

AN UNINTENDED CONSEQUENCE OF VOCATIONAL SCHOOLS' CONSTRUCTION: EVIDENCE FROM CEARÁ

Authors

Rafael B. Barbosa¹
Universidade Federal do Ceará

Alessandra Benevides
Universidade Federal do Ceará

Francisca Zilânia Mariano
Universidade Federal do Ceará

RESUMO

Este artigo analisa o efeito dos pares sobre os resultados educacionais de estudantes do ensino médio. Para identificar a relação causal, utiliza-se uma variação exógena na qualidade dos estudantes das escolas regulares no Ceará, decorrente da expansão das escolas profissionalizantes. Escolas profissionalizantes realizam um processo seletivo dentre os estudantes do primário e com isso afetam a qualidade dos estudantes das escolas regulares. Os resultados sugerem que reduzir a proporção de pares com elevado desempenho afeta negativamente as notas em testes padronizados de matemática e português dos estudantes das escolas regulares. Além disso, a qualidade dos pares impacta sobre a evasão escolar e a taxa de repetência durante o ensino médio. Tais resultados parecem ser explicados pelo aumento da distração dos estudantes, diminuindo seu foco nos estudos.

Palavras-Chave: Efeito dos pares, efeito de equilíbrio geral, Ensino Médio

ABSTRACT

This paper investigates the effect of reducing peer quality on high-school students' outcomes. To identify the causal effect, we exploit an exogenous variation in peer composition caused by a large program of selective public schools in Brazil. Results suggest that students exposed to an increase in the proportion of low-achievers diminish test scores, both math and language, and raise the rate of retention during secondary education. The channels that explain the underlying findings are associated with an increase in students' attention diversion and a reduction of persistence in learning. We do not find evidence that the teacher labor market and teacher pedagogical practices are affected by the change students' composition.

Keywords: Peer effect, General equilibrium effect, Secondary education
JEL: I2, I28, I24

ÁREA 4: Setor Público

¹ Email: rafael.barbosa@ufc.br/ telefone para contato: 85 99930-1938

1. Introduction

Social interactions are a fundamental determinant of several economic outcomes². However, the effect of peer in education, notably on secondary education, is fiercely debated. Peers can influence many aspects of an individual's economic decisions, especially during adolescence, when the student is subject to several behavioral biases. Furthermore, it is not clear whether high-achievers peers positively affect students, neither the mechanisms that could explain the potential effects. These debates are, in part, related to econometric identification's problems as self-selecting (sorting), endogenous, and reflection bias (Manski, 1993, Angrist, 1994). Recent literature has followed a series of new identification strategies to overcome these limitations.

In this paper, we consider a plausible quasi-natural experiment that changes the peer composition for some secondary schools in Brazil. Specifically, using a difference-in-difference empirical strategy, we find that increasing the proportion of low-achievers affects the students' test scores and other educational outcomes, as the rate of retention and drop-out. To measure exposure to low-achievers peers, we exploit a large state's program of selective public schools' construction in the Ceará, a Northeast state of Brazil. Since 2008, the state of Ceará has expanded the number of vocational schools (*Escolas Profissionalizantes*) relative to regular ones³. Unlike other secondary schools in Brazil, vocational schools in Ceará execute an admission process that selects high-achievers from middle education (Ensino Fundamental II). This selection process separates high and low-achievers during secondary education⁴ in municipalities that receive vocational schools.

The construction of vocational schools significantly reduces the proportion of high-achievers students attending regular education. We assume that the timing, location, and the share of students attending vocational schools are orthogonal to unobservable regular students' factors. We extensively test the validity of these assumptions. We use the proportion of students attending vocational schools as a measure of exposure to low-achievers peers for regular students.

Theoretically, the impact of change the peer composition is ambiguous (Duflo, Dupas, and Kremer, 2011; Lazear, 2001). Unfortunately, empirical pieces of evidence are also controversial. Lavy, Paserman, and Schlosser (2012) estimate that the proportion of repeaters harms the academic achievements of peers. Carrell, Hoekstra, and Kuka (2018) find similar results, indicating that the presence of disruptive students in elementary schools negatively affects long-run outcomes, as earnings and college attendance. However, Duflo, Dupas, and Kremer (2011) point out that homogeneous classrooms raise students' test scores, even with a large number of low-achieving students. In this case, the mechanism is explained by teachers' adjustment of pedagogical practices. We expect to shed light on this controversy providing robust empirical evidence of the mechanisms which peers potentially affect other students.

This paper has two main contributions. First, it estimates the effect of reducing peer quality on regular students' outcomes exploiting a plausible exogenous variation in the share of students enrolled in vocational schools during secondary education. We consider as the main outcome variable the student's test scores in math and language at the end of secondary education (12th grade). Moreover, we also assess the effect on long-run outcomes as the rate of student

² There are many evidences about peer effect in different areas as health (Fortin; Yazbeck, 2015), crime (Billings; Deming; Ross, 2016), job productivity (Georgeanas *et al.*, 2015, Falk; Ichino, 2006), and education (Sacerdote, 2014).

³ In Brazil, regular schools teach the standard high school curriculum and generally have four or five class hours per day.

⁴ In addition, vocational schools are preferred by parents and students in comparison to regular ones. Parents prefer vocational schools for two reasons. First, vocational school is full-time, contrary to the regulars, which is part-time. Therefore, parents can leave the students under the school's supervision during all workday. Second, vocational schools provide vocational skills, increasing the probability of students enters the labor market after the secondary. In some municipalities of Ceará, the supply of tertiary education is limited, then attending vocational schools can improve the economic opportunity.

drop-out and retention using a school-level version of the empirical strategy. Both outcomes are relevant to human capital accumulation and may affect the economic opportunities of the student's life.

Second, we extensively attempt to understand the mechanisms that explain underlying findings, analogously to Lavy, Paserman, and Schlosser (2012). We consider several potential channels as student engagement, interest in learning, student well-being, time allocation, personality traits, and teacher's pedagogical practices. These channels are related to main theories that explain how peers affect her colleagues.

The results suggest that the share of students attending vocational schools negatively affects the academic achievements of regular students. The estimate effect size is large in comparison with other educational policies. The magnitude of the impact in math is higher than in language test scores; however, the difference is small.

We also test the presence of a heterogeneous effect. First, we find that girls are more affected by peers' change in composition than boys in math test scores. Second, students residents in small municipalities are less impacted by the reduction of peer quality than students living in big cities. Third, the results also suggest the peer effect is highly nonlinear. Students with a low performance before secondary education (9th grade) are more negatively affected by their peers. Interestingly, we find a positive effect of the change in peer composition in high-achievers' students.

Furthermore, a concentration of low-achieving students raises the rate of drop-out and retention during secondary education. The impact on drop-out is more prominent in the small municipalities. This paper also contributes to evidence about the effect of peers on long-run outcomes as reported by Carrell, Hoekstra, and Kuka (2018), Gould, Lavy, and Paserman (2009), Bifulco, Fletcher, and Ross (2011), Bifulco et al (2014), Anelli and Peri (2017). Results for both short and long-run outcomes are submitted several robustness' checks, and we conclude that our estimates are unlikely to be biased.

