco. Supportive poles are mostly based in state public schools. By 2017, 53% were located in full time or semi-full time academic schools, 39% in state technical schools (Escolas Técnicas Estaduais; ETE), 7% in other state schools, and 1% in municipal schools (Dutra P. (2017)).

Finally, Pernambuco's *concurrent* online technical education offers a particular combination of characteristics, making in an interesting case for examining the schooling effects of online technical education. First, there are very few other massive programs that provide online technical education during high school. Second, the popularity of the courses and the government's emphasis on technical education has led to a high number of oversubscribed online technical tracks, yielding a sizable number of observations that allow us to obtain precise estimates. Third, online technical education has followed a centralized and consolidated admission process since 2012, which has remained stable over time.[13] Fourth, Pernambuco gathers rich student data, which allows us to follow pupils over time and merge different education databases. The state has also applied an annual standardized test since 2007 to all students in the 5^{th} and 9^{th} grades and in the 3^{rd} year of high school.

2.2 Admission Process

By 2012, technical education had become a flagship educational policy in Pernambuco, attracting an increasing number of students and recent high school graduates.[14] As a result, most online tracks were oversubscribed, leading the State Secretary of Education to institute an admission process to select candidates. All online technical tracks therefore now use an admission process to assign available seats to applicants, which

[13] Although most states in Brazil offer the *concurrent* modality, this track is not oversubscribed in all places. Furthermore, online modalities are often not available. Very few states that do online courses also have a formal admission process for technical education (e.g., São Paulo, Maranhão and Piauí).

[14] This has similarly been the case in other states in Brazil as well as in other countries such as the Dominican Republic, Costa Rica, and Argentina.

takes place once per semester and twice per year. The applicant chooses the supportive pole and the track. The one-hour admission exam includes 20 questions on Portuguese and mathematics. Students must score higher than 0 points on both the math and Portuguese sections of the exam. This criterion is binding, as 1,288 students in our sample do not pass the admission exam. Seats are assigned at each online technical track and supportive pole to the students with the highest scores until all vacancies in the track and supportive pole are filled. *Concurrent* and *subsequent* candidates apply and compete jointly for the same vacancies and attend the same online classes.

Since the admission exams for the *concurrent* and subsequent modalities are conducted jointly (and the administrative data does not classify candidates by modality), we use the student's age at the time of the admission exam as a proxy for the modality to which they are applying. The subsequent modality is only available for students who have already completed high school, meaning that students 18 years of age or younger are most likely applying as *concurrent* students. We use this criterion to identify students in each modality and separate out our sample of interest: the *concurrent* students.

About 5.3% of students in their 2^{nd} year of high school participated in the state admission exam for the *concurrent* modality between 2012 and 2017. In our sample, 55.6% of applicants took the exam in their 2^{nd} year of high school, 42.8% in their 3^{rd} year, and 16% took the exam 2 years or more after their 2^{nd} year of high school, indicating that the student may have applied to an online course after graduating from high school, i.e., as a subsequent student.[15] These percentages also reflect the fact that in Brazil, a student can graduate from high school before they turn 18 years old. We consequently reclassified all students who took the exam at least 2 years after their 2^{nd} year of high school, excluding them from all further analyses.

Note that an important rule in the admission process is that 75% of all seats are reserved for students who were enrolled in public schools in the 9^{th} grade. In practice, this means that there are two effective score cutoffs in each combination of online track and supportive pole. Candidates who score below the exam cutoff may still enroll in technical education during high school. Specifically, such students could potentially enroll in another online track that is not oversubscribed, or else opt for a technical course in a federal high school or even in the private sector.

[15] These percentages add up to more than 1 because students can apply to different tracks in different years. From 2012 to 2017, 11.7% of candidates in our sample that applied in their 2^{nd} year also applied in their 3^{rd} year.

3 Data and Descriptive Analysis

3.1 Data sources

We use four data sources: a) school censuses for the state, municipal, federal, and private school networks, which together represent the universe of students enrolled in high school in Pernambuco in a given school year; b) test scores from a standardized assessment; c) the applicant lists for state online technical courses; and d) administrative data for the students enrolled in *concurrent* online technical courses. With this information, we construct a rich panel dataset for Pernambuco from 2012 to 2018 in which we follow six cohorts of students in their last two years of high school (2^{nd} and 3^{rd} years).

The school census is an annual national survey that provides information on all students, teachers, and schools in each state. Using a unique student identifier, we created a panel of students in order to follow pupils' schooling trajectories.[16] These trajectories provide information on students' academic high school peers, teachers, and schools. For example, whether the school has a laboratory and other infrastructure, and teachers' level of education and experience. We use the school census from between students' 2^{nd} and 3^{rd} year of high school to identify whether the student repeated their 2^{nd} year of high school or dropped out of the Pernambuco education system.[17] We also use a follow-up survey of the school census (*Censo Rendimento*) to identify whether the 3^{rd} year students go on to graduate from high school.

The second database used is Pernambuco's statewide student assessment (*Sistema de Avaliação Educacional de Pernambuco*, SAEPE). SAEPE is a standardized exam that students take in the 5^{th} and 9^{th} grades and in the 3^{rd} year of high school.[18] It evaluates student achievement in language and mathematics in all state and municipal schools. The database also contains students' full names and a complete set of demographic and socioeconomic variables. We observe SAEPE information for the six cohorts of students in the 9^{th} grade between 2010 and 2015 and use their SAEPE score in the 3^{rd} year of high school between 2013 and 2018.

[16] In addition to the unique student identifier (*INEP ID*), the census includes the student's full name, date of birth and, for a minority of students, Brazilian ID number (CPF).

[17] We are unable to identify whether a student drops out completely from the education system or moves to another state. In both cases, the student would disappear from the sample.

[18] In Brazil, primary education lasts from grade 1 to 9, high school is counted as 1st-3d year.

Our third database consists of the applicant lists from online technical education admission processes from the 2^{nd} semester of 2012 to the 2^{nd} semester of 2017.[19] The applicant-list data provides information on the *supportive pole* and track the student applied to, the number of vacancies per track and supportive pole, the candidate's score on the admission exam, the student's rank within the online technical track and supportive pole of choice, whether the candidate was offered a seat in the online technical track, whether the student enrolled in the online technical track and supportive pole, and the student's background information.[20]

Finally, the fourth database reports student enrollment/completion status of the *concurrent* online technical courses. This database allows us to better identify the demographic and socioeconomic characteristics most associated with students dropping out or graduating from an online technical course in Pernambuco.

We start with 674,151 students enrolled in the second year of high school between 2012 and 2017 and 48,038 students who applied to a concurrent online technical course during these years, including 11 admission processes. We matched around 75% of these students to our other databases, for a total of 35,895 students.[21] Only 25,664 students in the original concurrent applicant lists have admission exam scores. Of these, we were able to match 71% of them (18,323 students) to our other databases. Our final sample excludes students without scores, students who were enrolled in private schools in 9^{th} grade, students who applied to track and supportive pole combinations that were not oversubscribed, and students identified as applying to the subsequent modality.[22] Our final sample consists of 16,928 students, as shown in Column 3 of

[19] The year of the admission process corresponds to the year in which students enroll in *concurrent* technical education. Students take the state admission exam in the semester they intend to enroll.

[20] Background information includes the student's full name, date of birth, gender, and a variable equal to one if the student was enrolled in a public school in 9^{th} grade and is thus eligible for a "public school quota" vacancy.

[21] Some students were not matched due to differences in the names reported in the two databases. For example, the mother's surname being reported in one database but not in the other. These errors are more likely for students who were not previously in the public education system, as their names were entered in the system for the first time with the admission exam. While this could imply that our results are not perfectly generalizable to all students in Pernambuco, it does not pose a threat to the internal validity of our empirical strategy.

[22] Students in private schools in 9^{th} grade do not have baseline socioeconomic and achievement information, since private schools are not required to administer the state standardized test. We exclude them as it is impossible to check the balance of the student characteristics across the track/school admission cutoff. Meanwhile, students who apply to under-subscribed concurrent track/supportive pole combinations do not have an admission exam cutoff, since they are accepted to the track/supportive pole combination. Given our empirical strategy, these students cannot be used in our analysis, nor can the 4,400 identified subsequent students.

Table 2.

3.2 Descriptive Statistics

Table 2 shows descriptive statistics for the population in Pernambuco's school census in the second year of high school from 2012 to 2017, as well as for the subpopulation who applied to *concurrent* technical education and for the final sample we will use in our analysis.[23] Column (1) presents descriptive statistics for the universe of students in the 2^{nd} year of high school, and Column (2) contains statistics for the subpopulation from Column (1) matched to the *online technical courses* applicant lists. Secondary school students are usually unmatched because they did not apply to a concurrent technical course, though this can occasionally be due to other matching challenges (e.g., names do not match), as explained in Subsection 3.1. Column (3) displays descriptive statistics for the subpopulation of Column (2) that excludes students who attended private schools in the 9^{th} grade, as well as students without admission exam scores, students who apply to under-subscribed tracks, and students identified as applicants to the subsequent modality.

In summary, from Column (1) to (2) the population was restricted to the subgroup of students who applied to a *concurrent* online course, whereas from (2) to (3), due to data identification and/or availability, we narrow the observations to our final sample. Panel A shows student characteristics, Panel B reports enrollment in technical and academic schools, Panel C presents the characteristics of the high schools attended by students in the subsample, and Panel D shows summary statistics for our outcomes of interest.

Comparing the student characteristics in Panel A, we see that the students who apply to *concurrent* online technical education (Column (2)) and those in our sample (Column (3)) are less likely to be male, more likely to be younger, more likely to come from a slightly higher socioeconomic status – i.e., they have slightly more educated mothers and are more likely to live on a paved street[24] – and had higher initial math and Por-

[23] Since we obtain standardized scores and socioeconomic information from the SAEPE standardized test database, information on these variables is conditional on having participated in the exam.

[24] In Brazil, whether a street is paved or not is highly correlated with the neighborhood's level of economic development. This variable is thus positively correlated with the socioeconomic status of the student's family. Other variables in our data that could indicate socioeconomic status, other than parents' level of education, are whether the student's family has a computer, whether the family has a cleaning person, and whether the student uses the internet frequently. However, these variables are not available for all years.

tuguese test scores in the 9^{th} grade relative to the population of students in the 2^{nd} year of high school (Column (1)). Other than the fact that *concurrent* online technical education involves a selective admission process, which may deter poorer students from applying, another possible explanation of the higher socioeconomic characteristics of students who apply to the concurrent courses may be that schools offering concurrent education have a longer school day. Students from more vulnerable backgrounds may need to work after school and, thus, would not have the option of enrolling in this type of program. These differences remain when comparing the population to the sample included in our analysis. Furthermore, students who believe they have a chance at getting a seat may be more likely to apply, which might explain why the students in Columns 2 and 3 have higher baseline test scores.

Panel B shows that, even though very small proportion of the overall population (Column (1)) enrolls in concurrent technical education (0.89%), as expected, a significantly larger proportion within the group of students who participate in the admission process (Column (2)) and within our sample (Column (3)) are enrolled in an online technical course. Less than one fifth of students in these groups enroll in concurrent education, which partly reflects the level of competition students face when applying to this type of education.

With respect to their academic high school choices, we see that while academic regular high schools account for the largest group of students, those who participate in the admission process (Column (2)) and in our sample (Column (3)) are more likely to be enrolled in full-time and semi full-time academic schools (compared to the population in Column (1)). Finally, private and federal high schools and technical integrated high schools (TIHS) are the least frequent academic enrollment choices. There are two main reasons why a large share of *concurrent* students are enrolled in full-time or semi full-time schools: 53% of the supportive poles for the in-person activities are located in such schools and, in 2016, the state of Pernambuco launched a program aimed at fostering innovation and entrepreneurship skills among high school students enrolled in full-time schools, including incentives to take technical courses offered online.

Panel C compares the characteristics of academic schools in which students from the three groups are enrolled. Summary statistics show that students who apply to *concurrent* technical education (Column (2)) and those in our sample (Column (3)) attend slightly smaller academic schools with higher per-pupil spending and higher teacher salaries. These schools are more likely to have laboratories and to employ more experienced teachers. These differences in school characteristics likely reflect the fact that students in Columns (2) and (3) are more likely to attend full-time and semi full-time academic high schools.[25]

Finally, Panel D compares our outcomes of interest for the three groups of students. Students in Column (3), in our sample, appear to have slightly better outcomes than the population of students who apply to *concurrent* online technical education (Column (2)) who, in turn, have better outcomes than those of the entire population of students in the 2^{nd} year of high school (Column (1)).

