Quality of education and human capital decisions: experimental evidence from Brazil*

Lucas Finamor† Bruno Ferman‡ André Portela Souza§

Sao Paulo School of Economics - FGV

Working Paper — April 19th, 2017

Abstract

This study makes use of the "*Jovem de Futuro*" program, an experimental intervention in Brazilian public high schools, to estimate the impact of quality of education on college and labor market entry. Our results indicate that high school students respond with more and better college enrollment after an increase in the quality of high school education. A better high school environment boosts the probability of students (i) to attend college, (ii) to be accepted into public colleges and into high quality and selective majors, (iii) and to be studying full time. These effects seem to be intermediated with admission through affirmative action policies (quotas). A reduction in working and only working (not studying) status is also detected. These results can be interpreted as suggestive evidence of the existence of credit constraints or heterogeneity in the returns to schooling function.

I. INTRODUCTION

The relationship between returns to education and schooling decisions is a cornerstone for the field of labor economics. The theoretical models looking for a rationalization of human capital investments were first developed in the sixties and seventies with Becker (1962), Ben-Porath (1967) and Mincer (1974), whose contributions are highly influential until today. In these approaches, schooling decisions were typically considered as the years of formal education acquired by an individual and, in general, no measure of education quality was not taken into account. Behrman & Birdsall (1983) made one of the first theoretical attempts to include quality of education in the mincerian equation. Based on the heterogeneity of quality across regions, their model predicts that the rate of returns to education is conditional on school quality, both being positively correlated. Moreover, in equilibrium, individuals who perceive higher returns to education also choose more years of schooling.

Several studies have tried to empirically estimate the causal impacts of a better schooling

*We would like to thank Ricardo Paes de Barros and the staff from the Instituto Unibanco for sharing with us the data from the experiment and INEP's staff for cooperation with their data. We are also in debt with Lycia Lima, Dalila Figueiredo, Marcel Fernandes, Amanda Arabage, Eduardo Ferraz, Sammara Cavalcanti, Vinicius Lima and seminar participants at EESP-FGV for their comments and suggestions.

†lucasfinamor@gmail.com

‡bruno.ferman@fgv.br

§andre.portela.souza@fgv.br

environment. The most common challenges these studies have to overcome are, in general, (i) to obtain an objective measure of education quality, (ii) to find a clear identification with a credible exogenous variation of schooling quality and (iii) to link adult outcomes with data on education. In this research, we try to overcome these challenges by using a randomized control trial (RCT) in Brazil called "*Jovem de Futuro*"(*Youth of the Future*). The program has been running since 2008, aiming at improving the quality of education in public high schools through improving management practices and allocating grants to the treated schools. Given the significant impacts on students' proficiency uncovered by several previous evaluations (Silva 2010, Barros et al. 2012, Oliva 2014) and the RCT structure of the program, we consider it as an exogenous shock of educational quality on those beneficiary schools and investigate the effects on students' labor market outcomes and higher education decisions.

The existent literature faced these challenges with different strategies. Some works used the experimental evidence from the STAR experiment[1], and positive effects of reduction in class size were found in the probability of students (i) to take a college-entrance exam (Krueger & Whitmore 2001), (ii) to attend college and high-quality colleges (Chetty et al. 2011), and (iii) to attend college, entering college on time and to earn a college degree (Dynarski et al. 2013). Other authors also made use of class size as a measure of educational quality and encountered positive effects on the decision to remain in school beyond the age of 16 (Dustmann et al. 2003) and on increasing the years of completed education (Fredriksson et al. 2013). Moving the focus to teachers rather than schools' inputs, Dearden et al. (2002) use the pupil-teacher ratio in England to assess its impact on adult outcomes and only find effect on wages at age 33, whereas no effect on human capital decisions is detected. Chetty et al. (2014), using a measure of value-added by teachers, reveal impacts of quality of education in college attendance and wages. Lavy (2016) explores experimental data on teachers' incentives and finds positive impacts of a teachers' incentive mechanism on years of schooling, enrollment in university, employment and earnings.

Other researchers took advantage of the high-quality American charter schools that used student lotteries to determine which students were admitted in oversubscription situations. The estimated effects are increasing the quality, but not necessarily the quantity of college attendance (Dobbie & Fryer Jr 2015, Angrist et al. 2016), increasing student proficiency, the probability of taking SAT and students shifting from 2-years to 4-years college (Angrist et al. 2016). Deming et al. (2014) also use high-quality schools with oversubscription and lotteries to reveal impacts of higher quality schools on the probability of high school graduation, post secondary attendance and degree completion.[2] With our strategy, we are closer to these last works, however instead of assigning students to better schools, we exploit an exogenous shock at school level.This way, we contribute to the literature estimating the effects of increasing quality in a given set of schools. While confound factors such as peer composition and geographical changes could be associated with the literature's previous results, our findings are closer to an exogenous quality shock for the students attending the same schools with the same peers. The evidence on improving the quality of schools is clearly distinct from the impact of assigning students to high-quality schools. The differences are relevant from both

[1]Details can be found in Krueger & Whitmore (2001), Chetty et al. (2011) and Dynarski et al. (2013).

[2]There is also experimental evidence from programs that combine education, health and other interventions such as the Head Start and Perry school programs. Since we are exclusively trying to estimate the causal impacts of quality of education we are not relating to this literature.

the Economics and policy perspectives.

We are also adding to the literature evidence on a developing country, where the labor-college decision has a much more importance than in developed countries. Barro & Lee (2013), using educational data from 146 countries, point out how higher education enrollment is much lower in the developed countries: in average 10.5% of the population over 15 years old, while in the developed countries this average is more than three times higher: 32.2%. Using the Brazilian household survey (PNAD) in 2014 we can also picture this argument in the country: among the adults aged 30 to 35 that completed high school, who are likely to have already finished the schooling investments, almost 59% did not progress to college. No more than 28% earned a college degree. The decision to simultaneously work and study is also very common. Among the individuals studying, more than 27% aged 15-18 (high school) study and also work. Considering the students expected to be in college, aged 19-22, this proportion reaches more than 40% (from PNAD 2014).

Lastly we have a highly diverse sample. The schools within the experiment are located in several different Brazilian municipalities and states. The initial quality-level is also very diverse, the sample comprises schools ranked along the entire distribution of proficiency in the National Exam of Upper Education (ENEM), which reduces the concerns about the external validity of our results.

In this paper we can assess how high school students in developing countries respond to the decision of continuing to invest in formal human capital (attending college), going directly to the labor market or acquiring both education and market experience at the same time. We find that freshman high school students in the initial year of the program, that are supposed to receive three years of treatment, are more likely to attend college. The effect is more than 3 percentage points (hereinafter pp), and significant for at least the first three years after high school completion. This corresponds to 7.5% increase in college attendance. The probability of being admitted to public colleges[3] (1.8 pp), to only study (1.7 pp) and to attend high-quality colleges (3.6 pp) are also increased. This effect of access on higher education seems to be enabled through the use of affirmative actions such as quotas to public school students (1.4 pp). A reduction in the probability of only working individuals is also detected, specially for men (-4.5 pp).

Therefore, students in Brazilian public schools are responding with more and better college enrollment after an increase in the quality of high school education. This is in line with the previous findings and the theoretical predictions. With this results we also commune with Angrist et al. (2016) and Lavy (2016) regarding how interventions in high school may produce positive effects and be cost effective. Reasonably, high schools interventions may not be too late. Additionally, these results could be driven by the expansion of the returns to schooling function or by the existence of credit constraints.

The next sections explore the details of the Program Jovem de Futuro (section II) and present the data and econometric strategy (section III). Section IV presents the results on college and labor decisions, college admission and selectiveness while section V explores het-

[3]As will be detailed later, in Brazil, public colleges are associated with an overall better reputation than the private colleges.

erogeneity in the results by baseline proficiency and gender. Section VI assess the sensitiveness of the results with alternative datasets and brief comments on the main findings conclude the article (section VII).

II. "JOVEM DE FUTURO" PROGRAM

Jovem de Futuro is an initiative of the *Instituto Unibanco*[4] (Unibanco Institute) that has been implemented in Brazilian public schools since 2008 in partnership with the states' secretariat of education. The aim is to improve the quality of education and to reduce the school dropout rates in high school. The intervention is three years-long, which is exactly the expected time for a student to complete high school. So, in each school there is a cohort that receives the full treatment for three years.

The program can be characterized by two pillars: (i) conditional cash transfers and (ii) technical support for school management. Each treated school receives an amount of R$ 100.00 (US$ 48.00) per student-year[5] as a conditional cash transfer. The conditions are to improve the students' proficiency in a standard-based exam applied by the Institute at the end of each school year and to implement a participatory budget process in the school (Barros et al. 2012). The expenditures are a discretionary choice of the school's principal, but the resources allocation has to follow some rules: no more than 40% can be used in infrastructure and at least 20% should be used in incentive mechanisms to teachers and students (20% for each group) (Rosa 2015). The second element of the program is the focus on school management: there is a transfer of management technologies[6] and the participation of students, students' parents and all school staff is stimulated. Agents from the Instituto Unibanco supervise the implementation of these new technologies and keep track of schools once a week.

The first phase of the program took place in the states of Minas Gerais (MG) and Rio Grande do Sul (RS). The treated schools received the program in 2008-2010 while the control group received the treatment immediately after the treatment group: 2011-2013. The second phase selected schools from two different states: Rio de Janeiro (RJ) and São Paulo. In the latter the program was implemented in two different areas: the metropolitan region of São Paulo (SP) and Vale do Paraíba (VP). The second phase intervention started in 2010 for the treated schools and in 2013 for the control group. Hereinafter we will refer to these localities (MG, RS, RJ, SP and VP) as the five areas where the experiment was implemented.

At each locality, state public schools could freely subscribe to the program, knowing in advance the programs' structure and also its experimental design[7]. The assignment to treatment and control was completely at random, using pairwise randomization, with stratification at school localization, number of students and proportion of students in high school. In the state of Minas Gerais the randomization procedure had minor differences: in this state the

[4]The Instituto Unibanco is a private organization founded in 1982 whose goal is to improve the quality of public education in Brazil.

[5]The rule that determined the amount of resources had a slight change from phase 1 to phase 2, in the first the number of students considered was the total number of students in school, meanwhile in the second phase only the number of students at high-school was considered but a minimum transfer of R$ 100,000 yearly was guaranteed for all schools. For details please check Rosa (2015).

[6]Such as the logic model and guidance to the adoption of an output oriented management.

[7]Rosa (2015) states how in São Paulo the secretariat of education established that only the schools with the lowest performance in the state educational evaluation would be eligible for the program.

number of control schools was larger than that of treatment schools, henceforth, strata with few schools were created instead of pairs[8]. The number of schools in each group and area are presented in Table 1 below.

Table 1: Randomization and Non-compliance

	Assignment (A)		Non-Compliance (B)		(A)-(B)		
Area	Treatment	Control	Treatment	Control	Treatment	Control	Total
MG	20	28	0	0	20	28	48
RS	25	25	3	0	22	25	47
RJ	15	15	3	0	12	15	37
SP	21	20	1	3	20	17	37
VP	20	20	0	0	20	20	40
Total	101	108	7	3	94	105	199

Source: Administrative data — Instituto Unibanco.