We then move to understand the channels that explain the underlying estimates, examining teacher and student surveys. Results suggest that girls exposed to a high proportion of low-achievers reduce the expectation about success in the future and increase the time spending on the internet. Boys are not affected by the change in peer composition. These conclusions support the evidence of large peer effects on girls. We do not find evidence that changes in peer composition impact teachers' pedagogical practices. We confirm our results using a teacher survey about her pedagogical practices and student perception about teachers.

At last, we verify the effect of reducing peer quality on students' personality traits. We measure the personality traits using a Big-Five instrument, called SENNA Instrument, implemented by Instituto Ayrton Senna, a Brazilian civil society organization focusing on citizen agency and public-service delivery. The SENNA Instrument measures the five constructs of the Big-Five inventory⁵ and the facets that make up each construct. We estimate the effect of vocational school constructions against the facets of Conscientiousness and Openness, both associated with educational achievement. Our estimates suggest that vocational school This paper contributes to three strands of literature. First, we contribute the literature about the importance of peer composition for student outcomes as Lazear (2001), Duflo et al (2011), Lavy et al (2012), Imberman et al (2012), Carrell et al (2018), Shiltz et al (2019), Bossavie (2020), and others. By using a quasi-experimental approach, we cope with the main econometric challenges in peer effect estimation. Also, our measure of the share of low-achievers does not depend on specific students' aspects, as repeaters (Lavy et al (2012)) or disrupt contemporaneous behavior and learning (Carrell et al (2010, 2018)) that may cause difficulty in the effect interpretation. Second, a few papers study how education policies might have general equilibrium effects on

⁵ Big Five Inventory is a five-factor model of personality traits formed by: Conscientiousness, Openness, extraversion, agreeableness, and neuroticism. For more details see Almond et al () and the Appendix A.#.

student outcomes (Duflo (2004), Bianchi (2020), Gilraine et al (2018)). This paper complements this literature by showing how the expansions of vocational schools have unintended spillover effects on students of different schools. Third, this paper sheds light on the channels through the social interactions impact students. A few papers study these mechanisms, especially in secondary schools (Lavy et al (2012)). We add to these studies by evidencing that the main channels of peer effect are related to diverting student attention, expectations about the future, and time allocation.

2. Backgrounds

2.1 Vocational schools in Ceará, Brazil

The state of Ceará is located in the Northeast of Brazil, one of the poorest regions in the country. Ceará's per-capita GDP is nearly USD 5,500, which is smaller than the average per-capita GDP in Brazil (close to USD 9,800). The state's population is approximately 9 million, and the Human Development Index (HDI) is 0.68, close to Brazil, 0.69⁶.

Secondary education in Ceará is segregated into three kinds of schools: regular, vocational, and indigenous education.⁷ Regular schools are part-time education and teach only the standard secondary nationwide curriculum, providing skills to students that aim to apply to higher education at the end of secondary. Indigenous schools teach only a specific indigenous curriculum and represent a small share of secondary school in Ceará, nearly 1%.

Since 2008, the government of Ceará has expanded the number of vocational schools. In 2017, there existed 120 vocational schools in Ceará, representing nearly 17% of the secondary education in the state. Vocational schools have remarkable differences comparing with regular ones. First, vocational schools have a full-time program and two different curriculums. Half of the day class, students learn a standard nationwide curriculum, similar to regular counterparts. In the other half time, students learn practical knowledge and occupation-specific skills that directly map into entering a particular occupation in the labor market. Second, and more relevant for our purposes, vocational schools in Ceará perform an admission process that cherry-picks the high-achievers. This admission process reduces the proportion of high-achievers in regular schools during secondary education. In general, vocational students have better socioeconomic characteristics, higher levels of cognitive and non-cognitive aspects than regular ones.

The admission process differentiates vocational from regular students in at least two aspects. First, vocational students show intrinsic motivation to study in these schools because they have to apply to compete for a slot. Demotivated students do not apply and do not participate in the selection process. Second, the admission process selects students based on high-stakes performances over the last three academic years. Therefore, vocational students have higher grades and also sustained their performance for a long time in comparison with other students. This aspect is associated with persistence and grit (Duckworth et al., (2009)).

Table 1 summarizes the differences between regular and vocational students. We consider test scores⁸ and many other demographics and socioeconomic aspects as the proportion of female, age, race, mother's education, and the share of students that at least one family member a conditional cash transfer (CCT) program. The average scores of vocational students are 27.32 and 13.21 points higher than regular students in math and language, respectively. Regular students are represented by a smaller proportion of girls and white students. In turn, vocational students are younger, their mothers are more educated, and they live in families that have a lower

⁶ To put in perspective, Ceará's GDP is similar to countries like Nicaragua, Mozambique, Albania. Moreover, the Ceará population is similar to Paraguay, a Latin-America country.

⁷ There exist a fourth school category since 2018, the full-time schools. However, our data span until 2014.

⁸ We measure cognitive skills using administrative data provide by Ceará's Department of Education (SEDUC). More details are described in section #. The distribution of these tests has a mean 250 and a standard deviation of 50 points.

proportion of parents attending the cash-transfer program. These results suggest that vocational students have high cognitive skills and better socioeconomic conditions than regular students.

During the implementation process of vocational schools, we observe variation in the exposure of students over time and geographical areas. The government smoothly scaled up the program across years and municipals. Figure 1 shows the municipality evolving of vocational schools along the time between 2010 to 2017.

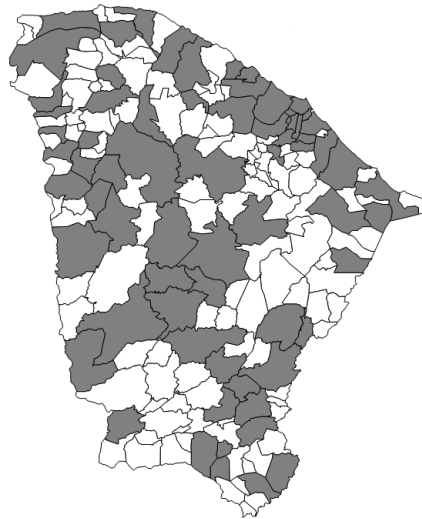
Vocational school is usually preferred by parents and students. The Ceará's state has a small offer of public higher education. Students that attend regular education and do not apply to public universities have few opportunities to enter the labor market. Vocational schools aimed to raise the opportunities for students by providing them occupation-specific skills, allowing enter the labor market just after the secondary.

3. Data and Descriptive Statistics

3.1 Data Sources

This paper requires considerable data sources. First, we create the share of students attending vocational schools in each municipality of Ceará's state from the annual School Census, a survey of every school in Brazil, conducted by the Ministry of Education. We consider only students that attended public middle education before enrolling in high-schools, excluding students from private schools. The School Census data also allows measuring some school quality indicators, as the proportion of teachers with higher education and the average number of students per classroom.

FIGURE 1: Vocational School construction in Ceará



Notes: Figure 1 presents the variation in time and geographic location of the expansion of vocational schools in Ceará's state, Brazil. The first vocational school was introduced in 2008 and in 2017 there exists 120 schools in Ceará, representing 17% of all secondary schools.