A student is deemed to repeat a year when their INEP ID (unique student identifier) appears at least twice or more in different school census years as attending the same high school year (specifically, 2^{nd} or 3^{rd} year of high school).[26] Students in our sample (Column 3) are significantly less likely to repeat a year (5% on average) than the full population of students in high school, of whom 10.9% of students repeat their 2^{nd} or 3^{rd} year of high school.

Math and Portuguese test scores in the 3^{rd} year are also higher in our sample compared to the population of students in the 2^{nd} year of high school (Column (1)). This is at least partially explained by the better baseline results in the 9^{th} grade, as reported in Panel A.

We estimate the dropout rate by examining whether a student's INEP ID disappears from Pernambuco's education census before the last year of our sample, 2018, when the last grade reported is lower than 3^{rd} year of high school. More specifically, we estimate whether a student drops out between the 2^{nd} and 3^{rd} years of high school. This dropout estimate could potentially be biased if student mobility across Brazilian states is high or if reporting errors are common. Fortunately, in Brazil, most students complete high school in their own state. In addition, "disappearing" from the educational census could be a result of an error in a student's INEP ID. Even though the school census must keep the same number for every student over time, a student may have her INEP ID switched due to administrative error. The dropout rate of 21% in Column (1) thus overestimates the actual percentage due to those data errors, as for example, when we are not able to follow up the same student (same INEP ID) from one school year to another even though the student continues in the school system. As an administrative error, however, it is independent of our identification strategy and should not be a

[25] Full-time and semi full-time schools in Brazil tend to attract teachers with better credentials and have better infrastructure.

[26] For all analyses, we only deem a student to have repeated a year if this occurs in the same year as when she applied to the *concurrent* online course admission exam or in the following year.

concern. Again, as we can see in Columns (2) and (3), the sample of matched students and our final sample have a significantly lower dropout rate than the entire population of students in their 2^{nd} year of high school.

4 Empirical Strategy

In Pernambuco, students who want to enroll in *concurrent* technical education must take an admission exam. Their score determines whether they can enroll in the track and supportive pole to which they applied. First, they must score higher than 0 points on both the math and Portuguese sections of the exam. Second, if the track and supportive pole combination is oversubscribed, seats are allocated according to student rank. This endogenously determines an admission exam score cutoff, with those students scoring below the cutoff unable to enroll in the program.

In this context, we use the discontinuity in the eligibility to enroll in a *concurrent* online course due to the admission exam cutoff in a regression discontinuity design (RDD) to analyze the effect of online technical education on standardized test scores and other schooling outcomes.

Out of 3,401 pole-year-semester-technical track combination that we observe for the admission processes to *concurrent* technical education from 2012 to 2017, 69% (2,355) have binding admission exam score cutoffs. This is important because the exam score is not always binding. In some cases, the demand for the program is low and students can enroll by simply scoring more than 0 on both sections of the exam. Intuitively, if we assume that admission test scores change smoothly with student characteristics,[27]

[27] Table 7 shows that there are no significant differences in student characteristics among those with scores on either side of the cutoff. Students on either side of the cutoff are just as likely to participate in the 9^{th} grade assessment, with no differences in their performance (main model, Column 12). However, we can observe that students above the cutoff are less likely to have mothers that completed high school (about 4 percentage points) and less likely to have mothers that attended university (about 4 percentage points at 10% significance). We interpret the decline in mother's level of education as an indication that students above the cutoff may be from a slightly lower socioeconomic background than students below the cutoff. Any bias caused by this difference would reduce our estimates, as test scores are highly correlated with socioeconomic status (Mizala et al. (2007)). In this case, our estimates can be interpreted as a lower bound for the effect of *concurrent* technical education. Finally, Column 4 indicates that there are more women above the cutoff. This might generate an upward bias in the estimations if women have better academic results. We ran linear regressions in order to estimate the correlation between the main outcomes and gender. Consistent with the literature, men are more prone to drop out and have higher math scores, while women have higher language scores. Regarding changes in density around the cutoff, there is no risk of manipulation of scores in this case since the cutoff is unknown ex-ante,

the discontinuous change in the probability of enrolling in *concurrent* technical education due to the score cutoff allows us to identify the causal effects of participating in this educational program. We can then use students in a small neighborhood below the cutoff as an adequate control group for students just above the score cutoff. Any difference in their educational outcomes can be attributed to the fact that they were more or less likely to access *concurrent* online technical education. This strategy allows us to address self-selection into state concurrent education since, for students close to the admission test score cutoff, scoring above or below the cutoff is almost random.

As shown in Figure 1, while a significant proportion of students with test scores allowing them to enroll in *concurrent* programs take the opportunity, not all do so. Additionally, even though candidates below the cutoff are not allowed to enroll in the (oversubscribed) *concurrent* track/school combination of choice, a few students with scores below the cutoff did manage to enroll in a *concurrent* track.[28] Since the probability of enrolling in a *concurrent* program changes by less than 1 over the admission exam score cutoff, the RDD in this analysis is a fuzzy RDD. We use the traditional RDD estimation, which identifies the average effect of "intent to treat," which means the effect of having the opportunity to be treated.

Therefore, we estimate the effect of having the opportunity to enroll in a concurrent program on educational outcomes using the following regression:

$$y_{ispt} = \alpha + \beta \mathbf{1}[s_{itr} - \overline{s_{sptr}} \geq 0] + f(s_{itr} - \overline{s_{sptr}}) + \theta_s + \lambda_p + \rho_t + \delta_r + \varepsilon_{itr} \quad (1)$$

where y_{ispt} is the educational outcome of interest for student i applying to the technical track p in state school s and in year t and semester r, $\mathbf{1}\left\{s_{itr} - \overline{s_{sptr}} \geq 0\right\}$ is an indicator equal to one if student i has an admission test score equal to or above the admission score cutoff $(\overline{s_{sptr}})$ for the concurrent track and school he applied to in year t and semester r, β is the parameter of interest that captures the effect of having the opportunity to enroll in a concurrent program, $f\left(s_{itr} - \overline{s_{sptr}}\right)$ is a flexible parametric specification that includes higher-order polynomials of the difference between the admission exam score and score cutoff and can vary on either side of the enrollment cutoff, and θ_s, γ_p, ρ_t, and δ_r are the fixed effect of the school, concurrent track, year

given that it depends on the number of seats and applicants.

[28] About 9% of students below the admission exam cutoff enrolled in a concurrent course. Unfortunately, we cannot observe in our data whether these candidates enrolled in the school/track combination they had applied to or in a different concurrent track/school combination. Anecdotal evidence suggests that these students may enroll in under-subscribed concurrent tracks in their initial school of choice.

and semester, respectively, of the student's application. In order to maximize statistical power, we pool data across *concurrent* tracks (p), schools (s), application years (t), and semesters (r) and use the distance of the student admission exam score to the relevant test score cutoff $(s_{itr} - \overline{s_{sptr}})$, as in Pop-Eleches & Urquiola (2013).

Our main specification uses bandwidth 3,[29] which is roughly twice the bandwidth suggested by the procedure in Imbens & Kalyanaraman (2012) – henceforth, IK – for the main outcomes.[30]

5 Results

In this section, we present the results from our analysis in three parts. First, we show how *concurrent* online technical education enrollment decisions change for students who obtain scores above the cutoff on the admission exam. Second, we examine the impact of being able to enroll in an online technical course on schooling outcomes. Finally, we explore potential heterogeneous effects behind our main results, focusing on the sub-populations of students most associated with higher completion rates of online technical education in an attempt to address the conservative ITT estimations.

5.1 Changes in Enrollment Decisions

The probability of enrolling in *concurrent* online technical education changes with the admission exam score and jumps discontinuously at the admission score cutoff. Specifically, it increases by 40 percentage points for students with admission scores above the cutoff. This is depicted in Figure 1, where each dot shows the proportion of students who enroll in a concurrent program relative to the distance of their admission exam score to the cutoff. Table 3 reports the same result: the probability of enrolling in a *concurrent* online technical course increases by about 40 percentage points for students

[29] Results using other bandwidths, including the IK bandwidth, are shown in our result tables and in the Robustness section. Results using alternative bandwidths are consistent with our main results.

[30] We used twice the IK bandwidth because the IK bandwidth is extremely narrow, although results using the IK bandwidth are always reported. Since our running variable is discrete, we cannot use alternative methods to select the bandwidth, such as the method described in Calonico et al. (2015). Additionally, we include a linear spline on admission exam score distance to cutoff interacted with a dummy indicating whether the student falls to the left or the right of the cutoff.

above the admission exam score cutoff (Column 2).[31] Such a narrow bandwidth does not allow us to include a flexible linear spline that could control for possible trends in the outcome. Therefore, Column (2) uses bandwidth 3, which is about double the IK bandwidth for the main outcomes of this paper, and a flexible linear spline on the distance of the admission exam score to the cutoff. Column (2) is our preferred specification and reports a slightly lower first stage (from 0.47 to 0.40).

Students at the cutoff are also more likely to enroll in a *concurrent* online technical track than those who score below the cutoff. In addition, to ensure that potential results on outcomes are not explained by differences in the academic high schools attended by students, Tables 8 and 9 explore whether there are differences in the schools attended by students above and below the admission exam cutoff. The results show that there are no significant differences in the type and characteristics of schools attended by students on either side of the cutoff and that this is not a potential threat of bias in this context.[32]

5.2 Impact on Schooling Outcomes

All estimations report intent to treat effects (ITT), which means that they show the effect on schooling outcomes of having the opportunity to enroll in *concurrent* online technical education administered by the state of Pernambuco. We report ITT as a more conservative method instead of "average treatment on the treated" (ATT) because ATT results would be biased if taking the treatment (i.e., enrolling in an online course) is not random conditional on being assigned to the treatment (i.e., scoring above the cutoff). This could happen, for example, if a higher proportion of students who would benefit the most from *concurrent* online technical education enroll after obtaining admission scores above the cutoff. In addition, our attrition analysis suggests that attrited students are unlikely to bias our results, as described in Section 1B of the Appendix.

[31] The optimal IK bandwidth for the concurrent education enrollment outcome is 0.8755. Since our running variable is discrete, this implies that this bandwidth only includes observations at the cutoff and is, thus, replaced with a bandwidth of 1, which includes one admission exam score bin at either side of the cutoff. In some cases, a few students with scores directly at the cutoff will be admitted into the technical track they applied to, while other students at the cutoff for the same concurrent track/school combination will not be admitted. This makes our first stage fuzzier, since it increases the proportion of treated students in the control group and may bias our estimates downward. To measure the magnitude of the bias, we repeat our main outcome regressions excluding students at the cutoff in the Robustness section.

[32] Teachers who teach math classes but do not have a college degree in math or a related field is the only variable with a statistical difference, but with a coefficient of a small magnitude.

The estimates of the effect of attending *concurrent* online technical education on academic outcomes are summarized in Table 4 and Figure 2. We explore five academic outcomes: the probability of repeating at least one year during high school after applying to *concurrent* programs (Columns (1) - (2) of Table 4 and Panel A in Figure 2), the probability of dropping out between the 2^{nd} and 3^{rd} years of high school after applying to *concurrent* programs (Columns (3) - (4) and Panel B), math and Portuguese SAEPE standardized test scores in the 3^{rd} year of high school (Columns (5) - (6) and Panel C, and Columns (7) - (8) and Panel D, respectively) and the probability of graduating from high school (Columns (9)-(10)).

Our estimates suggest that students who have access to *concurrent* online technical education in their 2^{nd} year of high school are, on average, 3 percentage points less likely to drop out between the 2^{nd} and 3^{rd} years of high school.[33] The result is also consistent with the literature (Elacqua et al. (2019); Kemple & Snipes (2000); Hall (2012)). The difference here is that, to the best of our knowledge, ours is the first study to estimate the educational effects of an online high school technical program.

In terms of the probability of repeating their 2^{nd} or 3^{rd} year of high school, students just above the admission cutoff are, on average, as likely as students just below the cutoff to repeat a grade. In Columns (5) to (8) of Table 4 and Panels C and D of Figure 2, we explore the impact of online technical high school on student learning, which shows that math and Portuguese SAEPE scores increase with admission exam scores. Once we take into account the existing positive relationship between these variables using a linear spline in Columns (6) and (8), the coefficients are still high and positive, but no longer statistically significant. Finally, there is no significant difference in the probability of graduating from high school between students just above and below the admission exam cutoff (Columns (9) and (10)).