As we can see in Table 1 there were problems with compliance on the treatment and control groups. There were two main problems in schools that did not comply with the assignment to treatment: (i) insufficient documentation or other bureaucratic problems, or (ii) insufficient performance of students in the program's intermediary evaluations. On the control group the main problem was that some schools ended up receiving the treatment by the Federal Ministry of Education. To avoid problems with selection regarding the schools' decisions to comply, we will focus only on the treatment assignment, limiting us to only identifying the intention to treatment (ITT) effect.

In the state of São Paulo (SP and VP areas) the program was intended to be implemented in 2009 but it only happened late in September and for more than 40% of the schools in November. Therefore, the Instituto Unibanco decided to consider 2010 as the initial year for the program in these areas. When assessing the baseline balance in 2009, the treatment and control groups were not balanced in terms of the *Prova Brasil*[9] student proficiency. The details will be explored in the section VI, however, in order to preserve our estimations, SP and VP were dropped from our analysis[10].

Table 2 presents the balance between treatment and control groups for several infrastructure and school features. Information comes from the School Census and ENEM datasets, which will be presented in the data section. As we can see, we cannot reject the equality of means for both groups for all variables except for the presence of Internet at schools. This does not raise our concern since it seems to be totally at random and it goes in the opposite direction of our findings (control group displays a higher proportion of schools with Internet connection). Important characteristics such as the number of students enrolled in school, in high school, in the senior year of high school and enrollment in the national exam (ENEM) are balanced among the schools in the experiment.

[8]The variables used in the stratification procedure were: school's location, number of enrolled students, proportion of high school students in the total number of students and proficiency of students in the senior year of high school.

[9]National exam for the 5th and 9th grades applied in all Brazilian public schools, every odd years since 2005— details are presented in section III.

[10]As a robustness check, we reestimate our main results including these areas and the results are qualitatively the same. Details are displayed in section VI.

Table 2: Baseline balance

	Treated schools		Control schools		Difference
	N	mean (sd)	N	mean (sd)	diff [se]
Computer lab	60	0.750	68	0.765	-0.026
		(0.437)		(0.427)	[0.080]
Science lab	60	0.750	68	0.779	-0.045
		(0.437)		(0.418)	[0.077]
Library	60	0.950	68	0.971	-0.018
		(0.220)		(0.170)	[0.027]
N existence classrooms	60	19.62	68	18.44	1.02
		(10.88)		(6.07)	[1.38]
N classrooms in use	60	19.23	68	18.46	0.703
		(9.14)		(6.00)	[1.23]
N of computers	60	20.88	68	22.66	-1.93
		(18.33)		(20.73)	[3.39]
Internet	60	0.833	68	0.985	-0.161***
		(0.376)		(0.121)	[0.051]
N of employees	60	118.75	68	100.69	20.61
		(132.15)		(34.29)	[18.64]
N of students	60	1619.15	68	1635.43	-7.07
		(578.69)		(543.54)	[84.03]
N of students in HS	60	990.37	68	1043.69	-40.42
		(471.15)		(503.93)	[55.68]
N of senior students in HS	60	230.52	68	247.81	-9.96
		(119.11)		(139.10)	[15.71]
N of students enrolled in ENEM	60	134.30	67	138.74	2.35
		(82.23)		(79.90)	[10.84]
% of students enrolled in ENEM	60	0.579	66	0.541	0.049
		(0.196)		(0.164)	[0.033]

Source: ENEM and School Census (2007 and 2009) — INEP. NOTE — The first two columns display respectively, the number of treated schools and the mean and standard deviation (in parenthesis) for each variable. Columns 3 and 4 exhibit the same for the control group. The fifth column displays the coefficient and standard error for the treatment variable in the following regression: $Y_{js} = \alpha + \beta T_{js} + \eta_s + \varepsilon_{js}$, where Y_{js} is the outcome for school j in strata s, T is the treatment indicator and η_s strata fixed effect.

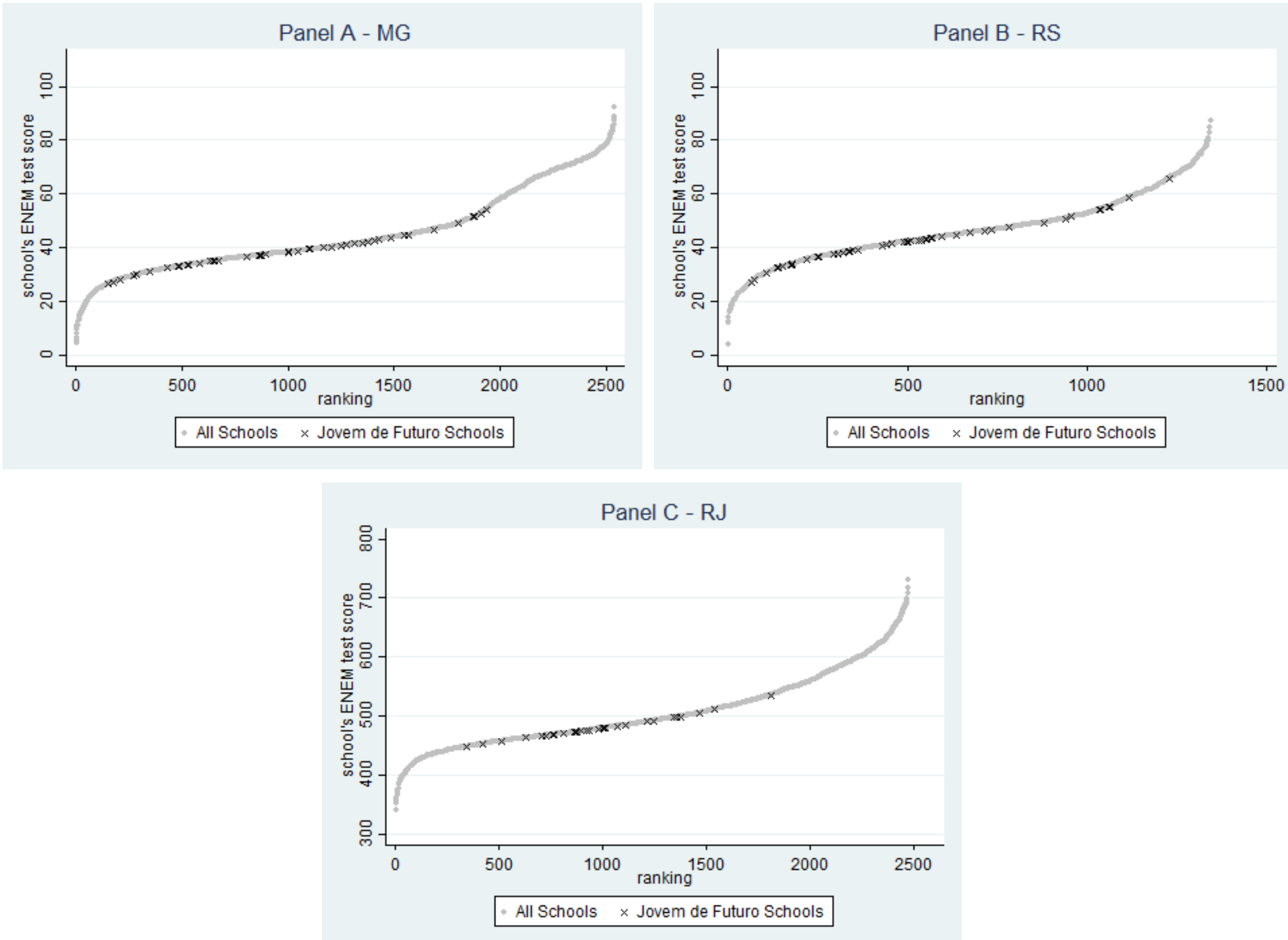
Using the average test score in ENEM as a measure of school quality we can assess whether the schools selected in the experiment are a representative subset of the non-selected schools in the state. Figure 1 displays ENEM's average at school level in the baseline year (2007 and 2009) and its ranking in the grade's distribution within each state, for all schools in the states of MG, RS and RJ. Schools participating in the experiment are evidenced with the X marker. It is possible to see that the schools are well dispersed along the grades distribution. This, added to the fact that the experiment takes place in 3 different states, several cities and covering almost 130 schools, gives substantial external validity to our results.

III. DATA AND ECONOMETRIC STRATEGY

A. DATA

The School Census is an annual census survey of students enrolled in Brazilian schools from kindergarten until high school produced by INEP (*Instituto Nacional de Estudos e Pesquisas*

Figure 1: ENEM's test scores at school level (baseline year)



Source: ENEM 2007 and 2009 — INEP. NOTE — Each point displays the school average (vertical axis) and its position in the ENEM's test score state ranking (horizontal axis). Panel A displays all schools from MG state and the schools participating in the PJF experiments are evidenced with the X marker. The same applies for RS state (panel B) and RJ (panel C). Since the exam changed its scale in 2009, the vertical axis are not the same. The writing exam is not considered.

Educacionais Anísio Teixeira), a research agency linked to the Ministry of Education. For every student-year observation it is possible to establish its enrollment status, grade and school. Using this dataset, we can generate the list of high school freshman students in the initial year of the intervention at each participant school. All students in this list are considered as the experiment participants with treatment status being defined by the school status in the randomization. This enables us to identify all individuals belonging to the three-years experiment cohort.

There is also a similar version of the School Census for higher education, the Higher Education Census, which is available since 2007 and comprises all students enrolled at any higher education college. The dataset allows us to observe the students' enrollment status, type of selection exam, funding, major and college information.

The Brazilian Ministry of Labor produces the Labor's Annual Social Information Report, the RAIS dataset, which contains administrative records of all formal labor contracts in the country. Every observation is a formal labor contract between a worker and a firm in a given year, which specifies the contract status at the end of the year (active or inactive), the duration of the work agreement, wage and the amount of contracted hours. Other information like workers' gender, age, educational attainment and other job-related characteristics are also available. The confidential identified dataset of RAIS spans 2002-2013. It is clear that when using RAIS, we are losing an important side of the labor market: the informal workers, in other words, those that do not have a legal working contract. However, due to their intrinsic nature, it is impossible to have administrative data on the informal contracts[11]. Using only secondary data, it is unconceivable to incorporate informal jobs into the analysis.

Unfortunately, there is not a direct link between the School Census and the datasets after compulsory schooling: Higher Educational Census and RAIS. The existing identification number (student ID) in the School Census is limited to the mandatory education informations such as the School Census and some standardized national exams. Higher Educational Census and RAIS use the individual social security number[12] as the primary identification key. However, in the end of high school students can take a National Exam called " National Exam of Upper Secondary Education" (ENEM), in which students have to fill their social security number to complete the enrollment. For almost all students enrolled in the exam, INEP is able to assemble the student unique student ID[13]. This enable us to link the formal education information with higher education and labor market datasets.