To measure test scores in math and language, we consider the administrative data from SPAECE⁹, a state test applied by Ceará's Department of Education (SEDUC). The SPAECE data span from 2008 to 2014 and include individual test scores for students in the 9th grade of the middle and 3rd grade of high-school (9th and 12th grades, respectively). The data contains 550.867 students at the end of secondary (12th grade), where nearly 93% of students attend regular schools.

⁹ Sistema Permanente de Avaliação da Educação Básica do Ceará (SPAECE).

The Ceará's Department of Education also applies, during SPAECE, a contextual survey to students that allows construct student demographic and socioeconomic characteristics. In some cases, we were not able to match student test scores to student's survey data because of misreporting. We test different specifications of empirical strategy to cope with this missing information.

The contextual survey also provides information about student engagement, perception about teacher practices, student well-being, and time allocation. We use this information to understand the potential mechanisms of the peer effect. SPAECE also applies a survey for teachers. We use the teacher survey to understand whether their pedagogical practices are associated with vocational school construction¹⁰.

We verify if the construction of the vocational school is associated with the student's personality traits. To measure the personality traits, we consider an instrument applied on census-basis to Ceará students only in 2015. The instrument, called the SENNA instrument, was conducted by the Instituto Ayrton Senna, a Brazilian civil society organization. This instrument allows capturing the Big-Five personality traits and its facets. We focus our estimates on the Conscientiousness and the Openness' facets because are more related to students' performance, according to literature. The facets of Conscientiousness are Determination, Focus, Organization, and Persistence. In turn, the facets of Openness are Curiosity to Learn, Creative Imagination, and Artistic Interest. Our results about personality traits should be seen as preliminary, however, since this relies on data for a single year.

Finally, we consider as long-run outcomes the rate of school drop-out and the rate of retention during high school. These variables are constructed from the Ministry of Education at the school level. Both variables represent the percentual of students that dropped out or repeated during one of the three years of secondary schools.

3.2 Descriptive Statistics

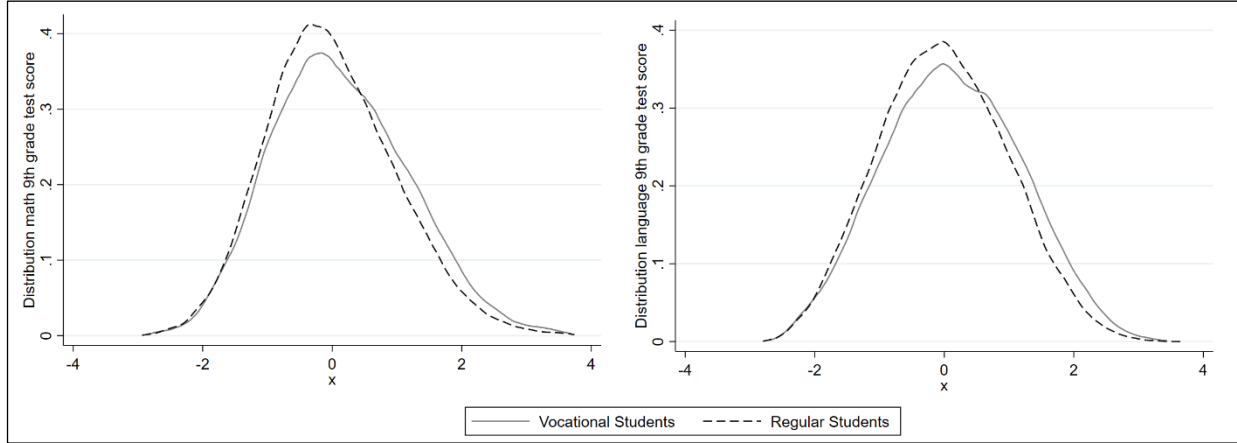
Based on the admission process, we posit that students attending vocational schools have better performance in comparison with regular students. Figure 2 displays the distribution of test scores in math and language at the ending of grade 9, before they enter in high-school. The scores are standardized to have mean zero and standard deviation one. As suggested by this graph, vocational students outperform regular students in math and language test scores. This evidences that regular students have low-abilities compared with vocational students before they attend vocational schools.

We are interested in the causal effect of high achieving peers on students left behind in regular education. Hence, the treatment group is composed of regular students exposed to vocational schools. Table A2, reported in appendix, compares treated and untreated students, i.e. the regular students that were not expose to vocational school construction.

Treated and untreated students are very similar. There is no statistical difference among scores in math and language at the end of middle education, grade 9. Treated students outperform untreated students on language scores at the ending of secondary. Untreated students are represented by a large share of black and brown and are younger than treatment group. In turn, the treated students are less poor, and their mothers have more years of instruction.

¹⁰ We do not have information about these surveys for all cohorts. Many questions change along the cohorts and we consider items with large size possible. Details of data, see Appendix A#

FIGURE 2: Distribution of math and language test scores in 9th grades



Notes: Figure 2 presents the distribution of math and language test scores in 9th grades for students that in the next year will attend regular (black) and vocational (gray) schools. It shows that vocational students have better performance in both subjects before enroll in secondary education.

In short, treated and untreated students has similar characteristics and we expect that the exposure to vocational schools is the only factor that differentiates short and long-run educational outcomes between the two groups.

Table 1: Descriptive statistics comparing vocational and regular students

Variables	Vocational Students		Regular Students		Difference
	Average	SD	Average	SD	
Test Scores Math 9 ^o grade	255.74	47.94	228.42	45.71	27.32***
Test Scores Portuguese 9 ^o grade	249.88	48.61	236.67	44.67	13.21***
Girls	0.57	0.49	0.55	0.49	0.02***
Black	0.11	0.31	0.14	0.35	-0.03***
Brown	0.11	0.31	0.12	0.33	-0.01
White	0.17	0.38	0.08	0.28	0.09***
Age	14.59	1.60	15.34	2.80	-0.75***
Mother education	2.94	1.07	2.85	1.07	0.09***
Attend Bolsa Familia	0.72	0.44	0.74	0.43	-0.02**

Note: Table 1 presents the descriptive statistics of vocational and regular students' factors. The last column shows the difference about the variables' averages. The stars represent the rejection of the null hypothesis of equal mean based on a test t. Significance levels: 1% ***, 5% **, 10% *.

4. Empirical Strategy

To overcome the usual problems of selection and sorting associated with the estimation of peer effects, we rely on exogenous variation in timing and location of vocational school construction. We also access the intensity of vocational school construction's exposure by using the share of students that, in some municipality and cohort, attend vocational school.

Our benchmark specification is the following

$$y_{ismt} = \beta + \gamma Share_voc_{mt} + \delta' X_{ismt} + \alpha' X_{smt} + \tau_t + \pi_s + \varepsilon_{ismt} \quad (1)$$

Where y_{ismt} is the achievement for student i , on the school s , in municipality m , on year t ; $Share_voc_{mt}$ is the share of students attending vocational schools in municipality m in year

t . π_s is a school effect, τ_t is a time effect, X_{ismt} is a vector of students controls that includes gender, race, age, mother's education, an indicator that a family is a recipient of a conditional cash transfer program (CCT), 9th grades test scores in math and language, polynomials third-order of these test scores, an indicator that student usage public transportation to go to school. X_{smt} is a vector of school covariates that includes the proportion of teachers with tertiary education and average class size. When presenting our estimates, we show different combinations of these covariates. ε_{ismt} is the error term. Standard errors are clustered at the school level.