Overall, the results suggest that the opportunity to enroll in an online technical course during the last two years of high school reduces the probability of dropping out of high school without negatively impacting achievement in core subject areas. We should bear in mind, though, that the students who dropped out of high school have, on

[33] Although, as discussed in Subsection 3.2, we cannot differentiate whether a student drops out of the education system or moves out of the state, since in both cases students would disappear from the state administrative records. In this context, dropping out is defined as disappearing from Pernambuco's educational census before the 3^{rd} year of high school and before the last year of our sample, 2018. Also, in our final sample, we were unable to accurately identify whether students dropped out or whether there were data inconsistencies for 8% of the observations. For these 8% of inconclusive observations, we assigned a value of "missing" for the dropout variable. The probability of being in this group is not statistically different between students just above and below the admission exam cutoff.

average, a lower socioeconomic background, which is associated with lower academic performance. Since students below the cutoff are more likely to drop out, we may be downward biasing our learning estimations. Second, as Table 7 reports, female students are in higher proportion above the cutoff. Empirical evidence has shown that in some countries male students are more likely than female students to drop out of high school (See for example Goldin et al. (2006)). If that was the case, the gender imbalance may upwardly bias our dropout estimation.[34] Third, given the low completion rate of *concurrent* online technical courses (see Section 1A in Appendix), the ITT estimations are, most likely, too conservative, and more susceptible to type II errors. Therefore, we take a closer look at student subsamples that are more likely to complete the online program in an attempt to tease out possible "hidden" effects of our main results, as reported in Section 5.3 on the heterogeneity analysis.

5.3 Heterogeneity Analysis

Our main results indicate that students who have the opportunity to enroll in a *concurrent* online technical education are less likely to drop out between their 2nd and 3rd years of high school, while their performance in core subjects is unaffected. On the other hand, as discussed in Section 1A in the Appendix, *concurrent* online technical courses have a low retention rate. Since our estimates include everyone who had the chance to enroll in the concurrent courses, who did not necessarily complete them, our analysis is probably too conservative, and might be neglecting some potential effects.

Using administrative data, we were able to identify the student characteristics most associated with the completion rates of *concurrent* online technical education (see Table 19). Specifically, these are **(i)** students with lower socioeconomic status (SES), **(ii)** students with higher baseline test scores in Portuguese on the 9th grade SAEPE test,[35] **(iii)** students attending a supportive pole in a full-time or *semi* full-time school. In this context, we repeated the same main outcome regressions restricting the analysis to these three subsamples in an attempt to explore the heterogeneous effects of *concurrent* online technical courses on schooling outcomes.

When we take a closer look at subsample **(ii)**, we find no significant differences in any

[34]Since the construction of our dropout variable disregards 8% of the observations containing inconsistencies and administrative errors that hinder a more accurate identification, we also run the balance check variables disregarding this same 8% and observe only a marginal gender imbalance in favor of women (p-value at 10%) around the cutoff. Results are available upon request.

[35]Scores are considered above average, if they are above the global average.

of the outcomes estimated. Thus, we focus the analysis on subsamples **(i)** and **(iii)**.

Table 10 reports our main outcome results for students from lower SES backgrounds who had the opportunity to enroll in *concurrent* online technical education. In general, the coefficients show no significant effects on any of the schooling outcomes, with the exception of Portuguese test scores in the 3^{rd} year of high school. The increase in Portuguese test scores is 0.14 standard deviations in our preferred specification (Column 8). This result should be taken with caution, however, as students above the cutoff in this sample had higher Portuguese test scores in the 9^{th} grade than those below the cutoff (see Table 16), which may increase the magnitude of the coefficient and drive the result.

Finally, we estimate the regressions of our main schooling outcomes for subsample **(iii)**, i.e., *concurrent* students who attended a supportive pole based in a full time or *semi*-full time school. It is worth recalling that supportive poles are spaces provided by the state of Pernambuco for the online technical students to attend each week for at least 3 hours of in-person activities. The supportive poles are mostly located in full-time or *semi* full-time public academic high schools or technical high schools. In other words, students in concurrent programs have to commute every week, for the entire duration of the course, to attend the in-person activities in order to graduate from the online technical course.[36]

Table 11 summarizes the results. The estimations show no difference in the probability of repeating a year in high school (2^{nd} or 3^{rd} years). Also, there is no effect on math and Portuguese test scores. The only outcome that shows a strong and significant effect is the probability of dropping out of high school. Column (4), our preferred specification, shows that students above the cutoff are 5.4 percentage points less likely to drop out during their 2^{nd} year of high school - a coefficient two percentage points larger than the result of the main sample. Furthermore, balance check estimations in Table 17 show that this subsample has no gender imbalance and students' mothers are less educated above the cutoff.

This finding suggests that attending a supportive pole in a full time school magnifies the impact of *concurrent* online technical courses on reducing the probability of dropping out of high school. Students who attend these supportive poles are more likely

[36]From 2018 on, the State Secretary of Education of Pernambuco universalized online technical education in the state, which led to two main changes: there is no longer an admission test (prospective students simply need to submit a few documents to apply and enroll in the course of interest) and attending the supportive poles is no longer mandatory (although the supportive pole network remains).

to graduate from the online technical course, which, in turn, should influence the students' decision to continue studying in high school (since, if they do not complete high school, they are not allowed to graduate from the *concurrent* technical course).

Why, then, do students attending a supportive pole based in a full-time school graduate at a higher rate from their *concurrent* online technical course compared to their peers? As mentioned above, students who attend high school in an academic full-time school are more likely to attend a supportive pole in the same full time school – especially because the State Secretary of Education of Pernambuco launched a program in 2016 to incentivize secondary students to apply to *concurrent* online courses with in-person poles in the same full-time school. Assuming this hypothesis is valid, having technical and academic modalities in the same place provides a *organizational ease* that may encourage the student to continue studying in the online technical course and remain in the academic high school. This assumption has already been shown to be effective at reducing dropout rates in other technical education modalities (Elacqua et al. (2019); Dougherty (2018)). On the other hand, as mentioned in Section 1A in the Appendix, students who do not attend in-person activities in full-time schools must do so in a technical high school. Most of these schools are located outside the urban core due to their size and infrastructure requirements, potentially posing logistical difficulties for the students.

6 Robustness Checks

Overall, our results show a reduction in the probability of dropping out between the 2^{nd} and 3^{rd} years of high school for students above the admission exam cutoff. In this section, we thus conduct several placebo exercises and explore alternative samples to verify the robustness of our results, particularly as this concerns the probability of dropping out of high school. We start with a replication of our main analysis, excluding students at the cutoff of the admission exam score. Specifically, the fact that some students at the cutoff are accepted to a concurrent online technical course may lead to a "contamination" of our "control group" by students that are actually "treated" and may underestimate the results. Removing these students from our sample, we expect students' outcomes to increase in magnitude. As Column (4) of Table 12 shows, students above the cutoff are now 3.93 p.p. less likely to drop out, compared to 3.11 p.p. in the main estimation (Table 4). The other outcomes increase in magnitude, but with no changes in the coefficient's statistical significance.

Secondly, we estimate our main outcomes including students who were enrolled in private schools in the 9^{th} grade. We preferred not to include this group in our main analysis since we do not have information on the previous socioeconomic conditions of students that did not participate in the SAEPE standardized test in the 9^{th} grade, making it impossible to carry out any socioeconomic balance checks. Table 13 shows that all of the results remain consistent with our main estimations. Columns (1) and (2) of Table 14 report similar results to our main estimation as we test alternative bandwidths (double the IK bandwidth and 5),[37] which is consistent with the dropout rates found using our our main specification.

Finally, Columns (3) and (4) of Table 14 suggest that when we estimate the probability of dropping out using fake cutoff points, the effect disappears. In other words, this indicates that we are not estimating effects “by chance” and that the discontinuity we explore appears to be well-identified.

7 Conclusion

The Covid-19 pandemic has sparked persistent policy debate over the potential impact of technology-mediated education delivery on student outcomes. However, the literature from which policymakers can draw to inform the design of policies and programs is relatively scant. In addition, rigorous empirical evidence on the effects of technical education in high school is limited and mostly available for developed countries with very different educational systems and methods of providing technical education. This study provides evidence in both of these areas by exploiting an oversubscribed large-scale *online* technical education program during high school in the state of Pernambuco, Brazil (known as *concurrent* technical education) to estimate the effect of having the chance to enroll in this type of technical education on student schooling outcomes.

Concurrent online technical education is a technical education modality offered to students in their 2^{nd} or 3^{rd} year of high school in addition to the academic high school curriculum. The program is free and aims to provide technical training to secondary students to better prepare them for the labor market. It consists of online technical courses that range from 800 to 1200 total hours with mandatory in-person activities. To complete the program and obtain a certificate, students must graduate high school

[37] A bandwidth of 5 was chosen because of the observed trend in Figure 2 Panel B. There seems to be an upward trend in dropout rates as we move away from our main estimation bandwidths (IK and 3).

and reach a minimum of 75% participation.

Overall, the results show that students just above the admission exam cutoff are less likely to drop out of high school, while their performance in Portuguese and math standardized exams remain unchanged relative to students below the cutoff. There is no effect on the probability of repeating a high school year. Our estimations are consistent with Elacqua et al. (2019) and provide suggestive evidence that students can take advantage of the benefits of a technical training (e.g., work readiness) while they are still in high school without hindering performance in core subjects, such as math and language.

It is worth highlighting that online technical education retention rates continue to represent a significant challenge. Only around 12% of enrolled students conclude the course. As a result, our high school dropout estimations are likely conservative. Our results also indicate that delivering technical education in an online format may be plausible as a cost-effective option in the current context where various countries are investing in online/hybrid education models and expanding technical education. More research- and policy-oriented investigations (e.g., Arias et al. (2020)) are needed to better understand the challenges and benefits brought by hybrid/online instruction and equip policymakers and governments to enact potential long-lasting or even permanent changes in the education delivery format.

References

Ahn, J. & McEachin, A. (2017). Student enrollment patterns and achievement in ohio's online charter schools. *Educational Researcher*, 46(1), 44–57.

Alpert, W. T., Couch, K. A., & Harmon, O. R. (2016). A randomized assessment of online learning. *American Economic Review*, 106(5), 378–82.

Arias, E., Brechner, M., Pérez, M., & Vásquez, M. (2020). De la educación a distancia a la híbrida: 4 elementos clave para hacerla realidad. *Banco Interamericano de Desarrollo (BID)*, Hablemos de Política Educativa, No. 2.

Bank, W. (2017). *World development report 2018: Learning to realize education's promise*. The World Bank.

Bettinger, E. P., Fox, L., Loeb, S., & Taylor, E. S. (2017). Virtual classrooms: How online college courses affect student success. *American Economic Review*, 107(9), 2855–75.

Brunner, E., Dougherty, S., & Ross, S. L. (2021). *The effects of career and technical education: Evidence from the Connecticut Technical High School System*. Technical report, National Bureau of Economic Research.

Bueno, C. (2020). Bricks and mortar vs. computers and modems: The impacts of enrollment in k-12 virtual schools. *Computers and Modems: The Impacts of Enrollment in K-12 Virtual Schools (July 3, 2020)*.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2015). Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association*, 110(512), 1753–1769.

Carbonaro, W. (2005). Tracking, students' effort, and academic achievement. *Sociology of Education*, 78(1), 27–49.

Chun, H.-K., Comyn, P., & Moreno da Fonseca, P. (2021). Skills development in the time of covid-19: taking stock of the initial responses in technical and vocational education and training.

Dougherty, S. M. (2016). Career and technical education in high school: Does it improve student outcomes?. *Thomas B. Fordham Institute*.

Dougherty, S. M. (2018). The effect of career and technical education on human capital accumulation: Causal evidence from massachusetts. *Education Finance and Policy*, 13(2), 119–148.

Dutra P., Bento G., R. S. B. T. (2017). A construção de uma política pública de educação a distância – ead pernambuco.

Elacqua, G., Navarro-Palau, P., Prada, M. F., & Soares, S. (2019). *Does technical education improve academic outcomes? Evidence from Brazil*. Technical report, IDB Working Paper Series.

Elacqua, G., Navarro-Palau, P., Prada, M. F., & Soares, S. (2020). Educación a distancia, semipresencial o presencial: ¿qué dice la evidencia? *Banco Interamericano de Desarrollo (BID)*, Hablemos de Política Educativa, No. 5.

Feng, W., Tang, J., & Liu, T. X. (2019). Understanding dropouts in moocs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33 (pp. 517–524).

Fitzpatrick, B. R., Berends, M., Ferrare, J. J., & Waddington, R. J. (2020). Virtual illusion: Comparing student achievement and teacher and classroom characteristics in online and brick-and-mortar charter schools. *Educational Researcher*, 49(3), 161–175.