ENEM has four objective exams based on the Item Response Theory (IRT): natural sciences, human sciences, language and mathematics as well as a writing exam. The exam was created in 1998 in order to evaluate the high school quality in Brazil, but had a completely different structure. In the course of time the exam started to be used as a part of the college admission process and culminated as being nowadays the main process to select students to public universities, to select beneficiaries for public scholarships in private colleges (Prouni) and also to choose students that can access the public funding systems that distributed loans to students in private colleges (FIES). Other public programs such as scholarships for academic exchanges also rely on the student score in ENEM. Students older than the ideal age to finish high school, can also obtain a high school degree if they achieve a minimum score in the exam.

Therefore, students have several incentives to participate in the national exam. As we can see on table 3, more than 60% of the students in the senior year of high school are enrolled in ENEM. The proportion is continuously increasing over time, reaching almost 80% of participation in 2014. We will use data from the exams of the years 2009-2014 and scores are available for the 5 components: natural sciences, human sciences, language, mathematics and writing.

From the list of freshman students of high school we retrieve 56,166 students. It is important to state that this approach reinforces our "intention to treat" strategy since our sample

[11] There is data such as the Brazilian demographic census or surveys such as PNAD that allow the estimation of the relative size of the informal market, the wage distribution, duration of work and other features of this important side of the labor market. Some information can be found in Dix-Carneiro & Kovak (2016).

[12] *Cadastro de Pessoa Física* - CPF

[13] For approximately 70% of students in ENEM's dataset the student ID is available.

Table 3: ENEM's enrollment

	2009	2010	2011	2012	2013	2014
Senior high school students (million)	2.147	2.198	2.214	2.226	2.213	2.225
ENEM's enrollment (million)	1.339	1.375	1.500	1.519	1.627	1.749
Proportion	0.624	0.625	0.678	0.683	0.735	0.786

Source: ENEM and School Census 2009-2014 — INEP. NOTE — Number of high school students is given by the School Census. ENEM's enrollment is restrict to senior high school students. Proportion of students enrolled in the senior year of high school that takes ENEM.

relies on the students' initial treatment status despite their schooling future decisions such as to complete high school, to drop out or to change school. Using their student ID number we look for them in the ENEM's dataset from 2009-2014 in order to obtain their social security number. The success in this procedure defines whether each student is present in our final sample (selected) or is considered as attrition (not selected). We are able to identify 14,216 students or 25.3% of our sample[14] . We have three sources of attrition: (a) students that dropout after the first year of high school, (b) students that do not take ENEM, and (c) students that do not have INEP's code in the ENEM dataset. While (c) is arguably random, (a) and (b) are likely to be affected by treatment and this will be considered in our empirical strategy.

We have information on access and characteristics of higher education for all selected students from 2011 to 2015. For labor market information we have data from 2011 through 2013. For phase 1 students (MG and RS areas) we will have 5 years of higher education information and 3 years of formal labor market information after expected high school completion (2010). Analogously, for phase 2 students (RJ area), higher education information will be available for 3 years after high school graduation (2013-2015) and 1 year (2013) for formal labor market.

Since 2005, in every odd year in Brazil there is a large national exam applied in every primary public school in the 5th and 9th grades[15] called *Prova Brasil* (Brazil's exam). Starting in 2009, identified individual data is available. Therefore, we can assess phase 2 students' proficiency with the aid of this exam. This information enables us to carry out some heterogeneity exercises on the students' proficiency prior to the program. Figure 3 in the appendix A details all the datasets, links and sources of attrition.

Table 4 below presents some descriptive statistics on the main variables. Our initial sample, that is, the list of freshman students in the first year of the program, is composed by 56,166 individuals. Selection indicator equals one if we can identify a student's social security number, we identify 25.3% of our sample. Approximately 46.9% of the students are in treated schools and 52.1% are women. For 17.2% (9,660) we have data on the ENEM's tests scores in the expected senior high school year[16]. For the RJ sample we have data on the *Prova Brasil* 2009

[14]It is important to state that we execute the analysis with an alternative data (with only senior students) where attrition is significantly lower (selection is above 75%) and the results are qualitatively the same. We chose as benchmark the data with freshman students since it exhibits less layers of attrition that could be affected by the treatment. Details will be given in section VI.

[15]Brazilian educational system is organized in 9 years of elementary school, through grades 1st to 9th, and in 3 years of high school (10th to 12th grades).

[16]For this assessment, we lose some students that were enrolled but did not take the exam and those students

proficiency (language and math) for more than 7,500 students.

Table 4: Descriptive Statistics

Variable	Nobs	Mean	sd
Sample			
Treatment	56166	0.469	0.499
Women	56166	0.521	0.500
Sample Selection	56166	0.253	0.435
ENEM's test scores			
ENEM's test score selection	56166	0.172	0.377
Nature Sciences	9585	472.028	68.304
Human Sciences	9585	524.726	78.200
Language	9364	506.960	66.017
Math	9364	506.475	97.290
Prova Brasil **baseline test scores [RJ]**			
Language	7594	248.201	43.106
Math	7586	250.734	44.092
Human capital decisions			
College	14216	0.475	0.499
Labor	14216	0.717	0.450
Only college	14216	0.144	0.352
Only labor	14216	0.386	0.487
College and labor	14216	0.331	0.471
College administration			
Public	14216	0.070	0.256
Private	14216	0.417	0.493
College admission			
Vestibular	14216	0.353	0.478
ENEM	14216	0.166	0.372
Quotas	14216	0.041	0.198

NOTE — Descriptive statistics. ENEM's test scores are from the expected senior year of high school (2010 in MG and RS and 2012 in RJ). Prova Brasil test scores are from 2009, only for RJ area. College indicator equals 1 if an individual is enrolled in any college-major in the analyzed period, Labor equals one if an individual has a formal work agreement during the analyzed period, Only college equals 1 if an individual is enrolled in a college-major and do not have an work agreement, the same holds for Only labor. College and Labor equals 1 if a individual is enrolled and have a formal work agreement. College administration and college admission can sum proportions higher than the proportion of students in college, because a student can be enrolled in more of one college-major during the analyzed period.

Considering the human capital decisions, 47.5% of the individuals enroll in any college-major in the analyzed period, among which 14.4% only study, i.e., do not have a formal work agreement and 33.1% study and work in the period. 71.7% have a formal work agreement in the period, among which 38.6% only work (do not enroll in any higher education major). It is worthy remembering that we have two more years of information on higher education than on

that only enrolled in the exam in later years.

the work agreements (RAIS). 7% of the students are enrolled in a public college while 41.7% are enrolled in private ones. Assessing the admission process, 35.3% use regular university entry exams (*vestibular*) and 16.6% use ENEM. 4.1% of students use quotas to dispute for a place (this does not exclude the ENEM and *vestibular* options)[17].

B. EMPIRICAL STRATEGY

Given the random assignment of treatment and control, our preliminary results rely on a simple regression controlling for strata:

$$Y_{ijs} = \alpha + \beta treat_{js} + \eta_s + \varepsilon_{ijs} \tag{1}$$

Where ijs indexes an individual i, in school j and strata s. Y_{ijs} is the outcome of interest and $treat_{js}$ is a dummy variable for treatment status. The fixed effect of strata, η_s is considered and all estimations are clustered at school level.

As addressed in the previous paragraphs, we only have data on the outcomes of interest for the students that have taken ENEM's exam and the probability of a student to enroll in the exam is likely to be affected by treatment. Table 5 below displays the selection average and also the effect of treatment on selection. The estimated treatment effect on selection is positive (1.7 pp) but we cannot reject the null hypothesis of a zero effect. Splitting the sample in the three areas of the experiment, we can reject the null-effect hypothesis only for RS, while the RJ exhibits a zero effect on average. Considering the gender, women seem to have a positive effect while men do not.

Despite the small and not significant treatment effect on selection, we also call upon a partial identification solution using Lee (2009), whose bounds estimators have desirable properties to be applied in this scenario. The procedure yields upper and lower bounds for the desired treatment effect, requiring mainly two assumptions: (i) independence of the potential outcomes from treatment, and (ii) monotonic effect of treatment on selection for all individuals. At each area (i) is directly verified given the random assignment of the treatment status. The second assumption cannot be tested, but it is strongly plausible — it is hard to think of a channel through which the treatment would discourage students to enroll in the national exam.

Since the randomization procedure in the first area (MG) was different from the remaining areas (not pairwised) and given the assumption (i) above, we apply the estimator separately in each area and then integrate the results using as weights the number of selected observations in the control group for each area. In order to use clusters at school level in the standard errors estimation bootstrap with 500 replications was used in the bounds procedure. We are not able to include strata's dummies in Lee's procedure, therefore in exceptional cases the bounds' estimation does not contain the point wise identification estimation[18].

[17] College administration and college admission can sum proportions higher than the proportion of students in college, because a student can be enrolled in more of one college-major during the analyzed period.

[18] The regression including strata fixed effect can differ from the regression excluding them if the treatment effect is heterogeneous and also the proportion of treated individuals is different among strata, what is likely to be true in our setup.

Table 5: Selection

Sample	Nobs	% Selected	Treatment effect	(SE)
All	56166	0.253	0.017	(0.014)
Areas				
MG	20744	0.283	0.022	(0.031)
RS	17195	0.252	0.033*	(0.018)
RJ	18227	0.220	-0.003	(0.022)
Gender				
Men	26927	0.206	0.005	(0.014)
Women	29239	0.297	0.027*	(0.015)
Student baseline proficiency [RJ]				
Below Median	3786	0.250	-0.024	(0.025)
3th quartile	1893	0.356	0.049	(0.034)
4th quartile	1893	0.491	0.007	(0.032)

NOTE — Proportion of selected observation and treatment effect on selection for different samples. The treatment effect comes from the following regression: $S_{ijs} = \alpha + \beta T_{js} + \eta_s + \varepsilon_{ijs}$, where S_{ijs} is an indicator variable whether student i in school j and strata s is selected. T is the treatment indicator and η_s strata fixed effects. Cluster at school level is considered. The student baseline proficiency is only available for RJ area. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Accounting for selection, the point estimated effect cannot be determined, but upper and lower bounds can be provided. The bounds can also help us to characterize our interpretation given some assumptions. For instance, assume that there is heterogeneity in the students' ability which is correlated with students' proficiency, probability of taking ENEM and attending college. The additional selected students that we can see among treated students are the compliers, that is, those that would not have taken ENEM without treatment but did it once having received treatment. If these compliers are among the high-ability individuals, then our point estimation is overestimated and the true effect is near the lower bound. Nevertheless, if the compliers are the low-ability students, than the point estimation is underestimated and the real effect comes closer the upper bound.