Our parameter of interest is γ , which measures the effect of the regular students being exposed to an increase of the proportion of high-achievers attending vocational schools. We assume that the rise of the share of students attending vocational schools implies a reduction of the quality of peers in the regular schools, measured by the proportion of low-achiever's students.

To control for potential confounding factors, we include in all specifications school and time fixed effects. However, one may be concerned that there are time-varying unobserved factors that are also correlated with the proportion of low-achieving students at school level. Therefore, we also estimate a model adding a full set of school-specific linear time trends to (1).

We also consider a school-level version of (1) that is used to estimate the effect of increase the share of low achieving students on the rate of student drop-out and retention. These variables have high opportunity costs for the students suggesting that the peer quality can have a persistent effect.

We estimate the following equation

$$y_{smt} = \beta_0 + \mu Share_voc_{mt} + \vartheta' X_{smt} + \tau_t + \pi_s + \varepsilon_{smt} \quad (2)$$

Where y_{smt} is the proportion of repeaters or students that dropped out the regular school s in the year t in the municipality m . Important, we consider these variables for all high school grades (10, 11, and 12 grades). In the Online Appendix, we estimate the same specification for the three grades of high-school separately.

To control for school potential confounding variables, we include specific school controls, X_{smt} , and school fixed effect. X_{smt} contains the proportion of teachers with tertiary education and average class size. Therefore, we also consider a model that includes school interacted with time fixed effects to capture time-varying school-specific unobservable factors.

4.1 The validity of the identification strategy

The validity of identification strategy depends on two key assumptions: (1) time and location of vocational school construction are exogenous to regular students, (2) the share of students attending vocational schools is unrelated to regular students' unobservable factors.

The first assumption is not validated if the decision about the time and location where constructed a vocational school is related to students' unobservable characteristics. To assess this possibility, we regress a logit panel model where the dependent variable is the time and location of vocational schools against a vector of controls that includes average students age, proportion of girls, racial shares, and average previous test scores in language and math for 9th and 12th grades. The results do not suggest that the decision to construct a vocational school is reasonably associated with these factors¹¹. The demographic factors are not significant and the estimates for average grades present contradictory implications. For example, the results indicate that vocational school constructions are correlated to municipalities with higher average language test scores and lower test scores in math.

¹¹ The results of these estimates are not reported in this paper by concision; however, it is available under the authors' contact by emails.

In turn, the share of students attending vocational schools are unrelated with regular students' unobservable factors is plausible assumption because the decision to be enrolled in a vocational school is taken before the secondary education. Thus, students accepted to enroll in vocational school do not previously know the quality of future peers.

A potential threat to this assumption is the students' capacity to anticipate the quality of future peers in high school based on the quality of current peers in middle schools. We consider this possibility unlikely for two reasons. First, students probably change the school during the transition from middle to secondary education because the number of middle schools is much higher than secondary education¹². Second, many students dropped out during the transition to middle to secondary education¹³. Both reasons affect the capacity of predict the quality of the peers on secondary.

The contributions of Manski (1993) has evidenced the fundamental problem of selection into peer groups which can contaminate peer effect estimates. First, students may self-select themselves into peer groups based on certain unobserved factors, called correlated effects. Second, peers may influence each other simultaneously, known as reflection problem. Third, it is difficult to distinguish between peer effects due to peers' achievement, endogenous effects, and peer effects due to peers' background, contextual peer effects.

Our empirical strategy overcomes the first and second fundamental problems. The construction of the vocational school has a municipal level impact, i.e. which is not school-specific. This minimizes the self-selection process, especially the exposure of the regular students to high-achievers. In addition, the inclusion of school fixed effects accounts for the most obvious source of student sorting between schools.

Table 2: Change in the students' characteristics

Table 2: Change in the students' characteristics						
Panel A: All municipalities	(1)	(2)	(3)	(4)	(5)	(6)
	Girls	Black	Brown	Age	EJA	Urb
Treat	-0.00304 (0.00187)	-0.0016 (0.00117)	-0.0081 (0.0135)	0.021 (0.044)	0.00609* (0.00333)	-0.00349 (0.00539)
Observations	2,601,977	2,601,977	2,601,977	2,601,977	2,601,977	2,601,977
R-squared	0.006	0.014	0.202	0.294	0.774	0.37
Panel B: Small municipalities	(1)	(2)	(3)	(4)	(5)	(6)
	Girls	Black	Brown	Age	EJA	Urb
Treat	-0.00132 (0.00218)	-0.00145 (0.00142)	0.000377 (0.0151)	0.0191 (0.0544)	0.00664 (0.00414)	-0.00732 (0.0071)
Observations	1,782,016	1,782,016	1,782,016	1,782,016	1,782,016	1,782,016
R-squared	0.005	0.012	0.2	0.239	0.687	0.214
School FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y

Notes: Table 2 presents the impact of vocational school construction on municipal students' factors. It evidences that only the proportion of students using public transport is significantly affected by the vocational school program. Each estimation includes municipal and time fixed effects. Significance levels: 1% ***, 5% **, 10% *.

The last concern is the possibility of students migrate when vocational schools are constructed. If the future regular or vocational student migrates to a different municipality when the vocational school is constructed, then the results could be biased. To test this possibility, we estimate the effect of the vocational school construction on the municipal students' demographic

¹² Particularly in Ceará, there existed 717 secondary schools and 4326 middle schools in 2017.

¹³ The rate of school progression from middle to secondary education is 86%.

and socioeconomic characteristics, like the proportion of girls, the racial shares, age, special educational programs¹⁴ (EJA), and the share of urban students. If the vocational school introduction produces a relevant migration process, we expect that municipal students' characteristics will be affected. Table 2 presents the estimates. All models include municipal and time fixed effects.

Panel A shows the estimates for all municipalities and Panel B consider only small municipalities¹⁵. We do not find evidence that vocational school construction affects the student demographic characteristics, suggesting an absence of migration caused by the presence of selective schools. In summary, the results corroborate the validity of the empirical strategy's assumptions.

5. Results

5.1 Effects on test scores

Tables 3 and 4 report the effect of the proportion of vocational students on high school achievement of regular students in math and language, respectively. We transform SPAECE's test scores into standardized z-scores to facilitate the interpretation of the results. We consider six specifications in which the differences stem from the additional covariates. The presence of missing data in the sample reduces the sample size as more covariates are incorporated.

Column 1 presents the average treatment effect and standard deviation of the outcome variables for regular students considering only additive school and time fixed effects. This sample has nearly half a million students at the end of high-school, spanning from 2008 to 2017. The columns 2-5 include additional controls. Column 2 considers regular controls as gender, age, racial status, and an indicator of the use of public transport. Column 3 adds 9th test scores in math and language to control for previous skills. Column 4 adds as control an indicator for students' families that receive conditional cash-transfer (CCT) and mother's education. Finally, Column 5 includes the third-order polynomials for math and language 9th test scores. Column 6 considers the same specification of column 5, however, it adds the school-by-time fixed effects which controls for potential time-varying unobservable factors.