Goldin, C., Katz, L., & Kuziemko, I. (2006). The homecoming of american college women: The reversal of the college gender gap. *Journal of Economic Perspectives*, 20(4), 133–153.

Hall, C. (2012). The effects of reducing tracking in upper secondary school evidence from a large-scale pilot scheme. *Journal of Human Resources*, 47(1), 237–269.

He, J., Bailey, J., Rubinstein, B., & Zhang, R. (2015). Identifying at-risk students in massive open online courses. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 29.

Hemelt, S. W., Lenard, M. A., & Paeplow, C. G. (2019). Building bridges to life after high school: Contemporary career academies and student outcomes. *Economics of Education Review*, 68, 161–178.

Hoeckel, K. & Schwartz, R. (2010). Learning for jobs oecd reviews of vocational education and training. *Austria: Organisation for Economic Co-operation and Development (OECD)*.

Imbens, G. & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, 79(3), 933–959.

Kemple, J. J. & Snipes, J. C. (2000). Career academies: Impacts on students' engagement and performance in high school.

Mizala, A., Romaguera, P., & Urquiola, M. (2007). Socioeconomic status or noise? tradeoffs in the generation of school quality information. *Journal of development economics*, 84(1), 61–75.

Neild, R. C., Boccanfuso, C., & Byrnes, V. (2015). Academic impacts of career and technical schools. *Career and Technical Education Research*, 40(1), 28–47.

Perna, L. W., Ruby, A., Boruch, R. F., Wang, N., Scull, J., Ahmad, S., & Evans, C. (2014). Moving through moocs: Understanding the progression of users in massive open online courses. *Educational Researcher*, 43(9), 421–432.

Pop-Eleches, C. & Urquiola, M. (2013). Going to a better school: Effects and behavioral responses. *American Economic Review*, 103(4), 1289–1324.

Rickles, J., Heppen, J. B., Allensworth, E., Sorensen, N., & Walters, K. (2018). Online credit recovery and the path to on-time high school graduation. *educational researcher*, 47(8), 481–491.

Ryan, P. (1998). Is apprenticeship better? a review of the economic evidence. *Journal of Vocational Education & Training*, 50(2), 289–325.

Ryan, P. (2001). The school-to-work transition: a cross-national perspective. *Journal of economic literature*, 39(1), 34–92.

Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. S. (2014). Student engagement in high school classrooms from the perspective of flow theory. In *Applications of flow in human development and education* (pp. 475–494). Springer.

Spence, M. (1978). Job market signaling. In *Uncertainty in economics* (pp. 281–306). Elsevier.

8 Tables and Figures

Table 1: Course loads in high school and technical education, total hours by modality in Pernambuco.

Modality of high school and technical education	**Starting in**	**Accumulated Academic course load (1^{st} to 3^{rd} year)**	**Technical course load (hours)**	**Total hours**
Academic regular	1^{st} year of high school	3,000	–	3,000
Academic semi full-time	1^{st} year of high school	4,200	–	4,200
Academic full-time	1^{st} year of high school	5,400	–	5,400
Technical integrated high school	1^{st} year of high school	4,200	1,200	5,400
Concurrent technical education	2^{nd} or 3^{rd} year of high school	–	800−1,200	800−1,200
Subsequent technical education	After graduate high school	–	800−1,200	800−1,200

Source: Diário Oficial do Estado de Pernambuco, Poder Executivo, 2012 and 2018.
Note: *Concurrent* technical education is provided by another institution that also belongs to the State Secretary of Education of Pernambuco and is not directly integrated into high school academic content. A student in a concurrent program can attend either regular or full-time high schools. There are also few cases of students attending both a technical integrated high school and a *concurrent* online technical course.

Table 2: (Continued on next page) Summary statistics of the population and the sample included in this analysis of students in their 2^{nd} year of high school in 2012-2017.

	Population in 2nd year of high school (1)	All matched applicants (2)	Sample in applicant lists (3)
Panel A: Student characteristics			
Number of observations	674,151	35,895	16,928
Male (%)	43.52%	39.93%	41.11%
	[0.496]	[0.49]	[0.492]
Average age	17.41	16.32	16.41
	[3.135]	[0.80]	[0.763]
Has SAEPE information (%)	68.43%	86.06%	88.53%
	[0.465]	[0.346]	[0.319]
Beneficiary of social program (%)	57.86%	55.58%	57.39%
	[0.494]	[0.497]	[0.495]
Attending school in Recife (capital city, %)	14.77%	11.17%	9.36%
	[0.355]	[0.315]	[0.291]
Mother completed high school (%)	37.12%	46.11%	44.34%
	[0.483]	[0.498]	[0.497]
Mother attended university (%)	8.55%	10.39%	9.76%
	[0.28]	[0.305]	[0.297]
Student lives in a paved street (%)	58.42%	63.65%	63.38%
	[0.493]	[0.481]	[0.482]
Average normalized math test score in 9^{th}	0.102	0.445	0.495
	[0.986]	[0.955]	[0.959]
Average normalized Portuguese test score in 9^{th}	0.108	0.466	0.499
	[0.983]	[0.938]	[0.94]
Panel B: Enrollment			
Student enrolled in state concurrent technical education (%)	0.89%	16.76%	16.65%
	[0.094]	[0.373]	[0.373]
Attends a technical integrated high school (TIHS,%)	2.50%	7.05%	8.54%
	[0.156]	[0.256]	[0.279]
Attends academic regular high school(%)	48.29%	34.72%	33.35%
	[0.5]	[0.476]	[0.471]
Attends full-time academic high school (%)	15.07%	24.46%	25.24%
	[0.358]	[0.43]	[0.434]
Attends semi full-time academic high school (%)	20.18%	27.01%	29.02%
	[0.401]	[0.444]	[0.454]
Attends private high school (%)	12.61%	5.94%	3.18%
	[0.332]	[0.236]	[0.176]
Attends federal high school (%)	1.35%	0.83%	0.67%
	[0.116]	[0.091]	[0.081]

Table 2 (Continued from previous page): Summary statistics of the population and the sample included in this analysis of students in their 2^{nd} year of high school in 2012-2017.

	Population in 2^{nd} year of high school (1)	All matched applicants (2)	Sample in applicant lists (3)
Panel C: School characteristics			
Average enrollment at the school level	872.81 [515.716]	769.37 [433.702]	751.07 [438.975]
Annual per pupil spending (R$)	3,304 [1665.874]	3,664 [1735.726]	3,613 [1616.789]
Hourly teachers' salary in school	16 [4.502]	17 [4.574]	17 [4.563]
School has a laboratory (%)	51.21% [0.5]	59.00% [0.492]	58.96% [0.492]
% Teachers have less than 5 years of experience	31.16% [0.404]	23.53% [0.373]	23.13% [0.37]
Teacher graduated from federaluniversity* (%)	10.88% [0.122]	10.87% [0.12]	10.10% [0.115]
Average students per teacher in school	27.29 [7.802]	26.16 [6.53]	26.12 [6.717]
Panel D: Outcomes			
Student repeats a year in HS (2^{nd} and/or 3^{rd} year)	10.89% [0.311]	5.4% [0.226]	5.5% [0.229]
Average normalized Math SAEPE score in 3^{rd} year of HS	0.036 [0.993]	0.386 [0.963]	0.415 [0.961]
Average normalized Portuguese SAEPE score in 3^{rd} year of HS	0.034 [0.984]	0.407 [0.885]	0.423 [0.871]
Student drops out between 2^{nd} and 3^{rd} year of HS	21% [0.209]	9.6% [0.095]	8.9% [0.089]

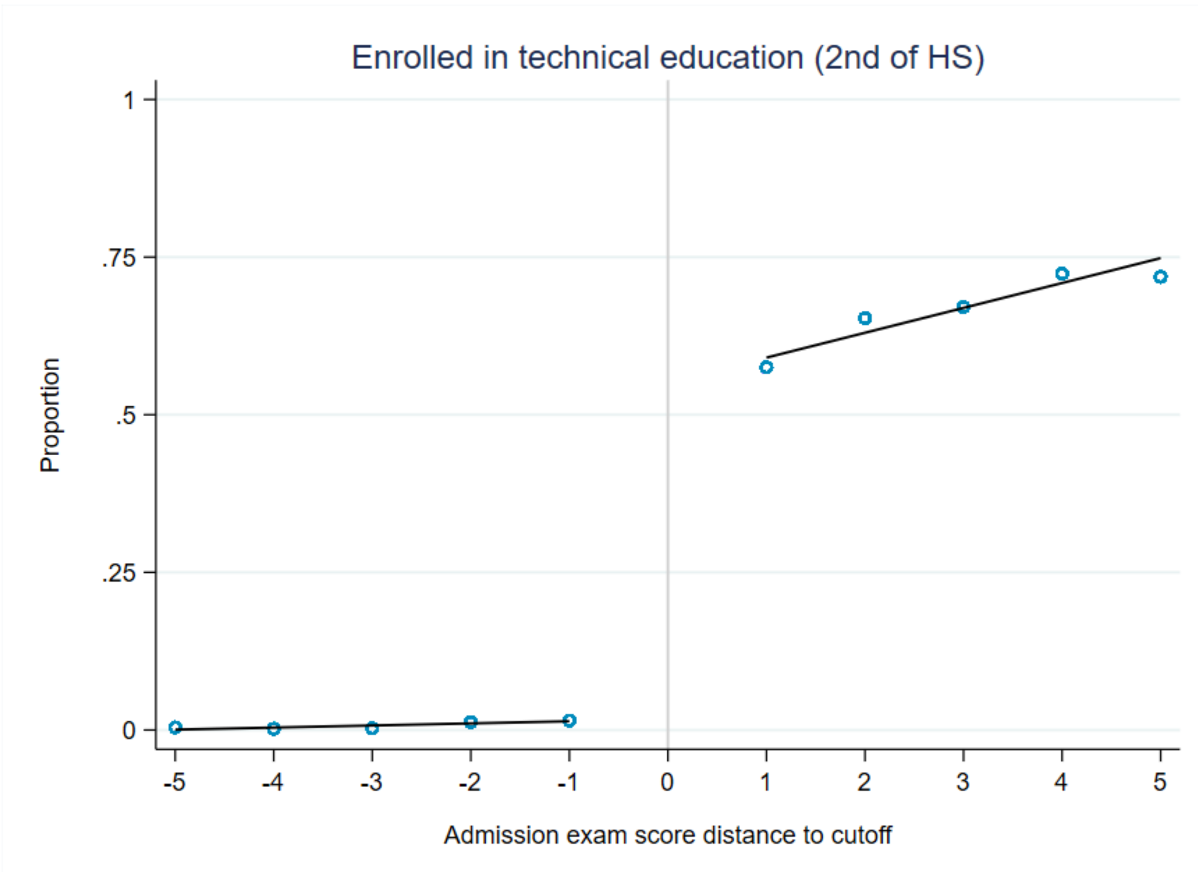
Notes: Column (1) shows summary statistics for all students in the population, Column (2) shows summary statistics for students in Column (1) who were matched to the online technical education applicant lists (using students' official identification number (CPF), date of birth and name) and Column (3) shows summary statistics for students in Column (2), excluding students who were in private school in the 92^{th} grade, students who applied to technical courses and schools that were not oversubscribed, students without admission exam scores and students later identified as studying technical courses under the subsequent modality (i.e., applied to online technical education after graduating high school). Test scores are normalized within years and grades and include all standardized exam (SAEPE) takers, some of which are not matched. Thus, normalized test scores have a mean of zero and a standard deviation of 1. Standard deviations are presented in brackets. *Only students in state and municipal schools participate in the standardized exam (SAEPE), from which we obtain socioeconomic information. Therefore, students' socioeconomic condition and proficiency level in the 9^{th} grade from Panel B do not include information from students of private and federal schools. Consequently, the socioeconomic information provided in the descriptive statistics for the population and for the list of students applying to concurrent online technical education may be slightly downward biased. **Federal Universities are public and considered the best universities in Brazil. To attend a course in a Federal University, a candidate must take an exam in a very competitive selection process.

Table 3: Changes in the probability of enrolling in *concurrent* online technical education in the 2^{nd} year of HS for students above/below the admission exam cutoff.