As stated before, for the RJ we have data on the students' proficiency in the Prova Brasil exam, one year before the treatment. On the one hand we can use this information to assess the plausibility of both scenarios relating to the compliers. On the other, as we can see in table 5, in RJ treatment seems to not affect selection. Even so, in the last part of table 5 we estimate the differential selection for students below median and in the 3rd and 4th quartiles[19]. For neither of them, treatment has a significant effect and the signal and magnitude do not exhibit a clear pattern. In addition, treatment effect on the upper quartile is essentially zero, which is interesting for two reasons: (i) it will be extremely useful for reaching tight bounds for

[19] As will be explained in section V, this division was chosen in order to increase the sample size in the lower tail distribution of baseline proficiency.

this subsample, and (ii) it reduces the concerns of overestimating the effects due to selection of high-ability individuals. This is also logically plausible since it is difficult to picture how the marginal students that are induced to enroll in ENEM would be in the upper tail of the baseline proficiency distribution.

IV. RESULTS

A. ENEM'S TEST SCORES

Previous researches have already established the positive and significant effects of the program on student's proficiency. Silva (2010), Barros et al. (2012) and Oliva (2014) present the effects using as a measure of students' proficiency an exam created and annually implemented by the Instituto Unibanco in the participant schools. The estimated effects are strikingly positive in all studies for all years and areas analyzed. But this measure of proficiency has two weaknesses. First, even though the exam was designed following the structure and scale of Brazilian standard exams, it is an exam created by the institute and therefore could be biased in any direction that could benefit the treated schools. The second vulnerability is due to the selection of a subsample of students to participate in the exam, that could generate some bias in the students' selection by the schools' principals. For example, one might imagine that stakes were higher for treated schools.

In order to avoid the two weaknesses mentioned above, Rosa (2015) uses ENEM to assess the program effects on students' proficiency. The results are very similar to the ones previously mentioned: the average effect on language and math scores are around 15% of standard error of the outcome variable. All the previous evaluations used only the pairs (or strata) where all schools complied with the assignment. Since we want to avoid any assumptions on the probability of schools not complying, and given the fact that we do have data for all the schools in ENEM's exam, we follow Rosa's strategy to assess the ENEM test scores but, in turn, we consider all schools in the experiment[20]. Our estimation is also different from Rosa (2015) since we consider all students from the first year of high school and not from the senior year[21]. Furthermore, we present bound estimations for the treatment effect. Table 6 below presents the results.

As we can see, the average intention to treat effect is positive and significant for all four objective exams. All the dependent variables are standardized by the grade distribution on the control group[22]. The effects for language and mathematics scores are approximately 12.5% standard deviation of the distribution of the outcome, around 3 and 4 percentage points less than the average treatment effects found in the literature. For natural and human sciences the magnitudes of the estimated treatment effect are smaller but still significant. Analyzing the estimated bounds, for language the estimated lower bound is positive. That is, even if all students selected only by virtue of the treatment were the students with the highest scores in the treatment group and are dropped from the sample, treatment would continue

[20] All schools in the MG, RS and RJ areas.

[21] Only students that were expected to receive the full treatment (three years) were considered in our estimations.

[22] 127 school are used in the analysis. One school is lost (from RS area) because it doesn't display desegregated information of students enrollment in the high school. Therefore, it was impossible to retrieve the list of freshman students for this school.

Table 6: ENEM's test scores

	Natural sciences	Human sciences	Language	Math
Treatment Effect	0.080*	0.086*	0.135***	0.116**
(SE)	(0.045)	(0.046)	(0.050)	(0.053)
Lower bound	-0.056	-0.062	0.026	-0.016
(SE)	(0.067)	(0.069)	(0.073)	(0.073)
Upper bound	0.176	0.178	0.235	0.179
(SE)	(0.095)	(0.098)	(0.092)	(0.096)
Trimming	0.061	0.061	0.057	0.057
Nobs	9585	9585	9364	9364
Ncluster	127	127	127	127

NOTE — The outcomes are normalized by the mean and variance of the control group. First line presents the linear regression estimates, standard error estimation takes into account cluster at school level. Lower and upper bound estimations follow Lee (2009), trimming proportion is also displayed, standard errors are computed by bootstrap with 500 replications and cluster at school level. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

to have a positive effect on language scores. Trimming proportion display the percentage of observations that are dropped (trimmed) in the Lee's bounds procedure. For mathematics, we can not discard negative effects with the lower bound, however any effect lower than -1.6% is statistically ruled out. As we analyzed in section III, it is likely to be the case that the true effect is far from the lower bound estimation.

B. COLLEGE AND LABOR MARKET ACCESS

In developing countries the labor market participation decision is fundamental after high school graduation, given the low access to college. Also, a large share of individuals study and work simultaneously. Using data from a Brazilian household survey (PNAD), table 7 presents the proportion of individuals in the categories: only studying, studying and working and only working, by age. When we consider the individuals aged trough 15 to 18 years, who were expected to be at high school, more than 20% study and work at the same time. Less than 30% of the individuals aged 20-22 are studying, among whom more than 60% share their time between college and labor. In panel B, we restrict the sample to the metropolitan areas where the program took place, the scenario is slightly better but not systematically different. One concern is that students that are simultaneously studying and working, may coincide with those students that are not in the expected grade, that is, entered into the school system lately, have dropped out at any point in its educational path or were retained in some grade. In panel C we restrict our statistics to those individuals studying in the expected grade for their age. Even though, more than 18% of high school students and more than half of college ones decide to study and work at the same time.

Therefore, the college-labor decision in developing countries such as Brazil seems to be an important decision after high school completion. Individuals face the options of: (i) continuing

Table 7: School-labor decisions — PNAD 2011

Age	15	16	17	18	19	20	21	22	23	24
Panel A - Brazil										
Studying only	0,744	0,602	0,457	0,259	0,168	0,124	0,092	0,071	0,056	0,042
Studying and Working	0,176	0,255	0,272	0,235	0,205	0,197	0,180	0,148	0,128	0,114
Working only	0,027	0,065	0,151	0,330	0,467	0,520	0,562	0,633	0,656	0,682
Panel B - Program metropolitan areas										
Studying only	0,839	0,730	0,575	0,319	0,217	0,162	0,125	0,078	0,060	0,045
Studying and Working	0,094	0,151	0,200	0,206	0,185	0,188	0,184	0,191	0,162	0,124
Working only	0,023	0,050	0,124	0,299	0,453	0,520	0,552	0,596	0,631	0,681
Panel C - Program metropolitan areas, only individuals studying in correct grade-age										
Studying only	0,897	0,833	0,733	0,731	0,478	0,435	0,404	0,281	0,255	0,250
Studying and Working	0,103	0,167	0,267	0,269	0,522	0,565	0,596	0,719	0,745	0,750

Source: PNAD 2011 (IBGE). NOTE — Proportion of individuals only studying, studying and working and only working by each age. Panel A present the results for all Brazil, meanwhile panel B is restricted to metropolitan areas of Belo Horizonte (MG), Porto Alegre (RS) and Rio de Janeiro (RS), where the program was applied. Panel C is also restricted for all individuals that are studying in the expected grade.

to exclusively invest in formal human capital going to college, (ii) acquiring both formal education and job experience or (iii) going directly to the labor market. Taking the traditional human capital models[23], some possible explanations to the existence of these cases are: (a) there is heterogeneity of individuals and for some of them the marginal return of only-studying is lower than the composite studying-working marginal return, or (b) some individuals are credit-constrained and can only study as long as they work to finance it.

Table 8 presents the results regarding the college-labor decisions. In the control group 46% of individuals are enrolled in higher education in any analyzed period, only 14% are exclusively studying while 32% study and work at the same time. Treatment increases the probability of going to college by 3.5 pp. This result is not statistically significant at the standard confidence levels, but is highly consistent across different exercises and databases. Taking into account the control mean, 46.2%, this effect represents an impact of 7.6%. Table 9 explores this effect for each year after the expected high school completion. In the first years, treatment boosts in 2.6 pp (17.8%) the probability of a student being enrolled in any college-major. The effect increases to 3.6pp (13.0%) in the second year and remains stable until the fifth year. For the first three years the null hypothesis can be rejected at the standard coefficient levels. Considering the bounds estimations, even if all students induced by the treatment to take ENEM went directly to college, treatment would have a positive effect on college entrance. In the annual results, the lower bounds for almost all years are negative but small.

The effect on college attendance is divided equally among studying exclusively and work-study. The effects are respectively 1.7 pp (significant) and 1.8 pp (not significant). The treatment effect on the proportion of individuals who work exclusively is negative (-3.1 pp) and marginally significant. Positive effects are also statistically ruled out by the upper bound estimation.

[23]Becker (1962), Ben-Porath (1967), Mincer (1974) and their developments such as Behrman & Birdsall (1983)

Table 8: College and Labor access

	College	Labor	Only College	Only Labor	College-Labor
Control Mean	0.462	0.721	0.136	0.395	0.326
Treatment Effect	0.035	-0.013	0.017*	-0.031	0.018
(SE)	(0.025)	(0.010)	(0.010)	(0.021)	(0.018)
Lower bound	0.005	-0.020	-0.029	-0.054	-0.016
(SE)	(0.031)	(0.020)	(0.025)	(0.038)	(0.028)
Upper bound	0.058	0.033	0.024	-0.001	0.037
(SE)	(0.038)	(0.032)	(0.021)	(0.031)	(0.028)
Trimming	0.046	0.046	0.046	0.046	0.046
Nobs	14216	14216	14216	14216	14216
Ncluster	127	127	127	127	127

NOTE — First line presents the control mean for each variable, second line exhibits linear regression estimates, standard error estimation takes into account cluster at school level. Lower and upper bound estimations follow Lee (2009), trimming proportion is also displayed, standard errors are computed by bootstrap with 500 replications and cluster at school level.* significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table 9: College effect by each year after high school completion

	Year 1	Year 2	Year 3	Year 4	Year 5
Control Mean	0.146	0.275	0.373	0.450	0.494
Treatment Effect	0.026**	0.036*	0.038*	0.035	0.037
(SE)	(0.012)	(0.019)	(0.023)	(0.033)	(0.033)
Lower bound	-0.019	-0.007	0.001	-0.008	-0.003
(SE)	(0.029)	(0.029)	(0.029)	(0.036)	(0.037)
Upper bound	0.034	0.047	0.054	0.062	0.067
(SE)	(0.020)	(0.028)	(0.033)	(0.046)	(0.048)
Trimming	0.046	0.046	0.046	0.064	0.064
Nobs	14216	14216	14216	10200	10200
Ncluster	127	127	127	97	97

NOTE — First line presents the control mean for each variable, second line exhibits linear regression estimates, standard error estimation takes into account cluster at school level. Lower and upper bound estimations follow Lee (2009), trimming proportion is also displayed, standard errors are computed by bootstrap with 500 replications and cluster at school level. Years 4 and 5 only have data for the MG and RS areas. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

C. UNIVERSITY CATEGORY AND ADMISSION PROCESS

It is important to state that admission in Brazil is college-major specific (henceforth, we will use the term "major" instead of "college-major") . Students must choose a major within college before attending classes and the selectiveness vary not only among colleges, but also across majors. With our data we can assess the selected major, type of admission process and university's funding sources (public or private). In this subsection we will explore the treatment effects in these dimensions.