Results show that all estimates are negative and statistically significant. This suggests that the share of students attending vocational school reduce the performance of students from regular schools. The estimates do not change in magnitude in different specifications, except for Column 3, which includes previous students' test scores. This result can be associated with the sample reduction. However, in more restrictive samples, as Columns 4, 5, and 6 the magnitude of initial specifications is restored.

The average effect size is -0.38σ for language and -0.40σ for math considering three years of exposure to low-achievers peers. The impact of reducing peer quality affects mathematical and reading skills, suggesting that the global effect on student ability can be larger¹⁶. To put these estimates into perspective, considering that, in 2017, 17% of secondary students in Ceará are enrolled in a vocational school, then the average effect is near -0.06σ for math and language.

¹⁴ Older students can opt to be enrolled in special educational programs (EJA) that reduce the time of secondary education.

¹⁵ We refer to a small municipality if their population is lower than 250 thousand inhabitants. Only three municipalities have population higher than 250 thousand inhabitants: Fortaleza (State's capital), Caucaia, and Juazeiro do Norte.

¹⁶ In Online Appendix, we present the estimation on the sum test scores. The point estimate is -0.79σ for our preferred specification similar to column 5 in both tables.

Table 3: Estimates of the proportion of vocational students on math achievement of regular school

Math	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.410*** (0.065)	-0.347*** (0.076)	-0.165* (0.093)	-0.419*** (0.131)	-0.444*** (0.128)	-0.498*** (0.049)
School fixed effects	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y
Regular control		Y	Y	Y	Y	Y
Prior test scores			Y	Y	Y	Y
Additional control				Y	Y	Y
Non Linearities					Y	Y
School-time fixed effects						Y
N. obs.	502.920	364.260	167.499	96.176	96.176	96.176

Notes: Table 3 reports the effect of share of vocational students on math test scores of regular students. The five specification change according to the number of covariates in each model. The specifications (1) to (5) contains school and time fixed effect. The column (6) presents results for the school-by-time fixed effects. The standard errors are estimated clustering by school. Significance levels: 1% ***, 5% **, 10% *.

These results evidence that peer quality is relevant to students' achievements. Specifically, our results indicate that peer quality matters to test scores at the end of secondary education in Brazil. Although the effect size is not large, we can compare it with other studies. Jackson (2014) analyzes the effect of teacher quality on high-school students in North Carolina State, US. The impact of a decrease by one standard deviation the teacher quality is 0.06, similar to our estimates. Then, to put in perspective, the effect of reducing peer quality is equivalent to reducing teachers' value-added by one standard deviation, according to estimates from Jackson et al (2014).

Table 4: Estimates of the proportion of vocational students on language achievement of regular school

Language	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	-0.490*** (0.061)	-0.379*** (0.066)	-0.258** (0.085)	-0.330** (0.118)	-0.337** (0.116)	-0.359*** (0.048)
School fixed effects	Y	Y	Y	Y	Y	Y
Year fixed effects	Y	Y	Y	Y	Y	Y
Regular control		Y	Y	Y	Y	Y
Prior test scores			Y	Y	Y	Y
Additional control				Y	Y	Y
Non Linearities					Y	Y
School-time fixed effects						Y
N. obs.	502.920	364.260	167.499	96.176	96.176	96.176

Notes: Table 4 reports the effect of share of vocational students on math test scores of regular students. The five specification change according to the number of covariates in each model. The specifications (1) to (5) contains school and time fixed effect. The column (6) presents results for the school-by-time fixed effects. The standard errors are estimated clustering by school. Significance levels: 1% ***, 5% **, 10% *.

5.1.1 Heterogenous effects on test scores

There exists a fierce debate about the presence of non-linearity of peer effect. For example, the monotonicity model posits that the lower the quality of the peers the bigger is the

negative effect on students. Sacerdote (2011) suggests that nonlinearities in the relationship between own and peer characteristics may help explain the heterogeneity evidence about peer effects. Several studies have found evidence of nonlinear peer effects, wherein lower ability students lose more from decreases in peer quality than do higher ability students (Hoxby and Weingarth (2005), Imberman et al. (2012), Burke and Sass (2013), and Fu and Mehta (2018)). We test the presence of non-linear effects on regular students due to the peer composition changes as a consequence of vocational school introduction.

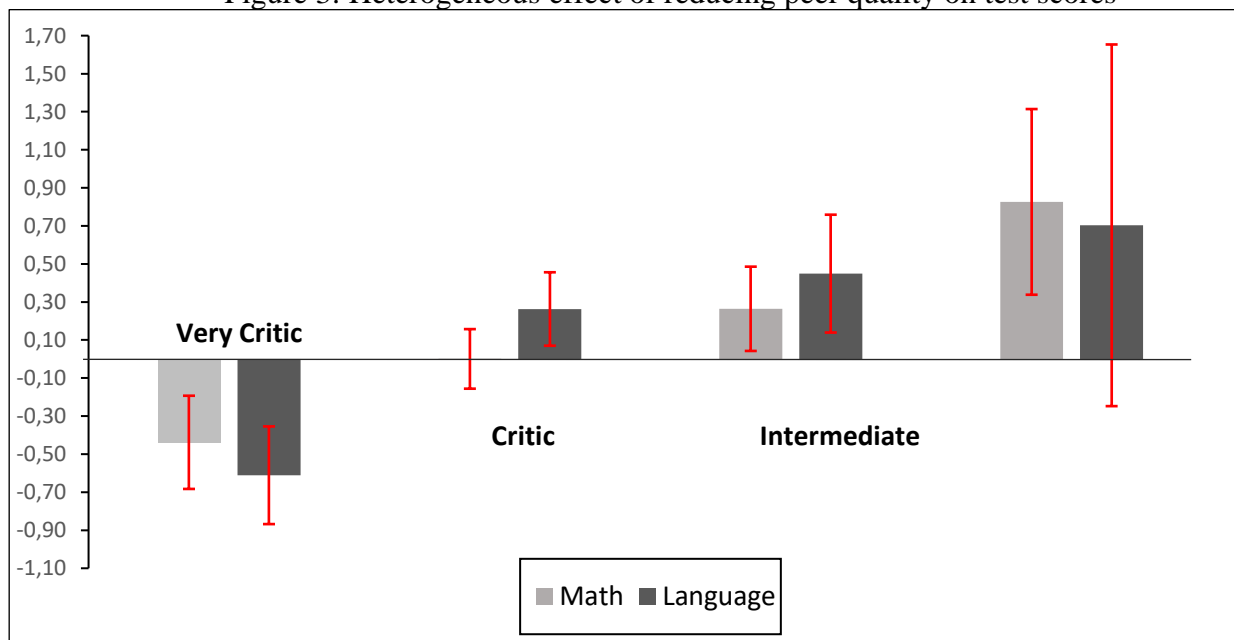
We verify the effect of being exposed to large peer quality reduction on regular students considering the student’s educational status in 9th grade. The Ceará’s Department of Education classifies the students in the 9th grade in four levels according to performance in test scores: Very Critic, Critic, Intermediate, and Advanced¹⁷.

We estimate our preferred specification that includes all covariates similar to Column 5 and 6 in Table 4. To assess the heterogeneous effect, we interact the exposure measure, *Share_voc_{mt}* with each of the student's prior educational levels, computed in 9th grade. Figure 3 depicts the effect on each student group level.