	Enrolled in *concurrent* online technical education	
	(1)	(2)
Above the cutoff	0.4743*** (0.0159)	0.4024*** (0.0248)
Flexible linear spline	No	Yes
Admission exam year FE, Admission exam semester FE, School FE and concurrent track FE	Yes	Yes
Total observations	4,144	8,581
R-squared	0.2994	0.4379
Adjusted R-squared	0.2828	0.4314
Control mean in bandwidth	0.0893	0.0498
Bandwidth	1*	3

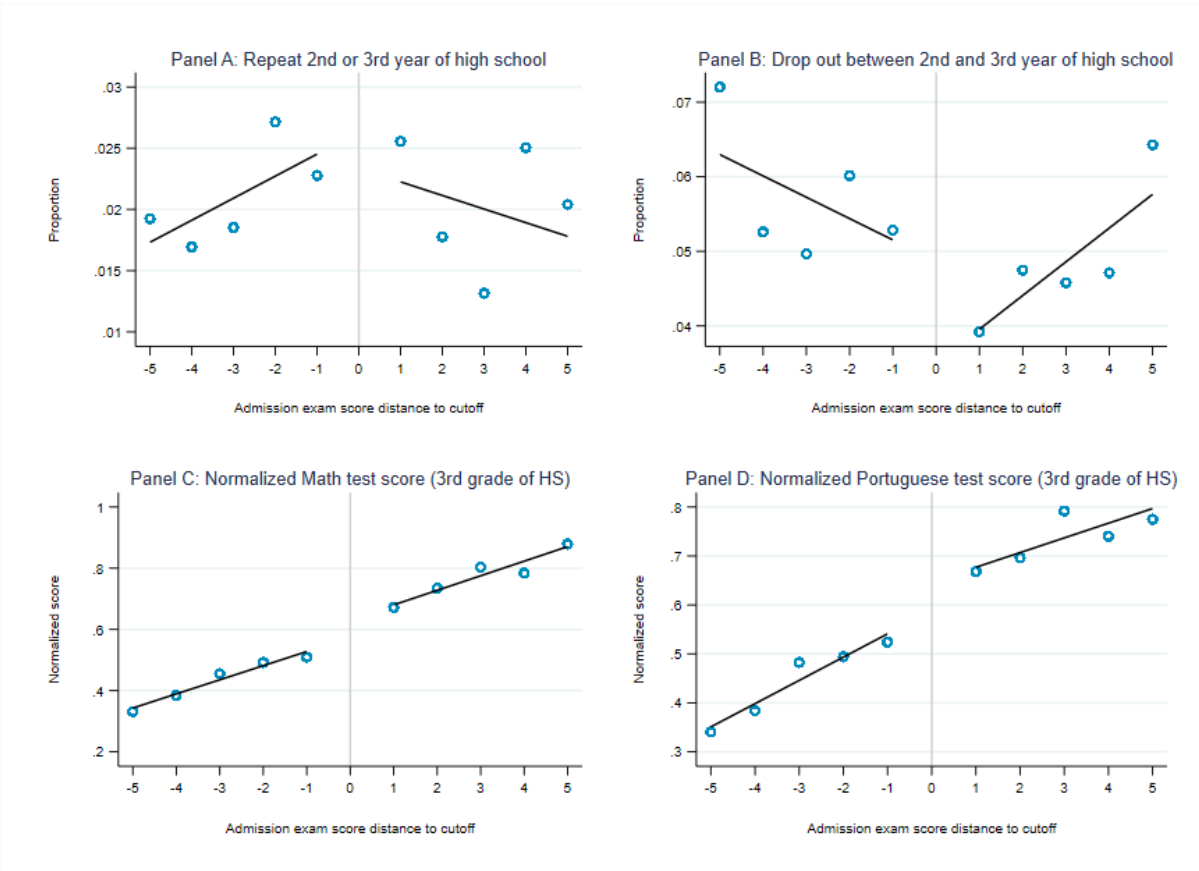
Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regression in the first column above uses bandwidth 1, since the bandwidth suggested by the procedure in Imbens and Kalyanaraman (2012) was smaller than 1 (0.8755) and did not include observations on both sides of the cutoff. The regression in the second column uses bandwidth 3. Additionally, they include application school, concurrent track, and application year and semester fixed effects. The sample excludes students who were in private school in the 9^{th} grade and students who applied to online technical courses that were not oversubscribed and had no cutoff.

Figure 1: Changes in the probability of enrolling in *concurrent* online technical education in the 2^{nd} year of HS for students above/below the admission exam cutoff.



Note: The graph above shows the proportion of students within each bin of distance to the admission exam cutoff went on to enroll in a concurrent technical program in their 2^{nd} year of high school. The admission exam cutoff is computed as the score of the first student who was not eligible to enroll in the program even though he/she had passed the admission exam. The sample excludes students who were in private schools in the 9th grade, as well as students who apply to technical courses and schools that are not oversubscribed, where there is no cutoff. Additionally, the graph above excludes students at the cutoff, since some of these students are administratively classified as having scored above the cutoff.

Figure 2: Change in outcomes in 3^{rd} year of HS for students above the admission exam score cutoff.



Note: The graph above shows the proportion of students within each bin of distance to admission exam cutoff that were enrolled in concurrent technical programs in state schools in their 2^{nd} year of high school. The admission exam cutoff is computed as the score of the first student who was not eligible to enroll in the program even though he/she had passed the admission exam. The sample excludes students who were in private school in the 9^{th} grade and students who applied to technical courses and schools that were not oversubscribed, where there is no cutoff. Additionally, the graph above excludes students at the cutoff, since some of these students were administratively classified as being above the cutoff.

Table 4: Changes in outcomes for students above the admission exam score cutoff.

	Probability of repeating		Probability of dropping out		Normalized math 3rd year of HS test score		Normalized Portuguese 3rd year of HS test score		Probability of graduating from HS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above the cutoff	0.0097 (0.0060)	0.0164* (0.0089)	-0.0241*** (0.0093)	-0.0311** (0.0140)	0.1923*** (0.0292)	0.0485 (0.0580)	0.1380*** (0.0336)	0.0686 (0.0499)	-0.0016 (0.0057)	0.0032 (0.0089)
Flexible linear spline	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Admission exam year FE, School FE and concurrent track FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	4,115	8,527	3,797	7,877	5,124	6,839	3,305	6,839	3,717	7,736
R-squared	0.0465	0.0353	0.0683	0.0519	0.1168	0.1156	0.0901	0.0695	0.0411	0.0219
Adjusted R-squared	0.0237	0.0240	0.0442	0.0399	0.0997	0.1026	0.0626	0.0558	0.0154	0.0093
Control mean in bandwidth	0.0225	0.0266	0.0812	0.0857	0.4957	0.4779	0.5062	0.4835	0.9766	0.9771
IK Bandwidth	1.2865	3	1.4746	3	2.2203	3	1.9953	3	1.3022	3

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regressions above use the bandwidth suggested by the procedure in Imbens and Kalyanaraman (2012) for the first regression of each outcome and 3, roughly twice the former bandwidth, for the second regression of each outcome. Additionally, they include application school, concurrent track and application year and semester fixed effects. The sample excludes students who were in private school in the 9^{th} grade, and students who applied to technical courses, schools that were not oversubscribed, and observations for which we could not identify whether the student dropped out or not between the 2^{nd} and 3^{rd} years of high school. Probability of repeating: student repeats at least one grade during high school after applying to a concurrent program; Probability of dropping out: student disappears from Pernambuco's educational census between the 2^{nd} and 3^{rd} years of high school and before the last year of our sample (2017); Normalized math 3rd year of HS test score: SAEPE score in math in the 3^{rd} year of high school; Normalized Portuguese 3rd year of HS test score: SAEPE score in Portuguese in the 3^{rd} year of high school. Test scores are normalized within year and grade and have a mean of zero and standard deviation of 1.

Table 5: High school and technical education modalities in Brazil and Pernambuco (2013 and 2018).

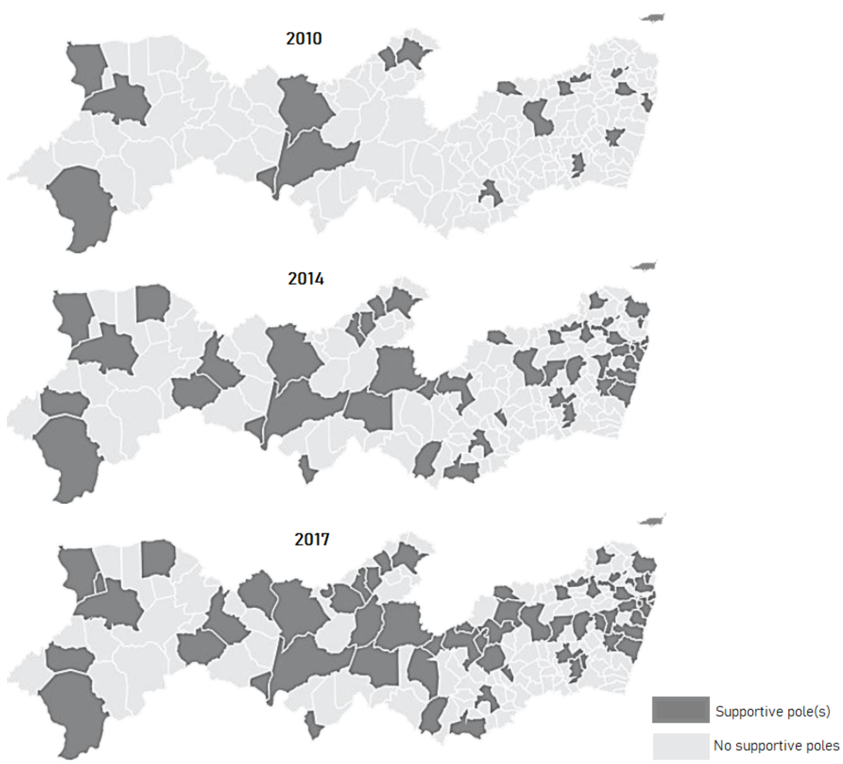
	2013				2018			
	Brazil		Pernambuco		Brazil		Pernambuco	
Academic regular	7,505,895	89.80%	275,557	71.10%	6,578,388	84.90%	179,115	52.60%
Academic full time	39,272	0.50%	74,420	19.20%	192,631	2.50%	125,188	36.80%
Technical integrated with HS	338,417	4.10%	12,265	3.20%	505,791	6.50%	22,068	6.50%
Concurrent technical education	**310,218**	**3.70%**	**7,048**	**1.80%**	**354,346**	**4.60%**	**10,481**	**3.10%**
Other technical education in HS	161,515	1.90%	18,410	4.70%	113,918	1.50%	3,629	1.10%
Total high school	8,355,317	100.00%	387,700	100.00%	7,745,074	100.00%	340,481	100%
Subsequent technical education	792,796	9%	34,727	9%	894,862	12%	67,918	20%

Source: School Census.
Notes: a. Includes adult and teacher education (EJA and normal/Magisterio). In Brazil, there is a type of technical training that is not included in the table, called *formação inicial e continuada* (FIC) or initial and continuing training courses. These courses are provided independently of any academic education level. The completion of these courses (FICs) typically does not count toward the completion of a formal academic education level, and thus such courses do not qualify students to take courses at the next level. "HS" is an abbreviation for high school.

Table 6: *Concurrent* online technical education weekly course load.

Average course load *per week*	
Online	**In-person**
12 hrs	3 hrs
3 hours of online, asynchronous classes administered by a teacher; **9 hours** of complementary activities and practice in the Virtual Learning Platform.	In-person mandatory class/activities conducted in a **Support Pole** equipped with laboratories, computers, and tutors.

Figure 3: Expansion in *supportive poles* for online technical education in Pernambuco in 2010, 2014, and 2017 (top to bottom).



Source: Dutra P. (2017)

Table 7: Change in the probability that students above the admission cutoff have a given socioeconomic characteristic.

	Took standardized test in 9th grade		Male		Age		Mother completed high school		Mother has university education		Normalized math 9th test score		Student lives in a paved street	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Above the cutoff	-0.0082	-0.0047	-0.0309*	-0.0510**	0.0050	0.0496	-0.0146	-0.0490*	-0.0135	-0.0401**	0.1969***	0.0663	-0.0085	-0.0060
	(0.0113)	(0.0171)	(0.0170)	(0.0255)	(0.0231)	(0.0454)	(0.0194)	(0.0290)	(0.0122)	(0.0185)	(0.0314)	(0.0604)	(0.0183)	(0.0277)
Flexible linear spline	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Admission exam year and semester FE and concurrent track FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	4,144	8,581	4,143	8,580	6,426	8,581	3,419	7,103	3,419	7,103	4,412	5,849	3,592	7,451
R-squared	0.0619	0.0455	0.1066	0.0922	0.0606	0.0572	0.0778	0.0631	0.0419	0.0266	0.0945	0.1025	0.0537	0.0401
Adjusted R-squared	0.0396	0.0344	0.0854	0.0816	0.0464	0.0462	0.0512	0.0498	0.0142	0.0129	0.0743	0.0871	0.0277	0.0272
Control mean in bandwidth	0.8864	0.8901	0.4345	0.4218	16.7779	16.7849	0.4617	0.4477	0.1175	0.1008	0.5567	0.5417	0.6367	0.6332
Bandwidth	1.5144	3	1.5853	3	2.1203	3	1.5520	3	1.5501	3	2.0406	3	1.7851	3

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regressions above use the bandwidth suggested by the procedure in Imbens and Kalyanaraman (2012) for the first regression of each outcome and 3, roughly twice the first bandwidth for the main outcomes for the second regression of each variable. Additionally, they include application school, concurrent track, and application year and semester fixed effects. Test scores are normalized within year and grade and have a mean of zero and standard deviation of 1. The sample excludes students who were in private school in the 9^{th} grade and students who applied to technical courses and schools that were not oversubscribed, where there is no cutoff.