Brazil has mainly two types of universities: the public ones, where all courses are completely free of charge (there is no tuition), and the private ones, where tuitions are charged and in only a minority of them few students can be awarded with scholarships. In general, private colleges have an overall lower reputation than their public counterparts. Therefore, the attractiveness and consequentially the selectiveness of public universities are higher than that of private colleges. In fact, Binelli et al. (2008) computed that, in 2003, there were on average 9 applicants for each place in public universities, while there were only 1.5 applicants for each place in private colleges.

In table 10 we estimate treatment effects on the university category and on the nature of the selection process. The effects are positive for both public and private universities' attendance, but is significant and large for the public ones: treatment boosts the probability of an individual to attend public courses in 1.8 pp, what represents more than 28.5% of effect. Given the higher quality and selectiveness of these courses, this is a strikingly positive effect. The estimated effect for private colleges is 2 pp or 4.8%, where we cannot reject the null hypothesis of zero effect. Since the proportion of students in public universities is quite small (6.3%), the bounds estimate is particularly conservative. To compute the lower bound the trimming procedure drops almost all students in public universities (4.6% of public college students are trimmed). This negative effect would be true in the unlikely event that all students that took ENEM only by virtue of the treatment were enrolled in public universities.

Table 10: College administration and selection process

	University		Selection process		
	Public	Private	*Vestibular*	ENEM	Quotas
Control Mean	0.063	0.410	0.349	0.155	0.035
Treatment Effect	0.018***	0.020	0.018	0.029*	0.014***
(SE)	(0.007)	(0.021)	(0.020)	(0.015)	(0.004)
Lower bound	-0.029	-0.011	-0.017	-0.011	-0.017
(SE)	(0.016)	(0.028)	(0.029)	(0.029)	(0.010)
Upper bound	0.022	0.043	0.036	0.043	0.016
(SE)	(0.013)	(0.032)	(0.031)	(0.022)	(0.008)
Trimming	0.046	0.046	0.046	0.046	0.046
Nobs	14216	14216	14216	14216	14216
Ncluster	127	127	127	127	127

NOTE —First line presents the control mean for each variable, second line exhibits linear regression estimates, standard error estimation takes into account cluster at school level. Lower and upper bound estimations follow Lee (2009), trimming proportion is also displayed, standard errors are computed by bootstrap with 500 replications and cluster at school level. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

College admission process is evaluated in two different dimensions: (a) type of admission exam, and (b) use of quotas, as part as affirmative actions. Nowadays universities can mainly use two kind of exams in order to rank applicant students: traditional entrance exams (*vestibular*) or ENEM. Prior to 2009, all universities had their own admission exam and applying students used to do several different exams in order to be eligible to more than one university. However in 2009, ENEM was reformulated and public universities were stimulated to use the national exam in their admission process, as a phase or even as the only test taken into account. This stimulus was also increased in the next year by the launching of SISU (*"Sistema de Seleção Unificada"*), a unified system that centralizes the selection of students in the public universities. Li (2016) documents this transition and also the importance of SISU and ENEM's selection in public, specially federal universities, in the last years. According to the author, in 2014, almost 25% of the students used ENEM to access college, among the public universities this number increased to almost 50%. In 2015, 57 public universities offered all places through SISU and other 51 used it partially in the selection process.

Affirmative actions in Brazil started in 2002 in few public universities as quotas that benefit black and public school students[24]. These actions increased over time and in 2012 a federal law was approved, the "Quotas Law", that establishes that 50% of all places in public universities administrated by the federal government should be reserved to public school students[25] (Li 2016).

Therefore the analyzed period (2011-2015) is marked by the increasing importance of public school quotas and ENEM as a college admission process. We also analyzed the treatment effects on the process of entrance for the admitted students. As we can see in table 10, the estimated effect of treatment is positive for all admission processes, but is larger and significant for ENEM (2.9 pp) and quotas (1.4 pp). The effect of admissions through quotas is remarkably large: an increase of 40% (from 3.5% to 4.9%). Since all students in the experiment, treated and control, are eligible to use quotas, the differential effect can be explained by a higher interest in college (public colleges) by the treated students and/or by higher ENEM's scores, since it is this grade that ranks students in order to compete for places in several public universities (with or without quotas).

D. COLLEGE QUALITY AND SELECTIVENESS

The evidence above indicates that treatment increased the amount of students attending college, and this increase is associated to both private and public universities. It is then important to evaluate how good this placement in terms of college-major quality is. We follow Moreira (2016) and construct a measure of quality and selectiveness of majors using freshmen ENEM's test scores. For each course we use use the average score of freshman students in all four subjects of ENEM. Since a student can take the exam in several years we only used the last score available prior to the university admission[26]. Therefore, for every college-major-year we have an average score and are able to rank all majors in Brazil in terms of this score. We will use the percentile of this distribution as a measure of major quality and selectiveness. In order to use representative measures, we eliminate all majors in which only

[24] While the higher education system is characterized by the higher quality of public universities, in the basic education and high school, the scenario is exactly the opposite. Private schools outperform public schools. For more details please check Binelli et al. (2008) and Assunção & Ferman (2015).

[25] The implementation was gradual. Details please check Li (2016).

[26] We only use ENEM's data starting in 2009.

less than 10% of students have available grades.[27]

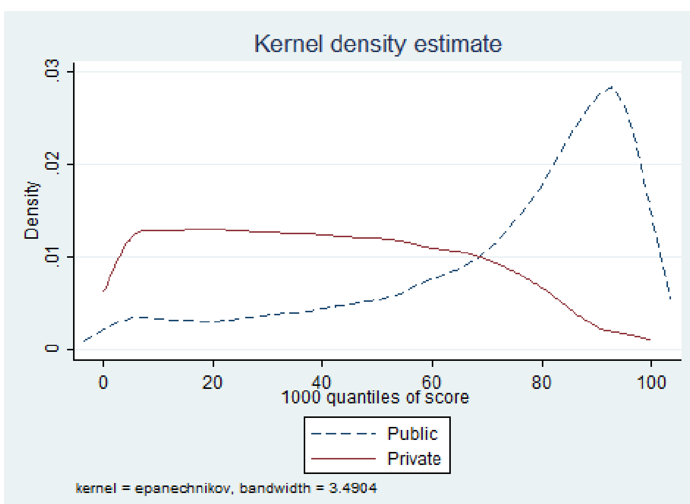
In table 11 we display how many college-majors and students in the higher education are covered by the proposed measure. According to this measure it is possible to assess more than 55% of college-majors in the country. When we take into account the number of students enrolled in majors for which the measure is available, 82.8% are covered in 2010, reaching more than 99% in 2015. The number of students in higher education that took the ENEM exam also increased continuously over the years: in 2015 more than 60% of higher education students had previously taken the exam. We can also use this measure to assess whether the prior believe that public colleges are more selective than the private holds. In the figure 2 below we display kernel density estimations for the quality measure, by type of administration. As expected the public majors exhibit a higher mean (more than 70), while the private distribution is almost uniform in the 5-70 range.

Table 11: Measure of quality and selectiveness of college-majors

Year	2010	2011	2012	2013	2014	2015
N of college-majors	21,456	24,312	25,830	26,464	27,265	28,303
% with quality-measure	0.527	0.559	0.589	0.578	0.563	0.563
N of students in higher eduation	2,508,223	2,642,891	2,873,971	3,080,674	3,313,852	3,123,011
% with ENEM's test scores	0.284	0.392	0.450	0.481	0.531	0.603
% of students with quality-measure	0.828	0.939	0.974	0.984	0.989	0.992

Source: Higher Educational Census (2010-2015), ENEM (2009-2014) — INEP. NOTE — Number of college-majors in Brazil and number of students in higher education. Proportion of majors with available quality-selectiveness measure and proportion of students studying in majors with available quality-selectiveness measure. Quality-selectiveness measure based on Moreira (2016).

Figure 2: Quality-selectiveness measure distribution — Public x Privates (2015)



Source: ENEM 2009-2014 and Higher Education Census 2015— INEP. NOTE — Kernel density estimate of the quality measure (in percentiles of the score ranking) for public and private colleges.

In table 12, we present the treatment effect on the probability of a student being admitted in

[27]We choose not to use ENADE (Student Proficiency National Exam) as measure of quality because of (i) the unavailability of annual data, (ii) the non existent evaluation for several courses, (iii) the lack of incentives to students execute the exam and (iv) the widely known criticism and boycott from several important Brazilian universities.

Table 12: College quality and selectiveness

	Below Median	Above Median	Above 70th	Above 80th	Above 90th
Control Mean	0.156	0.276	0.143	0.075	0.033
Treatment Effect	0.003	0.039*	0.036**	0.026***	0.010**
(SE)	(0.007)	(0.022)	(0.014)	(0.008)	(0.005)
Lower bound	-0.041	-0.001	-0.008	-0.023	-0.023
(SE)	(0.031)	(0.031)	(0.03)	(0.022)	(0.01)
Upper bound	0.013	0.053	0.045	0.03	0.014
(SE)	(0.019)	(0.033)	(0.023)	(0.016)	(0.008)
Trimming	0.046	0.046	0.046	0.046	0.046
Nobs	14216	14216	14216	14216	14216
Ncluster	127	127	127	127	127

NOTE — First line presents the control mean for each variable, second line exhibits linear regression estimates, standard error estimation takes into account cluster at school level. Lower and upper bound estimations follow Lee (2009), trimming proportion is also displayed, standard errors are computed by bootstrap with 500 replications and cluster at school level.* significant at 10% level, ** significant at 5% level and *** significant at 1% level.

a major in several percentiles of the quality-selectiveness distribution. Since we already know that treatment impacted positively ENEM's scores and the measure is comprised exclusively by these grades, we use a lag of the quality measure in order to avoid a mechanical effect on increasing the majors quality-selectiveness. Treatment does not impact admission in below-median majors, the point estimated is 0.3 pp, which is not statistically significant. The effect seems to be totally driven by above-median majors, where the effect is 3.9 pp, thus almost 15%. Our conservative estimation of the lower bound delivers a virtually zero effect estimation, increasing the confidence of a positive effect. Even in the highly selective majors, above the percentiles 70th, 80th and 90th the effects are positive and statistically significant, respectively 3.6 pp (25.4%), 2.6 pp (34.4%) and 1.0 pp (29.7%). Again, when we move upwards in the distribution the proportion of enrolled students falls significantly and the trimming in the lower bound estimation is more harmful. Take, for instance, the 90th percentile, there are 4.3% enrolled students among the treated, the trimming procedure will drop 4.6% of the highest observations, that is, all of the enrolled students.

Summing up all effects exhibited, we can see that treatment shifts students away from the "only working" status towards college, both in "studying exclusively" and "simultaneously studying and working" status. The increase in the number of students is partially split among public and private colleges, but almost all associated with majors ranked above the 50th percentile in the majors' quality-selectiveness measure. There is a positive effect on the admission throughout the use of quotas (1.4 pp). The magnitude of the admission effect in public universities (1.8 pp) almost coincides with the effects on the only studying status (1.7 pp). The same occurs with the effects on private colleges (2.0 pp) and the college-labor status (1.8 pp). Given the fact the public universities do not charge any tuition, we can argue that these students are able to exclusively dedicate to study, while the students attending private colleges have to work to be able to pay their fees. This can picture a scenario where credit constraints are present, but with the available data we cannot conclude for any direction in the credit constraint versus heterogeneity in returns to education debate.