Students at a Very Critic level are strongly negative affected in both math and language. This suggests that the validity of the monotonicity model, i.e., worsen peers lower the test scores of their schoolmates. The magnitude of the effect decreases according to the educational level raise, indicating the presence of weak monotonicity (IMBERMAN et al., 2012)).

Specifically for math, students at the Intermediate and Adequate levels benefit if more students attends a vocational school. This indicates the presence of the invidious model, which posits that a student’s performance increase by having less able peers. Consistent with the previous literature, we find that changing peer composition has a nonlinear effect on their colleagues, in which low-achievers’ students are more harmed and high-achievers are positively impacted.

Figure 3: Heterogeneous effect of reducing peer quality on test scores



Notes: Figure 3 reports the heterogeneous effect of share of vocational students on language and math test scores of regular students, considering the educational level achieved by students at 9th grade. The specification includes covariates in each model. All specification contains school and time fixed effect. The standard errors are estimated clustering by school. Significance levels: 1% ***, 5% **, 10% *.

¹⁷ This approach is more suitable to testing nonlinear models using our data. The standard approach is based on quartile specifications as Imberman et al (2012) or Hoxby and Weingarth (2006). Our sample has missing values in 9th-grade, which prevents obtaining the reliable ranking of the student in 9th grade.

5.2 Effect on student’s drop-out and retention

There are relatively few pieces of evidence of the long-run educational consequences of peers. Carrell, Hoekstra, and Kuka (2018), Anali and Peri (2017), and Bifulco et al (2014) are exceptions. We attempt to assess the long-run effect of reducing peer quality by estimating the impact on the rate of student dropout and the rate of student retention. Both variables have long-run consequences for secondary students.

We consider a school-level specification where the dependent variables are the rate of students’ drop-out – i.e. the proportion of students that dropped out the school in each cohort - and the rate of students’ retention – i.e. the ratio of students who remains in the same grade in each year. The specifications include school controls as the average class size and the proportion of teachers with tertiary. We also include school and time fixed effects.

Table 5 shows the results. Column 1 refers to the impact on school dropout considering all municipalities in Ceará, and column 2 presents the effect on school dropout just for small municipals. Column 3 refers to the impact on school retention for all municipals, and column 4 shows the effect on just small municipals. We restrict the sample for small municipalities in columns (2) and (4) because the stay at school can be affected by the local labor market opportunities.

The results suggest that reducing peer quality increase the likelihood of school drop-out in 5.3 percentage points, considering small municipalities (Column 2). On average, 11,4% of secondary students dropped out during secondary education in Ceará. Thus, the estimate represents an increase of almost 50% in the rate of school drop-out on average. The impact of peer on student drop-out is large, suggesting that peer quality matters for longer-run outcomes. Nonetheless, in column 1, the effect of reducing peer quality is not significant, although it is positive.

In turn, peer quality also affects the rate of school retention significantly. As reported in columns 3 and 4, the impact of the share of vocational students on regular school retention is 4.511 percentage points considering all municipalities and 4.259 percentage points restricting to the small’s one. On average, the rate of school retention is 7.08% during secondary education. This represents a rise of approximately 63% in the school retention.

Taken together, we conclude that reducing peer quality has relevant consequences for students’ outcomes in both short and long-run. It is also interesting to note that the construction of vocational school has an unintended effect on students that were not directly affected by these schools. This highlights the importance of public policy design.

Table 5: Effect on the rate of school drop-out and retention

VARIABLES	Dropout		Retention	
	(1) All Municipalities	(2) Small Municipalities	(3) All Municipalities	(4) Small Municipalities
Share of Vocational Students	3.978 (2.637)	5.350** (2.622)	4.511*** (1.486)	4.259*** (1.539)
Observations	2,930	2,000	3,893	2,674
R-squared	0.444	0.423	0.317	0.403
Municipal FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
School Controls	Y	Y	Y	Y

Notes: Table 6 reports the impact of share of vocational students on language and math test scores of regular students, considering the educational level achieved by students at 9th grade. The specification includes covariates in each model. All specification contains school and time fixed effect. The standard errors are estimated clustering by school. Significance levels: 1% ***, 5% **, 10% *.

6. Robustness

We conduct several robustness tests to verify if the estimates are robust. First, to check if the estimates capture a spurious correlation among the share of students attending the vocational school and the regular students' outcomes, we realize a falsification test using placebo regression. In short, we verify if the share of vocational students in the municipality m , in the year t affects the students' outcomes of regular students in the same municipality, in the year $t - 1$. If the estimates are significant, then potentially the treatment effects are driven by short-run trends.

Table 6 reports the placebo test estimations. We show no significant effect on the share of vocational students on previous regular students' test scores. The exception is the column (1) for language test scores. This specification does not include any covariates. The inclusion of the control variables eliminates the significance, suggesting that the validity of the estimates is conditioning to covariates in such a case.

Next, we test if the results depend on the measure of vocational school's exposure. We do not expect that distinct measures of the vocational school's intensity yield contrasting estimates. We consider two alternative measures of vocational school's exposure. First, it is a binary variable indicating the year that some municipality received the first vocational school. In this case, the model is interpreted as a standard difference-in-difference estimation with dynamic treatment. Second, we define the proportion of vocational schools for each municipal-by-year. Table 7 shows the results for both alternatives measures. The estimates are not insignificant and the signal of the parameters is negative, suggesting that the vocational school's exposure can negatively affect the students' outcomes.

Table 6: Falsification test

<u>Math</u>	(1)	(2)	(3)	(4)	(5)
Placebo Effect	-0.024 (0.016)	-0.007 (0.014)	0.029 (0.030)	0.038 (0.029)	0.041 (0.029)
<u>Language</u>	(1)	(2)	(3)	(4)	(5)
Placebo Effect	-0.033** (0.014)	-0.022 (0.031)	-0.022 (0.027)	-0.010 (0.031)	-0.008 (0.031)
No control	Y				
Regular control		Y	Y	Y	Y
Prior test scores			Y	Y	Y
Additional control				Y	Y
Non Linearities					Y

Notes: Table 6 reports falsification test. It tests if the share of vocational students in the municipality m , in the year t affects the students' outcomes of regular students in the same municipality, in the year $t - 1$. All specification contains school and time fixed effect. The standard errors are estimated clustering by school. Significance levels: 1% ***, 5% **, 10% *.

Finally, we test if the learning incentives depend on local economic opportunities. Big municipalities yield distinct local labor market incentives in comparison with small ones, potentially affecting students' effort. To test this possibility, we also estimate a model that ignores municipalities with more than 150 thousand inhabitants¹⁸.

¹⁸ We perform additional robustness checks do not reported by concision. For more details, contact the authors by email.