Table 8: Differences in schools attended by students above and below the cutoff in 2^{nd} year of HS.

	Enrolled in technical integrated high school (TIHS)		Enrolled in full-time academic school		Enrolled in semi full-time academic school		Enrolled in academic regular school		Enrolled in academic private school		Enrolled in federal school	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Above the cutoff	-0.0006 (0.0081)	-0.0024 (0.0121)	0.0098 (0.0131)	0.0123 (0.0196)	0.0139 (0.0133)	0.0229 (0.0200)	-0.0270* (0.0144)	-0.0230 (0.0216)	-0.0016 (0.0057)	-0.0136 (0.0090)	0.0054 (0.0035)	0.0027 (0.0054)
Flexible linear spline	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Admission exam year and semester FE, concurrent track FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	4,968	10,294	4,968	10,294	4,968	10,294	4,968	10,294	4,968	10,294	4,968	10,294
R-squared	0.2460	0.2203	0.2301	0.2264	0.2343	0.2247	0.1021	0.0959	0.0697	0.0644	0.0578	0.0409
Adjusted R-squared	0.2312	0.2126	0.2149	0.2188	0.2192	0.2171	0.0844	0.0870	0.0513	0.0553	0.0393	0.0315
Control mean in bandwidth	0.0842	0.0793	0.2717	0.2678	0.2832	0.2888	0.3141	0.3225	0.0368	0.0330	0.0086	0.0075
Bandwidth	1.4240	3	1.5250	3	1.6088	3	1.5422	3	1.2973	3	1.1216	3

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regressions above use the bandwidth suggested by the procedure in Imbens and Kalyanaraman (2012) for the first regression of each outcome and 3, roughly twice as much as the first bandwidth for the main outcomes for the second regression of each variable. Additionally, they include application school, concurrent track and application year and semester fixed effects. Test scores are normalized within year and grade and have a mean of zero and a standard deviation of 1. The sample excludes students who were in private school in the 9^{th} grade and students who applied to technical courses and schools that were not oversubscribed, where there is no cutoff.

Table 9: (Continued on next page) Differences in the characteristics of schools attended by students above and below the cutoff in 2^{nd} year of HS.

	Per pupil spending		School has a laboratory		Principal reported the school has good environment		% Teachers graduated in non exact science and teach math		% Teachers have less than 5 years of experience	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above the cutoff	14.6	32.4	0.0104	0.0020	-0.0106	-0.0172	-0.0172**	-0.0305***	-0.0209*	-0.0176
	(31.594)	(71.952)	(0.0163)	(0.0312)	(0.0154)	(0.0235)	(0.0079)	(0.0118)	(0.0112)	(0.0168)
Flexible linear spline	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Admission exam year and semester FE, concurrent track FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	9,827	9,827	3,791	5,061	4,302	8,873	4,718	9,788	4,654	9,644
R-squared	0.3146	0.3147	0.1749	0.1742	0.2507	0.2305	0.3055	0.3037	0.1575	0.1368
Adjusted R-squared	0.3077	0.3077	0.1544	0.1584	0.2338	0.2218	0.2910	0.2965	0.1398	0.1278
Control mean in bandwidth	3653.2	3653.2	0.5979	0.5915	1.3775	1.3797	0.4160	0.4160	0.2301	0.2337
Bandwidth	3.5407	3	2.0058	3	1.3775	3	1.4331	3	1.4649	3

Table 9 (Continued from previous page) Differences in the characteristics of schools attended by students above and below the cutoff in 2^{nd} year of HS.

	Hourly salary	teacher	% Teachers graduated in a Federal University		Proportion of mothers with university at baseline (3rd year of EM) in school in 2012		Average Math normalized test score at baseline (3rd year of EM) in school in 2012	
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Above the cutoff	0.1094	0.2167	0.0002	-0.0001	0.0003	-0.0058*	-0.0058	-0.0342
	(0.1420)	(0.3180)	(0.0025)	(0.0037)	(0.0022)	(0.0034)	(0.0165)	(0.0250)
Flexible linear spline	No	Yes	No	Yes	No	Yes	No	Yes
Admission exam year and semester FE, concurrent track FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	4,158	4,158	4,302	8,957	3,717	7,725	3,699	7,687
R-squared	0.2118	0.2122	0.5336	0.5245	0.2520	0.2269	0.2252	0.2113
Adjusted R-squared	0.1938	0.1937	0.5230	0.5192	0.2324	0.2169	0.2048	0.2011
Control mean in bandwidth	17.1945	17.1945	0.0908	0.0908	0.0793	0.0789	0.1502	0.1381
Bandwidth	3.9600	3	1.1688	3	1.0716	3	1.9001	3

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regressions above use the bandwidth suggested by the procedure in Imbens and Kalyanaraman (2012) for the first regression of each outcome and 3, roughly twice the former bandwidth, of the main outcomes for the second regression of each outcome. Additionally, they include application school, concurrent track and application year and semester fixed effects. The sample excludes students who were in private school in the 9^{th} grade and students who applied to technical courses and schools that were not oversubscribed, where there is no cutoff. For the purposes of this table, the information at baseline is the information gathered during the SAEPE standardized test in the 3^{rd} year of high school for students enrolled in that high school in 2012.

Table 10: Heterogeneity Analysis - changes in outcomes for students above admission exam cutoff for the online technical education – subsample restricted to low-SES students.

	Probability of repeating		Probability of dropping out		Normalized math 3rd year of HS test score		Normalized Portuguese 3rd year of HS test score		Probability of ever complete HS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above the cutoff	0.0110 (0.0076)	0.0141 (0.0120)	-0.0114 (0.0120)	-0.0139 (0.0182)	0.1909*** (0.0425)	0.1341 (0.0839)	0.1698*** (0.0386)	0.1483** (0.0713)	-0.0048 (0.0075)	0.0080 (0.0134)
Flexible linear spline	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Admission exam year and semester FE and concurrent track FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	1,854	3,849	1,767	3,665	2,426	3,239	2,426	3,239	1,658	3,471
R-squared	0.0882	0.0554	0.0786	0.0546	0.1318	0.1311	0.0925	0.0901	0.0919	0.0438
Adjusted R-squared	0.0406	0.0312	0.0279	0.0292	0.0971	0.1042	0.0563	0.0620	0.0385	0.0163
Control mean in bandwidth	0.0147	0.0205	0.0562	0.0586	0.4422	0.4294	0.4414	0.4322	0.9844	0.9839
IK Bandwidth	1.4110	3	1.6333	3	2.5747	3	2.4616	3	1.4381	3

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regressions above use the bandwidth suggested by the procedure in Imbens and Kalyanaraman (2012) for the first regression of each outcome and 3, roughly twice the former bandwidth, of the main outcomes for the second regression of each outcome. Additionally, they include application school, concurrent track, and application year and semester fixed effects. Test scores are normalized within year and grade and have a mean of zero and standard deviation of 1. A student is classified as low-socioeconomic status (SES) if her/his mother has not completed high school. This indicator has been demonstrated as a good socioeconomic proxy in Brazil and has significant correlation with other socioeconomic measures, such as household income and whether the family received aid through Bolsa Família. The sample excludes high-SES students, students who were in private school in the 9^{th} grade and students who applied to technical courses and schools that were not oversubscribed, where there is no cutoff.

Table 11: Heterogeneity Analysis - changes in outcomes for students above the admission exam cutoff – subsample restricted to students who applied to *supportive poles* in full-time or *semi* full-time schools.

	Probability of repeating		Probability of dropping out		Normalized math 3rd year of HS test score		Normalized Portuguese 3rd year of HS test score		Probability of ever complete HS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above the cutoff	0.0055	0.0163	-0.0388**	-0.0543**	0.1500***	-0.0223	0.1404***	0.0094	-0.0167*	-0.0171
	(0.0087)	(0.0121)	(0.0157)	(0.0237)	(0.0465)	(0.0915)	(0.0439)	(0.0810)	(0.0099)	(0.0150)
Flexible linear spline	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Admission exam year and semester FE, concurrent track FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	1,350	2,744	1,391	2,799	1,875	2,473	1,875	2,473	1,330	2,710
R-squared	0.0722	0.0788	0.0862	0.0634	0.1277	0.1228	0.0683	0.0657	0.0728	0.0485
Adjusted R-squared	0.0358	0.0607	0.0521	0.0453	0.1033	0.1035	0.0422	0.0452	0.0358	0.0295
Control mean in bandwidth	0.0181	0.0208	0.0990	0.0971	0.4774	0.4626	0.5060	0.4930	0.9826	0.9836
IK Bandwidth	1.5126	3	1.7787	3	2.6730	3	2.6279	3	1.5280	3

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regressions above use the bandwidth suggested by the procedure in Imbens and Kalyanaraman (2012) for the first regression of each outcome and 3, roughly twice the former bandwidth of the main outcomes for the second regression of each outcome. Additionally, they include application school, concurrent track, and application year and semester fixed effects. Test scores are normalized within year and grade and have a mean of zero and standard deviation of 1. The sample excludes students who did not apply to a supportive pole in a full-time or semi full-time school, students who were in private school in the 9^{th} grade and students who applied to technical courses and schools that were not oversubscribed, where there is no cutoff.

Table 12: Robustness check - changes in outcomes for students above the admission exam cutoff – excluding students at the cutoff

	Probability of repeating		Probability of dropping out		Normalized math 3rd year of HS test score		Normalized Portuguese 3rd year of HS test score		Probability of ever complete HS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above the cutoff	0.0065	0.0073	-0.0267**	-0.0393**	0.2379***	0.0993	0.2049***	0.0925	0.0014	0.0101
	(0.0067)	(0.0104)	(0.0109)	(0.0170)	(0.0316)	(0.0672)	(0.0285)	(0.0596)	(0.0066)	(0.0105)
Flexible linear spline	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Admission exam year and Semester FE and concurrent track FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	2,467	6,904	2,275	6,380	3,776	5,495	3,776	5,495	2,219	6,247
R-squared	0.0693	0.0409	0.0928	0.0565	0.1321	0.1230	0.0788	0.0713	0.0899	0.0285
Adjusted R-squared	0.0324	0.0270	0.0537	0.0416	0.1095	0.1070	0.0548	0.0543	0.0492	0.0128
Control mean in bandwidth	0.0269	0.0297	0.0858	0.0886	0.4547	0.4446	0.4703	0.4681	0.9727	0.9758
IK Bandwidth	1.2852	3	1.4713	3	2.2098	3	2.0038	3	1.3032	3

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regressions above use the bandwidth suggested by the procedure in Imbens and Kalyanaraman (2012) for the first regression of each outcome and 3, roughly twice the former bandwidth of the main outcomes for the second regression of each outcome. Additionally, they include application school, concurrent track, and application year and semester fixed effects. Test scores are normalized within year and grade and have a mean of zero and standard deviation of 1. The sample excludes students who had a score at the cutoff, students who were in private school in the 9^{th} grade and students who applied to technical courses and schools that were not oversubscribed, where there is no cutoff.

Table 13: Robustness check - changes in outcomes for students above the admission exam cutoff – including students in private schools in 9^{th} grade.