V. HETEROGENEITY

A. *BASELINE PROFICIENCY*

For RJ we have data on students that took the Prova Brazil in the 9th grade in 2009, one year before the intervention. Proficiency is missing for students that were not enrolled in the 9th grade in 2009 (were out of school or were retained in 10th grade), were enrolled in private schools or have missed the exam. We can use this student baseline proficiency to assess heterogeneous effects along the proficiency dimension, as a proxy for unobserved individual ability and quality.

The proportion of students that enroll in ENEM is increasing in the Prova Brazil scores distribution: high-ability individuals are more likely to enroll in ENEM than the low-ability students. Therefore, the proportion of selected students is higher for the students in the 3rd and 4th quartiles than for those in the 1st and 2nd quartile of the Prova Brazil scores' distribution. To perform the heterogeneity exercise we split the sample in quartiles, but to avoid estimations with a lower number of observations in the first quartile we pull it together with the second quartile[28]. The 4th quartile will be specially useful, since it is a subsample with the high-ability of individuals from treatment and control, where the proportion of attrition is small and the differential attrition is non-existent. The bound estimations for this quartile are particularly tight and informative.

Table 13 exhibits the main results by heterogeneity of initial proficiency. Panel A presents treatment effects for the below-median individuals, the point estimations are smaller in magnitude and we cannot reject the null hypothesis of zero effect for any variable. The estimated effect for college attendance is 1.3 pp (5.3%), which seems to be entirely coming from individuals choosing to simultaneously study and work (1.1 pp). Panel B display the results for students in the 3rd quartile of the distribution, the effects are greater in magnitude but again the null hypothesis cannot be rejected. The effect of higher education attendance is estimated in 3.5 pp (8.9%), concentrated in the study-work status (3.2 pp). Furthermore, a small negative effect for the only-labor status is also estimated, -1.3 pp. In this subsample the trimming proportion is very high leading to wide bounds estimations. Lastly, in Panel C, effects for the 4th quartile are presented. The college stock effect is 3.1 pp (5.8%), but there is a shift from the study-labor status (-2.2 pp) to the only study (5.3 pp) option. This indicates that there are effects in the extensive margin, namely that more students are attending college, but also in the intensive margin: students already in college but moving to the only-studying status. As expected, all bounds estimations are tight confirming the signal of the effects.

Table 14 conserves the same structure and exhibits the effects for college administration and selectiveness. Assessing the public-private effects, the point estimations for below-median individuals are concentrated in the private sector, while for the 3rd quartile they are divided between public (2.6 pp) and private (1.6 pp). In the upper quartile we verify a shift movement from the private (-2.6 pp) to the public sector (6.5 pp). The size of this effect is very impressive,

[28]Results are robust to the estimation with the original four quartile, or even terciles and quintiles. The results are available upon request.

Table 13: Main effects - Heterogeneity by baseline proficiency (*Prova Brasil*) [1]

	College	Labor	Only College	Only Labor	College-Labor
PANEL A - Below median					
Control mean	0.252	0.435	0.143	0.326	0.109
Treatment	0.013	0.001	0.002	-0.010	0.011
(SE)	(0.024)	(0.015)	(0.014)	(0.017)	(0.017)
Lower bound	-0.011	-0.030	-0.015	-0.034	0.001
(SE)	(0.049)	(0.061)	(0.030)	(0.047)	(0.029)
Upper bound	0.089	0.069	0.087	0.067	0.102
(SE)	(0.069)	(0.063)	(0.051)	(0.080)	(0.038)
Trimming	0.090	0.090	0.090	0.090	0.090
Nobs	946	946	946	946	946
Ncluster	30	30	30	30	30
PANEL B - 3rd quartile					
Control mean	0.388	0.428	0.226	0.266	0.162
Treatment	0.035	0.019	0.003	-0.013	0.032
(SE)	(0.033)	(0.043)	(0.029)	(0.020)	(0.033)
Lower bound	-0.071	-0.060	-0.132	-0.121	-0.108
(SE)	(0.038)	(0.065)	(0.053)	(0.083)	(0.061)
Upper bound	0.098	0.109	0.039	0.049	0.063
(SE)	(0.080)	(0.053)	(0.059)	(0.034)	(0.036)
Trimming	0.143	0.143	0.143	0.143	0.143
Nobs	673	673	673	673	673
Nclusters	29	29	29	29	29
PANEL C - 4th quartile					
Control mean	0.530	0.438	0.294	0.201	0.237
Treatment	0.031	-0.021	0.053**	0.001	-0.022
(SE)	(0.022)	(0.027)	(0.024)	(0.025)	(0.017)
Lower bound	0.007	-0.023	0.037	0.006	-0.032
(SE)	(0.045)	(0.040)	(0.039)	(0.037)	(0.031)
Upper bound	0.024	-0.006	0.055	0.025	-0.014
(SE)	(0.038)	(0.043)	(0.040)	(0.048)	(0.039)
Trimming	0.017	0.017	0.017	0.017	0.017
Nobs	929	929	929	929	929
Ncluster	30	30	30	30	30

NOTE — In each panel, first line presents the control mean for each variable, second line exhibits linear regression estimates, standard error estimation takes into account cluster at school level. Lower and upper bound estimations follow Lee (2009), trimming proportion is also displayed, standard errors are computed by bootstrap with 500 replications and cluster at school level. Panel A presents the effects for the below median students in the *Prova Brasil* distrubUtion, while panel B is restricted to students in th 3rd quartile and panel C for students in the 4th quartile. This baseline distribution is derived by the scores in *Prova Brasil* in 2009 for the RJ area. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

52.4%, and the bounds estimation yields the lower and upper values of 5.9 pp and 7.7 pp. The lower bound estimation is statistically different from zero.

Table 14: Main effects - Heterogeneity by baseline proficiency (*Prova Brasil*) [2]

	Administration		Selectiveness				
	Public	Private	Below Med	Above Med	Above 70th	Above 80th	Above 90th
PANEL A - Below median							
Control mean	0.030	0.224	0.124	0.119	0.043	0.021	0.009
Treatment	0.005	0.014	0.010	0.009	0.011	0.004	0.007
(SE)	(0.011)	(0.020)	(0.017)	(0.015)	(0.011)	(0.006)	(0.005)
Lower bound	-0.009	-0.002	0.011	-0.019	-0.004	-0.008	-0.001
(SE)	(0.013)	(0.044)	(0.030)	(0.030)	(0.016)	(0.010)	(0.007)
Upper bound	0.027	0.099	0.112	0.082	0.046	0.017	0.012
(SE)	(0.014)	(0.064)	(0.041)	(0.046)	(0.017)	(0.009)	(0.005)
Trimming	0.090	0.090	0.090	0.090	0.090	0.090	0.090
Nobs	946	946	946	946	946	946	946
Ncluster	30	30	30	30	30	30	30
PANEL B - 3rd quartile							
Control mean	0.061	0.33	0.171	0.193	0.098	0.064	0.024
Treatment	0.026	0.016	0.005	0.047**	0.019	0.010	0.000
(SE)	(0.020)	(0.034)	(0.028)	(0.021)	(0.013)	(0.013)	(0.005)
Lower bound	-0.061	-0.090	-0.121	-0.115	-0.098	-0.064	-0.024
(SE)	(0.022)	(0.049)	(0.048)	(0.066)	(0.031)	(0.02)	(0.006)
Upper bound	0.029	0.079	0.050	0.056	0.013	0.006	-0.001
(SE)	(0.023)	(0.067)	(0.050)	(0.038)	(0.027)	(0.021)	(0.010)
Trimming	0.143	0.143	0.143	0.143	0.143	0.143	0.143
Nobs	673	673	673	673	673	673	673
Nclusters	29	29	29	29	29	29	29
PANEL C - 4th quartile							
Control mean	0.124	0.421	0.166	0.356	0.170	0.109	0.050
Treatment	0.065**	-0.026	-0.013	0.049*	0.090***	0.053***	0.026**
(SE)	(0.027)	(0.023)	(0.022)	(0.027)	(0.029)	(0.020)	(0.010)
Lower bound	0.059	-0.049	-0.024	0.029	0.072	0.042	0.027
(SE)	(0.032)	(0.036)	(0.032)	(0.034)	(0.034)	(0.029)	(0.026)
Upper bound	0.077	-0.031	-0.005	0.046	0.090	0.061	0.046
(SE)	(0.039)	(0.034)	(0.039)	(0.037)	(0.046)	(0.04)	(0.025)
Trimming	0.017	0.017	0.017	0.017	0.017	0.017	0.017
Nobs	929	929	929	929	929	929	929
Ncluster	30	30	30	30	30	30	30

NOTE — In each panel, first line presents the linear regression estimates, standard error estimation takes into account cluster at school level. Lower and Upper bound estimations follow Lee (2009), trimming proportion is also displayed, standard errors are computed by bootstrap with 500 replications and cluster at school level. Panel A presents the effects for the below median students in the *Prova Brasil* distribution, while panel B is restricted to students in th 3rd quartile and panel C for students in the 4th quartile. This baseline distribution is derived by the scores in *Prova Brasil* in 2009 for the RJ area. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

The small effect of college attendance for below-median students is barely detected when we divide the college by the quality-selectiveness measure. For the 3rd quartile, the effect is concentrated in the above 50th majors, a 4.7 pp (24.3%). Since the effect for the above 70th is lower, students seems to be attending mostly majors in the 50th-70th interval. For the high-ability individuals the size of the effects is also large: 9.0 pp (53.2%), 5.3 pp (48.8%) and 2.6 pp (51.9%) respectively for above 70th, 80th and 90th college-majors. All students seem to improve with treatment, the estimated effect on college attendance is positive for all quartiles. However, high-ability students seem to really make progress throughout the treatment, increasing their chances to attend high-quality and selective institutions, specially public ones.

B. GENDER

The literature presents mixed effects when assessing heterogeneity by gender. One group of articles reveals effects concentrated on women (Deming et al. 2014), another presents evidence in favor of greater effects for men (Dustmann et al. 2003, Chetty et al. 2011, Dynarski et al. 2013), while yet another finds no significant differential effects by gender (Angrist et al. 2016, Fredriksson et al. 2013). There is also evidence, such as Lavy (2016), where this heterogeneity is not monotonic even within the paper, the schooling effect is larger for girls while the gains in earnings are greater for boys.

Despite larger negative effects in labor market for men (-2.3 pp for labor decision and -4.5 pp for only labor status, both statistically significant) than for women (-0.6 pp and -1.9 pp, respectively), we find no heterogeneity by gender. The results are displayed in the appendix B.