Table 7: Alternative measures to vocational school exposure

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Math	Math	Math	Language	Language	Language
Binary Treatment	-0.0991*** (0.0142)	-0.0916*** (0.0143)	-0.0943*** (0.0146)	-0.0686*** (0.0122)	-0.0586*** (0.0119)	-0.0613*** (0.0121)
School Treatment	-0.000164** (0.000)	- (0.000)	- (0.000)	-0.000170*** (0.000)	- (0.000)	- (0.000)
School FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Student Controls	N	Y	Y	N	Y	Y
School Controls	N	N	Y	N	N	Y

Notes: Table 7 reports similar estimation of tables 3 and 4, however with alternative measures vocational school exposure. Binary treatment represents an indicator variable with one to municipalities m that receive a vocational school in time t , and School Treatment represent the share of vocational school in relation to regular ones. All specification contains school and time fixed effect. The standard errors are estimated clustering by school. Significance levels: 1% ***, 5% **, 10% *.

7. Mechanisms

Few papers address empirically the channels that explain the peer effects on education, exceptions are Duflo, Dupas, and Kremer (2011), and Lavy, Paserman, and Schlosser (2012). From a theoretical point of view, two concurrent theories attempt to predict the mechanisms of peer effect. First, reducing peer quality can affect classroom behavior, impacting student's effort to learning as in Lazear (2001). In turn, Duflo, Dupas, and Kremer (2011) argue that student composition can change teacher pedagogical practices, and thus affecting the quality of learning. Interestingly, theories predict contradictory effects for the importance of peers in secondary education.

This paper contributes to this discussion by assessing the channels of peer effects in an exploratory way. Using the contextual survey of SPAECE, we select 20 items referring to the following categories: student engagement, social skills, interest in learning, school well-being, noise at school, expectations, and time allocation. We test the effect of the proportion of students attending vocational school on these indicators¹⁹.

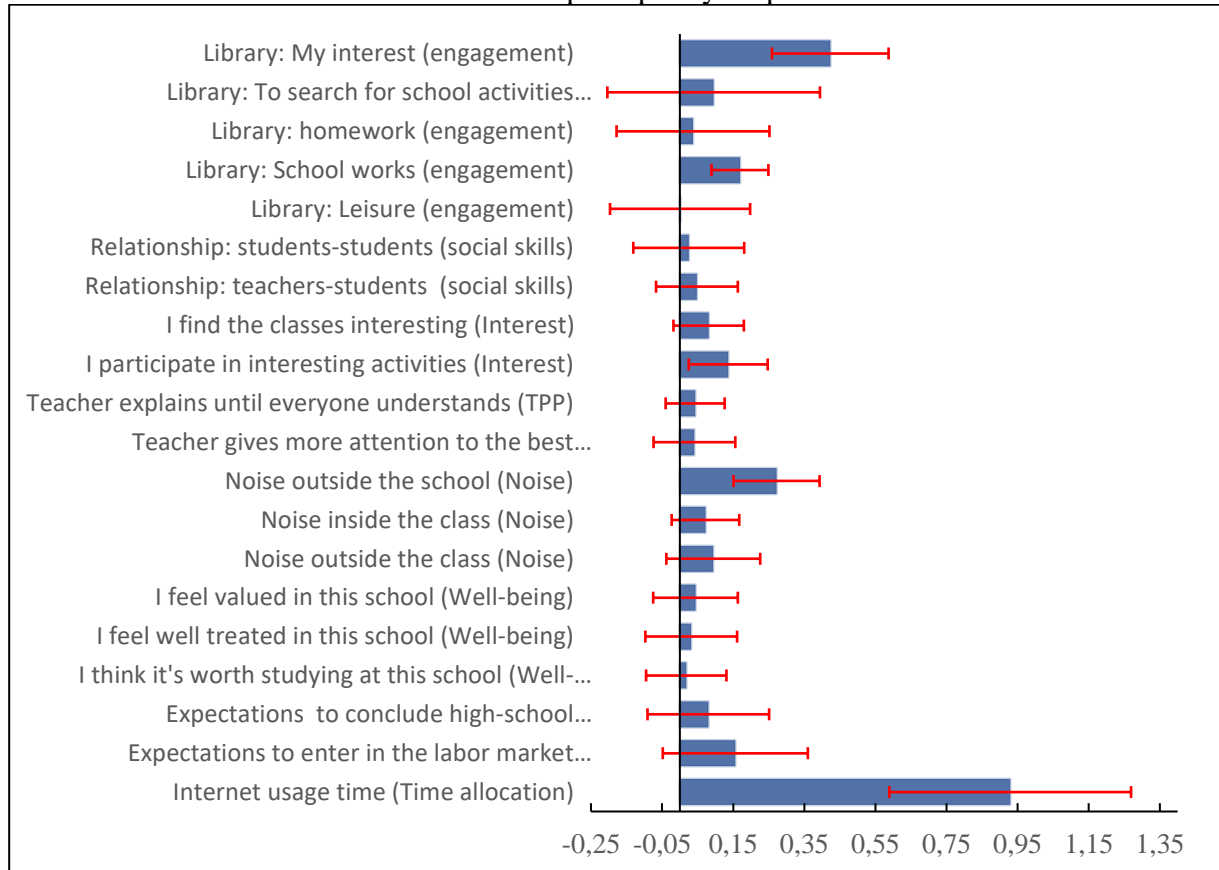
We consider a model that includes covariates as the age, gender, race, an indicator if the student's family attends a conditional cash-transfer program, and mother's education. Also, we include time and school fixed effects to control for time-varying and time-invariant unobservable factors. Figure 4 displays the estimates for each item. In parenthesis is the categories' names. For example, "*Library: My interest (engagement)*" refers to the question "*I go to the library to my own interest*" and is related to school engagement. All items are standardized, then the estimates represent standard deviations.

Only four variables present significant effects: going to the library by my interest, going to the library because of school works, the noise outside the school, and internet usage. The last two variables suggesting that peer quality affects student attention. The share of students attending vocational school affects the student time spending on the internet, i.e., reducing peer quality changes the student's time allocation. Another relevant factor is the noise outside the school, which is negatively affected by the share of vocational students. Both results suggest that reducing peer quality diverting the student focus on learning.

¹⁹ Although the contextual survey of the SPAECE to be applied in all year of the sample, there are substantial difference among the items across the years. Therefore, we consider only the cohorts 2010-2013 and 2011-2014. To detailed information about this sample, see the Online Appendix available by email to the authors.

Interestingly, our results contrast with Duflo, Dupas, and Kremer (2011), and Lavy, Paserman, and Schlosser (2012) because we do not support the notion that a high proportion of low-achieving students induce teachers to modify their pedagogy and their personalized attention to better students. Students do not perceive that teacher diverting her attention to specific students.

FIGURE 4: Effect of peer quality on potential mechanisms.



Notes: Figure 3 presents the impact of the share of vocational students on several variables that explains the mechanisms of why the reduction of peer quality affect students' outcomes. We regress each variable against additive school and fixed effects and the share of vocational students. The standard errors are estimated clustering by school level.

8. Conclusion

This paper investigates if peer composition affects secondary students' outcomes. To identify the causal effects, we exploit the variation of vocational schools in time and location across municipals of Ceará from 2008 to 2017. Vocational schools in Ceará realize an admission process to cherry-pick the high-achievers from primary education. This selective process reduces the share of high-achievers in regular secondary schools.

We find that regular students exposed to vocational school construction decrease their performance in math and language test scores. The effect is large, significant and presents small variations in different specification, suggesting that the estimates are reliable.