	Probability of repeating		Probability of dropping out		Normalized math 3rd of HS test score		Normalized Portuguese 3rd year of HS test score		Probability of completing HS	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above the cutoff	0.0040	0.0055	-0.0273***	-0.0326***	0.1917***	0.0226	0.1726***	0.0613	-0.0042	0.0022
	(0.0051)	(0.0075)	(0.0078)	(0.0115)	(0.0251)	(0.0495)	(0.0224)	(0.0429)	(0.0052)	(0.0081)
Flexible linear spline	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Admission exam year and Semester FE and concurrent track FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	5,574	11,486	5,217	10,765	6,803	9,066	6,803	9,066	5,012	10,396
R-squared	0.0359	0.0307	0.0552	0.0534	0.1048	0.1070	0.0605	0.0620	0.0394	0.0217
Adjusted R-squared	0.0188	0.0222	0.0373	0.0445	0.0916	0.0970	0.0466	0.0514	0.0203	0.0121
Control mean in bandwidth	0.0268	0.0286	0.0834	0.0859	0.5237	0.5060	0.5148	0.5040	0.9757	0.9766
IK Bandwidth	1.2409	3	1.4031	3	2.1700	3	2.0043	3	1.2491	3

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regressions above use the bandwidth suggested by the procedure in Imbens and Kalyanaraman (2012) for the first regression of each outcome and 3, roughly twice the former bandwidth of the main outcomes for the second regression of each outcome. Additionally, they include application school, concurrent track, and application year and semester fixed effects. Test scores are normalized within year and grade and have a mean of zero and standard deviation of 1. The sample excludes students whose score was at the cutoff and students who applied to technical courses and schools that were not oversubscribed, where there is no cutoff.

Table 14: Robustness check – alternative bandwidths and placebo cutoffs for dropping out of high school between the 2^{nd} and 3^{rd} years of HS.

	Probability of dropping out (Alternative bandwidths)		Probability of dropping out (placebo cutoff)	
	(1)	(2)	(3)	(4)
Above the cutoff	-0.0416**	-0.0264**	-0.0127	0.0071
	(0.0195)	(0.0110)	(0.0113)	(0.0132)
Flexible linear spline	Yes	Yes	Yes	Yes
Admission exam year and semester FE, concurrent track FE	Yes	Yes	Yes	Yes
Total observations	5,948	11,152	7,931	7,931
R-squared	0.0544	0.0524	0.0520	0.0516
Adjusted R-squared	0.0386	0.0438	0.0401	0.0397
Cutoff	0	0	-2	-3
Control mean in bandwidth	0.0869	0.0910	0.0899	0.0811
IK Bandwidth	2.9426	5	1.4963	1.6080

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. Column (1) reports double the IK bandwidth; Column (2) uses an alternative bandwidth of 5. Columns (3) and (4) use alternative cutoffs of the admission exam score that is 2 and 3 points to the left of the actual cutoff (0). We also only focus on dropouts as an outcome since it is the only outcome that was statistically significant. Additionally, the regressions include application school, concurrent track, and application year and semester fixed effects. Test scores are normalized within year and grade and have a mean of zero and standard deviation of 1. The sample excludes students who had scores at the cutoff, and students who applied to technical courses and schools that were not oversubscribed, where there is no cutoff.

Table 15: Change in the probability that students have a given socioeconomic characteristic for students above the cutoff *excluding students at the cutoff*

	Took standardized test in 9th grade		Male		Age		Mother completed high school		Mother has university education		Normalized math 9th test score		Student lives in a paved street	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Above the cutoff	-0.0118	-0.0121	-0.0393**	-0.0663**	-0.0014	0.0317	-0.0160	-0.0645*	-0.0155	-0.0549**	0.2158***	0.0867	-0.0144	-0.0059
	(0.0130)	(0.0203)	(0.0197)	(0.0303)	(0.0253)	(0.0540)	(0.0224)	(0.0345)	(0.0143)	(0.0220)	(0.0342)	(0.0709)	(0.0210)	(0.0328)
Flexible linear spline	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Admission exam year and semester FE, concurrent track FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	2,467	6,904	2,466	6,903	4,749	6,904	2,042	5,726	2,042	5,726	3,281	4,718	2,142	6,001
R-squared	0.0831	0.0486	0.1321	0.0991	0.0639	0.0586	0.0918	0.0637	0.0669	0.0304	0.0964	0.1039	0.0811	0.0453
Adjusted R-squared	0.0468	0.0348	0.0977	0.0860	0.0446	0.0449	0.0480	0.0473	0.0218	0.0133	0.0694	0.0849	0.0389	0.0293
Control mean in bandwidth	0.8868	0.8916	0.4349	0.4171	16.7842	16.7918	0.4709	0.4454	0.1207	0.0956	0.5436	0.5274	0.6450	0.6345
Bandwidth	1.5144	3	1.5853	3	2.1203	3	1.5520	3	1.5501	3	2.0406	3	1.7851	3

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regressions above use the bandwidth suggested by the procedure in Imbens and Kalyanaraman (2012) for the first regression of each outcome and 3, roughly twice the former bandwidth for the main outcomes for the second regression of each variable. Additionally, they include application school, concurrent track, and application year and semester fixed effects. Test scores are normalized within year and grade and have a mean of zero and standard deviation of 1. The sample excludes students who were in private school in the 9^{th} grade and students who applied to technical courses and tracks that were not oversubscribed. Additionally, this sample also excludes students at the cutoff.

Table 16: Balance check of demographic characteristics and previous learning for students above the cutoff *restricted to low-SES students.*

	Male		Age		Normalized Math 9th test score		Normalized Language 9th test score	
	(3)	(4)	(5)	(6)	(11)	(12)	(11)	(12)
Above the cutoff	-0.0351	-0.0584	0.0345	0.0699	0.2027***	0.1481*	0.2680***	0.2446***
	(0.0252)	(0.0379)	(0.0349)	(0.0673)	(0.0437)	(0.0832)	(0.0427)	(0.0803)
Flexible linear spline	No	Yes	No	Yes	No	Yes	No	Yes
Admission exam year and semester FE, concurrent track FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	1,868	3,876	2,914	3,876	2,362	3,131	2,362	3,131
R-squared	0.1142	0.0856	0.0727	0.0649	0.1194	0.1147	0.1008	0.0939
Adjusted R-squared	0.0683	0.0624	0.0421	0.0412	0.0837	0.0867	0.0644	0.0653
Control mean in bandwidth	0.4007	0.3876	16.8442	16.8543	0.5135	0.5079	0.5061	0.4972
Bandwidth	1.9531	1.9531	2.4566	2.4566	2.3546	2.3546	2.1702	2.1702

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regression in the first column above uses bandwidth 1, as the bandwidth suggested by the procedure in Imbens and Kalyanaraman (2012) was smaller than 1 (0.8755) and did not include observations on both sides of the cutoff. The regression in the second column uses bandwidth 3. Additionally, they include application school, concurrent track, and application year and semester fixed effects. The sample excludes students who were in private school in the 9^{th} grade and students who applied to technical courses and schools that were not oversubscribed. Additionally, this subsample is restricted to students with low socioeconomic status. To assign a designation of low socioeconomic status, we use the student's mother education. Students' mothers who have only elementary education or less are considered low socioeconomic status. We found similar results using receipt of aid through Bolsa Familia as a socioeconomic proxy.

Table 17: Change in the probability that students have a given socioeconomic characteristic for students above the cutoff *restricted to students who applied to supportive poles in full time or semi- full time schools.*

	Took standardized test in 9th grade		Male		Age		Mother completed high school		Mother has university education		Normalized math 9th test score		Student lives in a paved street	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Above the cutoff	-0.0157	-0.0302	0.0073	-0.0144	-0.0583	-0.0458	-0.0263	-0.0541	-0.0366*	-0.0829***	0.1365***	0.0022	-0.0359	-0.0216
	(0.0186)	(0.0272)	(0.0223)	(0.0424)	(0.0393)	(0.0775)	(0.0326)	(0.0481)	(0.0207)	(0.0317)	(0.0520)	(0.0970)	(0.0304)	(0.0458)
Flexible linear spline	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Admission exam year and semester FE, concurrent track FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	1,506	3,029	2,305	3,028	2,306	3,029	1,243	2,545	1,243	2,545	1,610	2,117	1,323	2,683
R-squared	0.0627	0.0366	0.0869	0.0903	0.0732	0.0726	0.0735	0.0569	0.0470	0.0339	0.1018	0.1114	0.0647	0.0535
Adjusted R-squared	0.0305	0.0194	0.0666	0.0741	0.0527	0.0561	0.0346	0.0368	0.0070	0.0133	0.0730	0.0885	0.0279	0.0344
Control mean in bandwidth	0.9006	0.9058	0.4108	0.3999	16.7427	16.7452	0.5022	0.4660	0.1408	0.1131	0.5648	0.5391	0.6254	0.6111
Bandwidth	1.9213	3	2.2078	3	2.5429	3	1.6942	3	1.7629	3	2.1943	3	1.8736	3

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regression in the first column above uses bandwidth 1, as the bandwidth suggested by the procedure in Imbens and Kalyanaraman (2012) was smaller than 1 (0.8755) and did not include observations on both sides of the cutoff. The regression in the second column uses bandwidth 3. Additionally, they include application school, concurrent track, and application year and semester fixed effects. The sample excludes students who were in private school in the 9^{th} grade and students who applied to technical courses and schools that were not oversubscribed. Additionally, this subsample is restricted to students who applied to supportive poles in full-time or semi full-time schools.

Appendix

1A. A closer look at student retention in online technical education

Massive open online courses (MOOCs) have experienced a consistent increase in enrollment numbers over the recent years. Despite the successful trend, MOOC retention rates continue to be a major challenge: only around 5% of initial applicants complete the courses (Perna et al. (2014); He et al. (2015); Feng et al. (2019)). On this topic, there is an incipient body of literature that attempts to better understand student behavior in online courses.

Consistent with this global trend, Pernambuco's online technical courses have high dropout rates. According to the online learning administrative database of the Secretary of Education, 69.3% of *concurrent* students dropped out at some point of the online course and 11.8% of the enrolled students completed it, as shown in Table 18[38]. In practice, students who drop out of an online course continue attending secondary school and might apply again to another online course in the following semester(s). Students with "cancelled enrollment" make up 4.5% of the sample, which represents the proportion of students that cancelled the course subscription before having started it, while 14.3% are classified as "active enrollment" which means that the student has not finished the online course. The State Secretary of Education of Pernambuco allows a five-year period to graduate. Most of the students with "active enrolment" have finished all modules and have only the final evaluation pending.

To better understand who is more likely to drop out and complete the course, we estimate a linear probability model controlling for a set of student and course characteristics restricting our final sample. Table 19 reveals that female and older students are more likely to drop out, as well as students with higher socioeconomic status[39]. On

[38]Qualitative interviews with managers from Pernambuco's State Secretary of Education reveal that around 50% of students drop out before completing the 1^{st} module of the course. The sample we analyze in Table 18 corresponds to all students above the cutoff in our final sample that we matched with the administrative database. The total number of students (3,897) mentioned in Table 18 is higher than the total number of enrolled students in our sample (16% of 16,928, or approximately 2,819). This probably happens because some students on applicant lists with a status different from "enrolled" (such as "approved" or "not classified") managed to enroll later in their course of interest and have their status updated to "enrolled" in the final administrative database of the Secretary of Education.

[39]We considered different models for the socioeconomic proxy: i) a dummy variable for a mother with a university degree; ii) a dummy variable for a mother with completed high school or more; iii) whether the student's family is a beneficiary under Bolsa Família. All of these were positively correlated with dropout rates from online courses.

the other hand, students with higher academic proficiency in 9^{th} grade – especially in language – are less likely to drop out and have greater chances of completing the online course. Also, there is no statistically significant correlation between the year the student started the course (2^{nd} or 3^{rd}) and the probability of dropping out the online course. Finally, students attending a supportive pole (for the in-person activities) located in a full-time or *semi* full-time school tend to drop out at lower rates and are more likely to complete the online course than students attending a supportive pole in technical high schools (*Escolas Técnicas de Pernambuco*; ETE), which are the same schools that provide technical integrated high schools in the state.

The positive correlation between the online course graduation rate and the supportive pole being located at a full-time or semi full time school may indicate some advantages of attending in-person activities in these poles, as a logistic ease for the students who also attend high school classes in a full time school (recall that 54% of our sample attend high school in a full time or semi-full time school[40]). On the other hand, as already shown in Table 2, only 8% of students in our sample also attends academic classes in a technical integrated high school, therefore, our sample's majority attends secondary education either in regular or in full time or semi-full time schools. Therefore, we might expect logistics difficulties to commute to a technical high school only for the in-person activities. Most technical high schools are located out of the urban perimeter, which represents an extra effort for the students not only in terms of time, but also financially. *Concurrent* modality do not have free public transportation available, so distance and transportation costs are a barrier for attendance. It is worth recalling that, until 2017, if a student did not attend 75% of the in-person classes in the supportive pole, she not graduate from the online course.

Finally, in 2017, the State Secretariat of Education conducted a survey to explore the main reasons that led students to drop out of the online technical course. As Figure 4 shows, the two main causes reported are, respectively, "*have found a job*" (31%)[41] and "*began to attend higher education*" (26%). Interestingly, these two main reasons might be themselves, a potential *"side effect"* of attending a technical course. The first hypothesis, based on signaling theory (Spence (1978)), is that students see their employment prospects improve while studying the technical course and decide to drop out even

[40] A student may attend high school in a full time school and the supportive pole in another full time school. We are not able to confer whether the student's academic full time school is the same as the one where the supportive pole is located.