These results are different from the findings of Anderson (2008) which reveal a clear pattern of greater effects for women when evaluating early childhood interventions. These mixed results may indicate that the diversity of interventions and environments where they are applied may produce setups that are favorable for each gender. Another possible conclusion is that when assessing this heterogeneity without any strong theoretical guidance, we are susceptible to give importance on arguably random realizations. It appears that some work within the topic is necessary to shed some light across the gender heterogeneity findings.

VI. ADDITIONAL EXERCISES

Aiming at checking the sensibility of our results we perform two additional exercises. The first is based on the construction of a different panel with more observations. As detailed in section III, we can only link individuals in ENEM through their unique student identification, and this student ID is absent for at least 30% of the individuals in ENEM's dataset. We can use an alternative strategy that avoids this specific losing data mechanism. In ENEM's registration, students may inform the school they belong and also their status: such as completing or not completing high school at the end of the year. With this information we can use ENEM's information in the last year of each intervention (2012 and 2010) and track senior high school students that belong to the schools participating in the experiment. This procedure allows us to identify more than 27,000 students, among which 19,675 (or 72.2%) are selected. We are increasing the number of observations but losing the initial treatment list in the freshman years in high school. The results are qualitatively the same and are displayed in the appendix D. It is worth to highlight how the overall effect in college attendance is very similar: 3.4 pp instead of 3.5 pp in the benchmarking dataset.

As we discussed and it is presented in the appendix C, using students proficiency in the Prova Brasil exam at the baseline year, SP and VP sample are unbalanced, treated students exhibit higher averages in the baseline. There are several reasons that may explain these findings: (i) since treatment was actually initiated at the end of 2009, it had also produced effects in that same year, (ii) since the final allocation in treatment and control groups was known, better students decided to move to treated schools, or (iii) unlucky casual result from the randomization.

If (iii) is the explanation then the schools would be unbalanced in previous exams. We tested balance in the Prova Brazil in 2007, two years before the intervention early-beginning[29]. We reject the null hypothesis of equal means for the VP area, indicating that selected schools had already higher proficiency previous to the randomization. We can also indirectly test (i), comparing schools' score in 2009, therefore controlling for the 2007 performance. We again reject the null-hypothesis, evidencing that (i) may also explain the unbalance in the exercise. Results are displayed in appendix C. We could not test (ii), but is likely to be also one of the driving forces of the detected unbalance.

The second exercise reintroduces the areas of SP and VP into the original sample and reestimate all the effects. The results are presented in the appendix E. Again, the qualitative results are all aligned with the presented results. As a comparison, the overall effect in college attendance is estimated in 3.2 pp (significant at 10% level), while our benchmark estimations displays a 3.5 pp effect. All the remaining results also display similar estimations. This expanded panel is also useful to compare with the previous proficiency-heterogeneity exercises, since they could only be done with RJ information. Attention must be paid to the previously mentioned unbalance. The results are weaker than we expected, nevertheless, the overall picture remains unchanged.

VII. CONCLUSION

In this research, we make use of *Jovem de Futuro*, an experimental intervention in Brazilian public high schools aiming at improving the quality of education through improving management practices and allocating endowments to beneficiary schools. The program is associated with a significant increase in students proficiency, already established by several previous evaluations. Therefore we assume the experiment as an exogenous increase of educational quality and use the randomization procedure in order to evaluate the impacts of a better education on the individual human capital decisions, namely the decision to work and/or study after high school completion.

Our estimations are based on an ITT strategy, considering as treated the students enrolled in the freshman year of lottery-winners high schools in the first year of the experiment. This status is regardless the future decisions of the student (to be or not at this high school) and of the school (to comply or not with treatment assignment). The same holds for the control group. We investigated outcomes after high school completion, students were expected to be

[29] We use school averages here, since in 2007 other students were enrolled in the 9th grade.

treated through the three years of high school.

Students from schools that received the positive educational quality shock are more likely to attend college (3.5 pp), to attend public colleges (1.8 pp), to study full time (1.7 pp) and to attend high quality and selective majors[30] (3.6 pp). The effect on college admission seems to be specially driven by ENEM (2.9 pp) and quotas (1.4 pp), being the latter strongly related to the gains in ENEM's scores (in average 10.42% of standard-deviation). The increase in the stock of students in college is split among those only studying (1.7 pp) and those in the studying-working status (1.8 pp). There is also a reduction in the probability of individuals with only working status, specially concentrated in men (-4.5 pp).

When assessing the heterogeneity by initial proficiency, the gains in college enrollment for individuals in the lower tail distribution of initial proficiency are lower and associated with private colleges, study-work status, although neither results are statistically significant. For individuals in the 3rd quartile, the magnitude of the higher education attendance is greater split equally between public and private colleges and in only studying and work-study status. Their gains are related to majors above the median of quality-selectiveness distribution. Lastly, for individuals in the upper quartile, thus above the stock effect on college attendance, treatment shifts individuals from private to public college, from work-study toward study full-time and from below 70th to above 70th majors (quality-selectiveness measure). Some effects such as the increase by more than 52% in public college enrollment and in majors above 70th in the quality-selectiveness measure are very impressive.

Despite the small and not significant effect of treatment in selection we also present bounds estimations based on Lee (2009) to deal with selection for all results. We use this procedure to qualify how several estimated results survive even under the unlikely possibility that all compliers individuals take the highest results in the evaluated outcomes. Using the RJ area we could also generate some evidence that this is not the case.

These results indicate decisions towards more and better college enrollment associated with a higher education quality. Students in a better schooling environment are more likely to move away from the market in direction to college, in the full-time study or study and work simultaneously. For high ability individuals we also can detect a movement from the study and work simultaneously to the only-studying status. We are not able to specify if this is related to the expansion of the returns to schooling function or to credit constraint, since students are more likely to attend public colleges where no tuition is charged.

The results in the stock of college attendance (more college effect) are arguably to hold in general equilibrium, that is, if all Brazilian public schools receive treatment. However, the same is not true for the effects associated with students moving to higher quality majors (better college effect). As we could see, individuals are attending public schools specially through the use of quotas for public students (and possibly for black and other minorities). If treatment would be extended to all public schools we cannot foresee how the equilibrium in the public college admission would be.

Therefore, students in Brazilian public schools are responding with more and better college

[30] Above the 70th percentile of the quality-selectiveness distribution

enrollment after an increase in the quality of education. This add with the recent findings in Angrist et al. (2016) and Lavy (2016), in order to identify that high school interventions indeed effect students decisions to continue to invest in formal human capital. High school is not too late. We contribute to the literature estimating the effects of increasing quality in a given group of schools (rather than moving students to better schools) and in a developing country where the labor-college decision is remarkably relevant. We are not able to estimate the treatment effects on college completion and subsequently on earnings, which are interesting questions whose answers depends on the time expected for college graduation and market entrance after college.

References

Anderson, M. L. (2008), 'Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects', *Journal of the American statistical Association* **103**(484), 1481–1495.

Angrist, J. D., Cohodes, S. R., Dynarski, S. M., Pathak, P. A. & Walters, C. R. (2016), 'Stand and deliver: Effects of boston's charter high schools on college preparation, entry, and choice', *Journal of Labor Economics* **34**(2), 275–318.

Assunção, J. & Ferman, B. (2015), Does affirmative action enhance or undercut investment incentives? evidence from quotas in brazilian public universities. Working paper.

Barro, R. J. & Lee, J. W. (2013), 'A new data set of educational attainment in the world, 1950–2010', *Journal of development economics* **104**, 184–198.

Barros, R., de Carvalho, M., Franco, S. & Rosalém, A. (2012), 'Impacto do projeto jovem de futuro.', *Estudos em Avaliação Educacional* **23**(51), 214–226.

Becker, G. S. (1962), 'Investment in human capital: A theoretical analysis', *The journal of political economy* pp. 9–49.

Behrman, J. R. & Birdsall, N. (1983), 'The quality of schooling: quantity alone is misleading', *The American Economic Review* **73**(5), 928–946.

Behrman, J. R., Rosenzweig, M. R. & Taubman, P. (1996), 'College choice and wages: Estimates using data on female twins', *The Review of Economics and Statistics* pp. 672–685.

Ben-Porath, Y. (1967), 'The production of human capital and the life cycle of earnings', *The Journal of Political Economy* pp. 352–365.

Binelli, C., Meghir, C. & Menezes-Filho, N. (2008), Education and wages in brazil. Working paper.

Card, D. (1999), 'The causal effect of education on earnings', *Handbook of labor economics* **3**, 1801–1863.

Card, D. & Krueger, A. B. (1992), 'Does school quality matter? returns to education and the characteristics of public schools in the united states', *Journal of Political Economy* **100**(1), 1–40.

Card, D. & Krueger, A. B. (1996), *Labor market effects of school quality: theory and evidence*, Brooklings Institution.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W. & Yagan, D. (2011), 'How does your kindergarten classroom affect your earnings? evidence from project star', *The Quarterly Journal of Economics* **126**(4), 1593–1660.

Chetty, R., Friedman, J. N. & Rockoff, J. E. (2014), 'Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood', *The American Economic Review* **104**(9), 2633–2679.

Coleman, J. S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F. & York, R. (1966), 'The coleman report', *Equality of Educational Opportunity* .

Dearden, L., Ferri, J. & Meghir, C. (2002), 'The effect of school quality on educational attainment and wages', *Review of Economics and Statistics* **84**(1), 1–20.

Deming, D. J., Hastings, J. S., Kane, T. J. & Staiger, D. O. (2014), 'School choice, school quality, and postsecondary attainment', *The American economic review* **104**(3), 991–1013.

Dix-Carneiro, R. & Kovak, B. K. (2016), Trade liberalization and regional dynamics, Technical report, mimeo, Duke University, https://sites. google. com/site/rafaeldixcarneiro/research.

Dobbie, W. & Fryer Jr, R. G. (2015), 'The medium-term impacts of high-achieving charter schools', *Journal of Political Economy* **123**(5), 985–1037.

Dustmann, C., Puhani, P. A., Schönberg, U. et al. (2012), 'The long-term effects of school quality on labor market outcomes and educational attainment', *Centre for Research and Analysis of Migration (CReAM), Department of Economics, University College London CReAM Discussion Paper Series* **1208**.

Dustmann, C., Rajah, N. & Soest, A. (2003), 'Class size, education, and wages', *The Economic Journal* **113**(485).

Dynarski, S., Hyman, J. & Schanzenbach, D. W. (2013), 'Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion', *Journal of Policy Analysis and Management* **32**(4), 692–717.

Ferreira, R. S. (2014), Alocação de recursos financeiros e a função de produção escolar, Master's thesis, EESP-FGV, Brasil.

Fredriksson, P., Öckert, B. & Oosterbeek, H. (2013), 'Long-term effects of class size', *The Quarterly Journal of Economics* **128**(1), 249–285.

Gill, K. B. T. S. B., Zimmer, R. et al. (2008), Going beyond test scores evaluating charter school impact on educational attainment in chicago and florida, Technical report, Mathematica Policy Research.

Heckman, J. (1974), 'Shadow prices, market wages, and labor supply', *Econometrica: journal of the econometric society* pp. 679–694.