In addition, we also test the effect of peer composition on the rate of school drop-out and the rate of retention during secondary education using a school-level version of the empirical strategy. Our results indicate that reducing the peer quality increase the rate of students that drop-out the school (this result is restricted to large municipalities) and the rate of repeaters. We realize a battery of robustness' checks and we conclude that our estimates are unlikely to be biased.

Moreover, we are interested to understand the underlying mechanisms that explain the results. Using a contextual survey applied to students, we test different potential channels: student

engagement, social skills, teachers' pedagogical practices, student's well-being, school and classroom noise, and student's time allocation.

The vocational school exposure affects student time allocation and outside school noise. Specifically, the share of students attending vocational schools increases the time spending by regular students on the internet. Therefore, we conclude that the mechanism that explains our results is the students' diversion, following the "bad apple" theory of Lazear (2001).

Finally, this paper contributes to literature investigating the unintended effect of some policies. In our case, the construction of selective public schools changes the composition of regular schools, reducing the quality of their peers. The effect harms the regular students affecting short and long-run outcomes. Policy-makers should account for this spillover effect when evaluating the overall impact of the vocational school expansion in Ceará.

References

- ANELLI, Massimo, and PERI, G. The Effects of High School Peers' Gender on College Major, College Performance and Income. *Economic Journal*, 2017.
- ANGRIST, J. D. (2014). The Perils of Peer Effects. *Labour Economics*, 30, 98-108.
- ATKIN, David, "Endogenous Skill Acquisition and Export Manufacturing in Mexico," *American Economic Review*, August 2016, 106 (8), 2046–85.
- BIANCHI, N. The Indirect Effects of Educational Expansions: Evidence from a Large Enrollment Increase in University Majors, *Journal of Labor Economics*, v. 38 (3), 2020.
- BIFULCO, Robert, FLETCHER, J.; OH, S. S. and ROSS, S. Do High School Peers Have Persistent Effects on College Attainment and Other Life Outcomes?" *Labour Economics* 29: p. 83–90, 2014.
- BILLINGS S, DEMING DJ, ROSS S. Partners in Crime: Schools, Neighborhoods and the Formation of Criminal Networks, NBER Working Paper n° 21962, 2016.
- BOSSAVIE, L. The Effect of Immigration on Natives' School Achievement: Does Length of Stay in the Host Country Matter? *Journal of Human Recourses*, 55(2), p. 733-766, 2020.
- CARRELL, S.; HOEKSTRA, M. and KUKA, E. The long-run effects of disruptive peers. *The American Economic Review*, 108 (11), p. 3377-3415, 2018.
- CARRELL, Scott E., and James E. WEST. Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors. *Journal of Political Economy* 118 (3): p. 409–32, 2010.
- CARRILLO, B. Present Bias and Underinvestment in Education? Long-run effects of childhood exposure to booms in Colombia, *Journal of Labor Economics*, (forthcoming), 2020.
- CHARLES, Kerwin Kofi, HURST, E., and NOTOWIDIGDO, M. Housing Booms and Busts, Labor Market Opportunities, and College Attendance, *American Economic Review*, October 108 (10), 2947–94, 2018.
- DUCKWORTH, A. L., PETERSON, C., MATTHEWS, M. D., & KELLY, D. R. Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), p. 1087-1101, 2007.
- DUFLO, Esther. The medium run effects of educational expansion: Evidence from a large school construction program in Indonesia. *Journal of Development Economics*, 74(1), p. 163–197, 2004.
- DUFLO, E., DUPAS, P., and KREMER, M. Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *The American Economic Review*, 101(5), 1739-1774, 2011.
- FALK, A. and ICHINO, A. Clean Evidence on Peer Effects. *Journal of Labor Economics*, 24 (1), 39-57, 2006.
- FORTIN, Bernard and YAZBECK, Myra, Peer effects, fast food consumption and adolescent weight gains. *Journal of Health Economics*, vol. 42(C), p. 125-138, 2015

GEORGANAS, S.; TONIN, M and VLASSOPOULOS, M. Peer pressure and productivity: the role of observing and being observed. *Journal of Economic Behavior and Organization*. Vol 11(c), p. 223-232, 2015.

GILRAINE, M; MACRTENEY, H. and McMILLAN, R. Education Reform in General Equilibrium: Evidence from California's Class Size Reduction, NBER Working Paper No. 24191, 2018.

GLEWWE, Paul and KASSOUF, Ana Lucia, The impact of the Bolsa Escola/Familia conditional cash transfer program on enrollment, dropout rates and grade promotion in Brazil, *Journal of Development Economics*, vol. 97(2), p. 505-517, 2012.

GOULD, Eric D., LAVY, V. and PASERMAN, M. Daniele. Does Immigration Affect the Long-Term Educational Outcomes of Natives? Quasi-Experimental Evidence. *Economic Journal* 119 (540): p.1243–69, 2009.

HOXBY, C. M., and WEINGARTH, G. Taking Race out of the Equation: School Reassignment and the Structure of Peer Effects: Working paper, 2005.

IMBERMAN, Scott A., KUGLER, A. and SACERDOTE, B. Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees. *American Economic Review* 102 (5): p. 2048–82, 2012.

JACKSON, K. Teacher Quality at the High-School Level: The Importance of Accounting for Tracks, *Journal of Labor Economics*, V. 32(4), 2014.

LAVY, V., PASERMAN, M. D., and SCHLOSSER, A. Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom. *The Economic Journal*, 122(559), p. 208-237, 2012.

LAZEAR, E. P. Educational Production. *The Quarterly Journal of Economics*, 116(3), p. 777-803, 2001.

MANSKI, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 60 (3), 531.

SACERDOTE, B. Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward? *Annual Review of Economics*, 6 (1), 253-272, 2014.

SHILTZ, F.; MAZREKAJ, D.; HORN, D. and De WITTE, K. Does it matter when your smartest peers leave your class? Evidence from Hungary, *Labour Economics*, Volume 59, p. 79-91, 2019.

Appendix

Table A1. Descriptive statistics

Variables	Treated		Untreated		Difference
	Average	SD	Average	SD	
<i>Test Scores</i>					
Test Scores Math 9°	236.80	44.50	236.70	44.83	0,10
Test Scores Portuguese 9°	229.00	45.75	227.72	45.50	1,28
Test Scores Math 3°	259.24	49.08	258.72	49.82	0,52
Test Scores Portuguese 3°	258.07	47.08	251.48	46.79	6,59*
<i>Race</i>					
Black	0.223	0.416	0.130	0.336	0,09*
Brown	0.125	0.331	0.081	0.274	0,04*
White	0.206	0.404	0.134	0.341	0,07*
Other	0.443	0.496	0.653	0.475	0,21*
<i>Mother education</i>					
Incomplete Primary	0.023	0.150	0.012	0.109	0,01*
Primary	0.169	0.289	0.093	0.216	0,07*
Secondary	0.055	0.228	0.031	0.174	0,11*
Higher	0.019	0.139	0.011	0.106	0,008
<i>Others</i>					
Conditional Cash-Transfer	0.754	0.430	0.742	0.437	0,01
Public Transport	0.290	0.453	0.510	0.499	-0,22
Girl	0.551	0.497	0.566	0.495	0,01