[41] As the survey was applied to both modalities (*concurrent* and subsequent), the answer "*have found a job*" is most likely driven by students in subsequent modality, given that 91% of students in *concurrent* modality does not work.

before completing it[42]. A second possibility is that students might feel more motivated or engaged to continue studying (Ryan (1998)) and drop out to attend tertiary education. Dougherty (2016), for example, shows that students more exposed to a career and technical education program in high school in Arkansas are more likely to enroll in a two-year college.

The third main cause of drop out pointed out by the interviewees indicates that the supportive pole may play an important role in a student's decision to drop out – as 12% reported dissatisfaction with this initiative ("*I did not like the Supportive pole*").

[42]Daniel Barros, coordinator at the State Secretary of Vocational Education of São Paulo, observed in a Webinar about the future of technical education, that students from technological courses in São Paulo that were able to find a job still during the technical course have higher chances to drop out the course before finish it. See in *Webinar Desafios e o futuro do ensino técnico e profissional no Brasil*.

Table 18: Online technical education enrollment status, class shift, and track

	Population of students in online technical courses (*subsequent* and *concurrent* modalities) (1)	Matched sample of students in online technical courses (*concurrent* modality) (2)
Number of students	60566	3897
Panel A: Concurrent online technical education enrolment status and class shift		
Active enrollment in concurrent	0.145	0.143
	[0.352]	[0.35]
Canceled concurrent	0.064	0.045
	[0.244]	[0.207]
Concluded concurrent	0.158	0.118
	[0.364]	[0.322]
Dropped out concurrent	0.632	0.693
	[0.482]	[0.461]
Concurrent morning shift	0.047	0.036
	[0.016]	[0.186]
Concurrent afternoon shift	0.298	0.214
	[0.457]	[0.41]
Concurrent night shift	0.655	0.750
	[0.475]	[0.432]
Panel B: Concurrent online technical education tracks		
Business Administration	0.208	0.321
	[0.405]	[0.467]
Librarian	0.050	0.009
	[0.218]	[0.095]
Interior design	0.042	0.026
	[0.201]	[0.159]
System developer	0.166	0.218
	[0.371]	[0.413]
Logistics	0.099	0.071
	[0.298]	[0.257]
Multimedia teaching	0.018	0.001
	[0.133]	[0.022]
Human resources	0.132	0.099
	[0.337]	[0.299]
School secretary	0.060	0.028
	[0.238]	[0.166]
Workplace safety	0.217	0.223
	[0.411]	[0.416]
Restaurant and bar attendant	0.009	0.003
	[0.093]	[0.05]
Student concurrent track - STEM careers	0.472	0.611
	[0.499]	[0.487]
Student concurrent track - humanities careers	0.303	0.164
	[0.459]	[0.37]
Student concurrent track - general services careers	0.225	0.226
	[0.417]	[0.418]

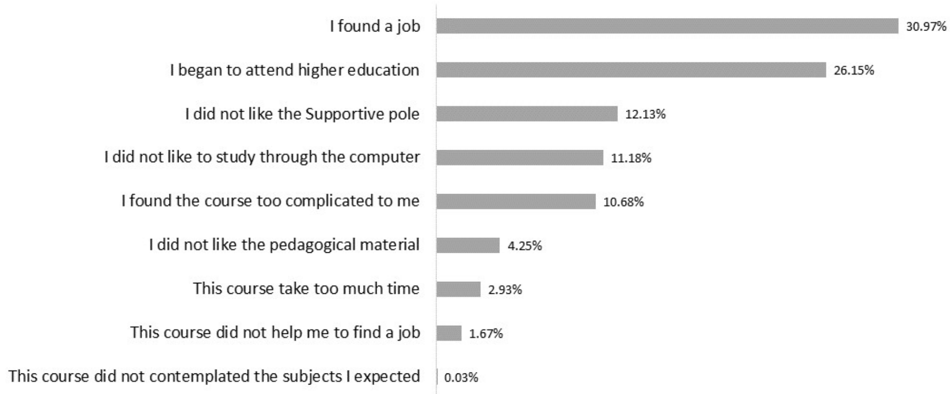
Notes: We matched administrative data of all students enrolled in any online technical course between 2012 and 2017 with our final dataset. Column (1) includes the total applicant population that passed the admission exam and enrolled in any online technical course. Column (2) considers all matched students between our final database and administrative database. It excludes students who were in private school in the 9^{th} grade and those who applied to courses that were not oversubscribed. To keep only one observation per student, we restricted our analysis to the 2^{nd} year of high school. Missing information accounts for 4% of data in column (2) and was omitted.

Table 19: Probability of dropping out and completing a *concurrent* online technical course.

	Student **dropped out** from a *concurrent* online technical program	Student **completed** a *concurrent* online technical program
	(1)	(2)
Age	0.0289** (0.0142)	-0.0249** (0.0100)
Male = 1	-0.0496** (0.0224)	0.0133 (0.0160)
Mother completed high school	0.0703*** (0.0221)	-0.0421*** (0.0159)
Household is a beneficiary of Bolsa Família	0.0168 (0.0222)	0.0154 (0.0158)
Student attends a part time high school	-0.0481 (0.0433)	0.0564* (0.0292)
Student attends a full-time high school	-0.0491 (0.0440)	0.0210 (0.0298)
Student attends a semi full-time high school	-0.00603 (0.0432)	0.0153 (0.0288)
Student started concurrent pgm in 3rd year of HS	0.00547 (0.0232)	0.00255 (0.0166)
Online technical course - STEM (Services omitted)	-0.165 (0.192)	0.180*** (0.0307)
Online technical course - Humanities (Services omitted)	-0.316 (0.206)	0.204*** (0.0615)
Normalized math 9th test score	-0.00988 (0.0133)	0.0149 (0.00957)
Normalized Portuguese 9th test score	-0.0290** (0.0131)	0.0229** (0.00906)
Student completed high school	0.00326 (0.0627)	0.0289 (0.0427)
Supportive pole based in full-time or semi full-time school	-0.0201 (0.0233)	0.0426** (0.0175)
Admission exam year FE and concurrent track FE	Yes	Yes
Constant	0.473 (0.318)	0.331* (0.178)
Total observations	1,876	1,876
R-squared	0.051	0.052

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regressions above include all students we could merge from administrative data of enrolled students and our final sample. Not all students have information on 9^{th} grade standardized test scores, which explains the lower number of total observations in each regression compared to the total number of enrolled students we have in our final sample.

Figure 4: Drivers behind the decision to drop out from an online technical course



Source: Dutra P. (2017). This graph is the result of a survey carried out in 2017 by the State Secretary of Education of Pernambuco. The survey was administered to students who dropped out of any online technical course in 2017.

2A. Attrition

One concern for our analysis is that 6.6% of students who do not drop out of our sample do not have information on one of the outcomes of interest (SAEPE test score) in 3^{rd} year of high school, as reported in Table 20 bellow. There are three main reasons why this information may be missing: some schools do not participate in SAEPE, despite the effort to universalize SAEPE, some small schools still do not participate; the student did not attend school on the day the SAEPE standardized test took place; or, the student was attending a private or federal school. These school networks are not included in Pernambuco's evaluation system. As presented in Table 20, most attrition in our sample is explained by students not showing up to school on the day SAEPE took place (5.1% of students in 3rd year of high school), followed by students enrolled in federal or private high schools (1%) and, finally, by students enrolled in state schools that did not participate in SAEPE (0.5%).

As 6.6% of students in our sample in 3^{rd} year of high school attrit the SAEPE standardized exam, we further analyze whether attrited students present similar characteristics around the admission cutoff, otherwise, this may bias our results. First, the probability of attriting is not statistically different around the cutoff, as shown Table 21. Second, differences in characteristics for students who attrited are, generally, balanced. As reported in Table 22, Column (4), students who attrit our sample have a lower probability of living on paved streets - indicating worse socioeconomic conditions. Finally, Table 23 indicates that among students who attrit, sociodemographic characteristics are balanced around the cutoff. This evidence suggests that attrited students from 3^{rd} year of high school on the standardized test in 3^{rd} year of high school are unlikely to bias our results.

Table 20: Student attrition from 3^{rd} year of HS to standardized test in 3^{rd} year of high school.

	2012.2 (1)	2013.1 (2)	2013.2 (3)	2014.1 (4)	2014.2 (5)	2015.1 (6)	2015.2 (7)	2016.1 (8)	2016.2 (9)	2017.1 (10)	2017.2 (11)	**All years** (12)
Total students	800	1,319	2,540	2,726	1,145	1,351	1,054	1,136	1,470	1,639	1,748	16,928
Total attrition from 3^{rd} to standardized test in 3^{rd} year of HS	0.4%	2.1%	1.9%	10.6%	10.4%	11.0%	10.6%	5.3%	5.9%	3.9%	6.4%	6.6%
School did not participate in standardized test in 3^{rd} year of HS	0.0%	0.2%	0.1%	0.8%	0.9%	0.7%	0.6%	0.2%	0.3%	0.7%	0.3%	0.5%
Student was in a private or federal school and did not take the standardized test in 3^{rd} year of HS	0.2%	0.4%	0.2%	1.4%	1.5%	1.6%	1.6%	1.9%	0.8%	0.6%	1.2%	1.0%
Student did not take standardized test in 3^{rd} year of HS	0.2%	1.4%	1.7%	8.5%	8.1%	8.7%	8.5%	3.2%	4.8%	2.6%	5.0%	5.1%

Notes: This table reports the percentage of students in our sample by year and semester in the applicant lists that leave the sample by not taking the standardized exam in the 3^{rd} year of high school. The sample excludes students who were in private school in the 9^{th} grade and students who apply to technical courses and schools that are not oversubscribed. HS: high school

Table 21: Change in the probability of attriting for students above the admission exam cutoff.

	Student in 3rd year of HS without standardized test information
	(1)
Above the cutoff	0.0017 (0.0092)
Above the cutoff x distance to score	0.0024 (0.0024)
Distance to score	0.0008 (0.0008)
Admission exam year FE	Yes
School FE and concurrent track FE	Yes
Total observations	12,661
R-squared	0.0433
Adjusted R-squared	0.0355

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Regression includes concurrent track, and application year and semester fixed effects.

Table 22: Differences in characteristics for students who attrited in our sample.

	Male	Age	Mother completed University	Student lives in a paved street
	(1)	(2)	(3)	(4)
Attrited from 3rd year to standardized test in 3rd year of HS	-0.0143	0.0348	0.0312*	-0.0658***
	(0.0138)	(0.0236)	(0.0187)	(0.0175)
Total observations	15,370	15,371	12,225	12,658
R-squared	0.0001	0.0001	0.0002	0.0012

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regressions above include the full sample.

Table 23: Balance of student characteristics within students who attrit from 3^{rd} year of high school to standardized test in 3^{rd} year of high school.

	Male		Age		Mother completed high school		Mother has some university education		Normalized math 9th test score	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Above the cutoff	-0.0729	-0.1701	0.2043*	0.1713	0.1141	-0.0132	-0.0284	-0.0847	-0.1594	-0.4018
	(0.0891)	(0.1233)	(0.1147)	(0.1916)	(0.0772)	(0.1710)	(0.0623)	(0.0919)	(0.1634)	(0.3075)
Flexible linear spline	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Admission exam year and semester FE and concurrent track FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total observations	221	454	335	454	259	259	189	259	220	300
R-squared	0.4152	0.2609	0.3327	0.2788	0.3687	0.3721	0.4508	0.4039	0.3608	0.338
Adjusted R-squared	0.1127	0.0951	0.1294	0.117	0.1148	0.1099	0.1396	0.155	0.0541	0.1084
Control mean in bandwidth	0.503	0.449	16.604	16.668	0.431	0.431	0.104	0.097	0.738	0.655
Bandwidth	1.8708	3	2.0795	3	3.1587	3	2.7003	3	2.7641	3

Notes: *** p<0.01, ** p<0.05, * p<0.1. Robust standard errors are in parentheses. Each column represents a separate regression. The regressions above use the bandwidth suggested by the procedure in Imbens and Kalyanaraman (2012) for the first regression of each outcome and 3, roughly twice the former bandwidth for the main outcomes for the second regression of each variable. Additionally, they include application school, concurrent track, and application year and semester fixed effects. They exclude students who were in private schools in the 9^{th} grade and students who applied to technical courses and schools that were not oversubscribed, where there is no cutoff.