Heckman, J. J. (1979), 'Sample selection bias as a specification error', *Econometrica* **47**(1), 153–161.

Instituto Unibanco (2010), 'Relatório de atividades 2009'.

Instituto Unibanco (2011), 'Relatório de atividades 2010'.

Instituto Unibanco (2012), 'Relatório de atividades 2011'.

Krueger, A. B. & Whitmore, D. M. (2001), 'The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project star', *The Economic Journal* **111**(468), 1–28.

Lavy, V. (2016), 'Teachers' pay for performance in the long-run: The dynamic pattern of treatment effects on students' educational and labor market outcomes in adulthood'.

Lee, D. S. (2009), 'Training, wages, and sample selection: Estimating sharp bounds on treatment effects', *The Review of Economic Studies* **76**(3), 1071–1102.

Li, D. L. (2016), O novo ENEM e a plataforma Sisu: efeitos sobre a migração e a evasão estudantil, Master's thesis, FEA-USP, Brasil.

Mincer, J. (1974), ‘Schooling, experience, and earnings. human behavior & social institutions no. 2.’.

Moreira, D. (2016), Success spills over how awards affect winners’ and peers’ performance in brazil. Working paper.

Oliva, B. T. (2014), Três ensaios em economia da educação, PhD thesis, EESP-FGV, Brasil.

Rosa, L. S. N. (2015), Avaliação de Impacto do Programa Jovens de Futuro, Master’s thesis, IPE-USP, Brasil.

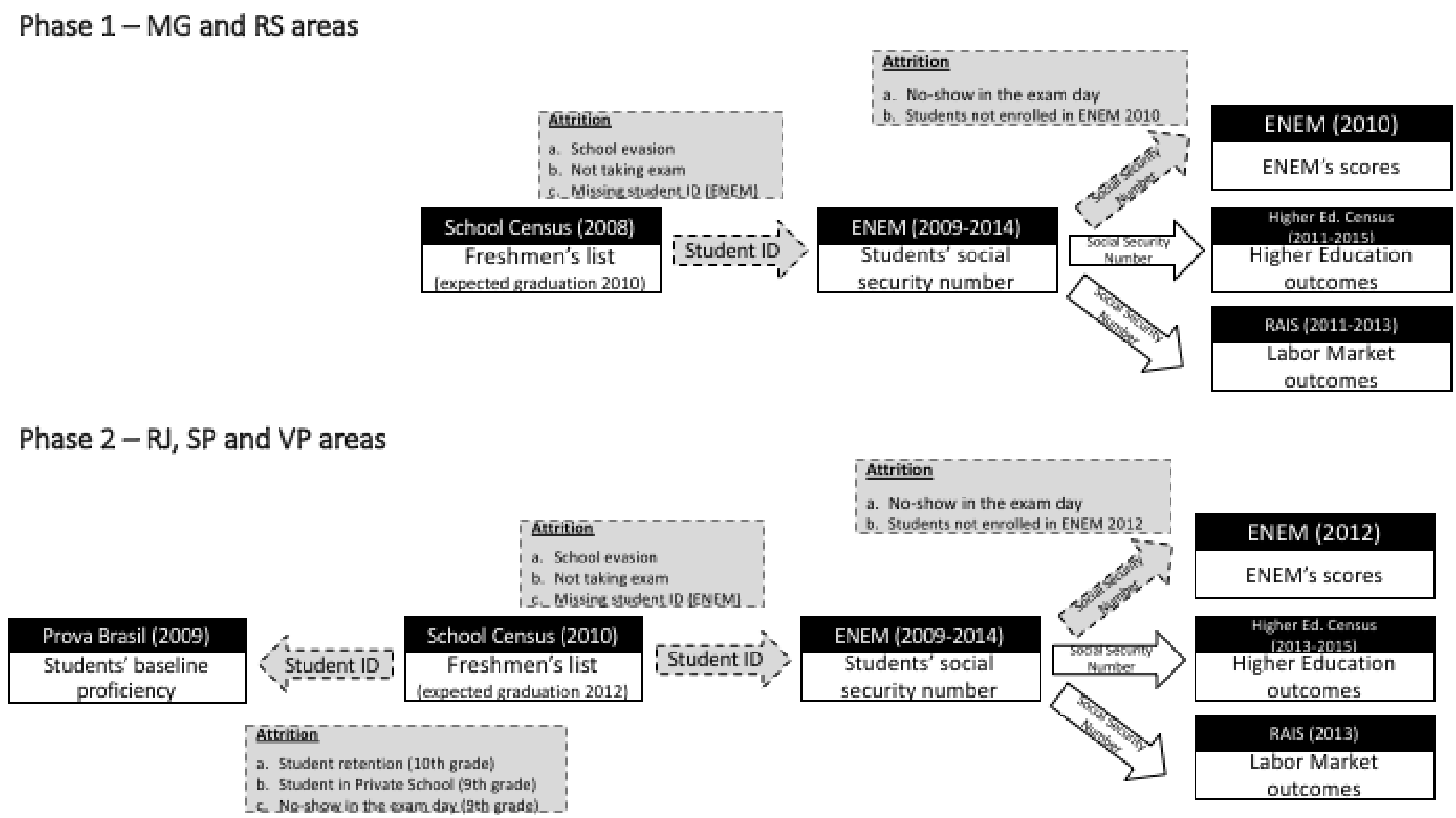
Silva, S. (2010), Análise dos efeitos de programas educacionais: o caso projeto Jovem de Futuro do Instituto Unibanco, Master’s thesis, EESP-FGV, Brasil.

Takeuti, C. (2014), Os efeitos da gestão escolar com autonomia financeira, Master’s thesis, EESP-FGV, Brasil.

I. APPENDIX A

DATASETS

Figure 3: Datasets



II. APPENDIX B

HETEROGENEITY BY GENDER

Table 15: Main effects - Heterogeneity by gender [1]

	College	Labor	Only College	Only Labor	College-labor
PANEL A - Men					
Control mean	0.440	0.737	0.127	0.423	0.313
Treatment	0.037	-0.023**	0.015	-0.045*	0.022
(SE)	(0.029)	(0.011)	(0.011)	(0.027)	(0.022)
Lower bound	0.022	-0.031	0.001	-0.052	0.008
(SE)	(0.038)	(0.024)	(0.026)	(0.038)	(0.032)
Upper bound	0.047	-0.006	0.026	-0.027	0.033
(SE)	(0.040)	(0.031)	(0.024)	(0.039)	(0.031)
Trimming	0.027	0.027	0.027	0.027	0.027
Nobs	5545	5545	5545	5545	5545
Ncluster	127	127	127	127	127
PANEL B - Women					
Control mean	0.477	0.711	0.143	0.377	0.334
Treatment	0.031	-0.006	0.017	-0.019	0.014
(SE)	(0.024)	(0.012)	(0.011)	(0.02)	(0.017)
Lower bound	-0.007	-0.014	-0.047	-0.054	-0.032
(SE)	(0.027)	(0.021)	(0.023)	(0.039)	(0.026)
Upper bound	0.065	0.058	0.026	0.019	0.041
(SE)	(0.039)	(0.032)	(0.02)	(0.028)	(0.030)
Trimming	0.060	0.060	0.060	0.060	0.060
Nobs	8671	8671	8671	8671	8671
Ncluster	127	127	127	127	127

NOTE — In each panel, first line presents the control mean for each variable, second line exhibits linear regression estimates, standard error estimation takes into account cluster at school level. Lower and upper bound estimations follow Lee (2009), trimming proportion is also displayed, standard errors are computed by bootstrap with 500 replications and cluster at school level. Panel A presents the effects for men and Panel B for women. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table 16: Main effects - Heterogeneity by gender [2]

	Administration		Selectiveness				
	Public	Private	Below Med	Above Med	Above 70th	Above 80th	Above 90th
PANEL A - Men							
Control mean	0.062	0.390	0.109	0.300	0.163	0.089	0.041
Treatment	0.018**	0.024	0.002	0.039	0.038**	0.023**	0.008
(SE)	(0.008)	(0.025)	(0.009)	(0.025)	(0.018)	(0.011)	(0.007)
Lower bound	0.000	0.011	-0.010	0.024	0.021	0.006	-0.006
(SE)	(0.018)	(0.034)	(0.024)	(0.036)	(0.032)	(0.025)	(0.013)
Upper bound	0.026	0.035	0.015	0.049	0.046	0.031	0.019
(SE)	(0.015)	(0.036)	(0.023)	(0.037)	(0.030)	(0.021)	(0.011)
Trimming	0.027	0.027	0.027	0.027	0.027	0.027	0.027
Nobs	5545	5545	5545	5545	5545	5545	5545
Ncluster	127	127	127	127	127	127	127
PANEL B - Women							
Control mean	0.063	0.424	0.187	0.259	0.130	0.065	0.028
Treatment	0.019**	0.014	-0.001	0.039*	0.037***	0.028***	0.012**
(SE)	(0.007)	(0.021)	(0.009)	(0.021)	(0.014)	(0.008)	(0.005)
Lower bound	-0.029	-0.025	-0.06	-0.015	-0.025	-0.033	-0.016
(SE)	(0.015)	(0.025)	(0.031)	(0.029)	(0.028)	(0.018)	(0.009)
Upper bound	0.023	0.047	0.013	0.058	0.048	0.034	0.015
(SE)	(0.014)	(0.032)	(0.018)	(0.033)	(0.022)	(0.015)	(0.008)
Trimming	0.06	0.06	0.06	0.06	0.06	0.06	0.06
Nobs	8671	8671	8671	8671	8671	8671	8671
Ncluster	127	127	127	127	127	127	127

NOTE — In each panel, first line presents the control mean for each variable, second line exhibits linear regression estimates, standard error estimation takes into account cluster at school level. Lower and upper bound estimations follow Lee (2009), trimming proportion is also displayed, standard errors are computed by bootstrap with 500 replications and cluster at school level. Panel A presents the effects for men and Panel B for women. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

III. **APPENDIX C**

BASELINE BALANCE — *PROVA BRASIL* 2009 - RJ, SP AND VP AREAS

Table 17: Baseline balance [1]

	RJ		SP		VP	
	Language	Math	Language	Math	Language	Math
Treatment effect	0.119	0.092	0.261***	0.271***	0.357***	0.369***
(SE)	(0.155)	(0.136)	(0.080)	(0.097)	(0.132)	(0.121)
Nobs	9950	9950	12359	12359	8514	8514
Ncluster	30	30	41	41	40	40

NOTE — Results based on the following regression: $Y_{ijs} = \alpha + \beta T_{js} + \eta_s + \varepsilon_{ijs}$, where Y_{ijs} is the *Prova Brasil* test score for student i, in school j and strata s, T is the treatment indicator and η_s strata fixed effect. Cluster at school level is considered. Outcome of interest is the student scores at *Prova Brasil* 2009. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table 18: Baseline balance [2]

	2007 SP		2007 VP		2009 SP		2009 VP	
	Language	Math	Language	Math	Language	Math	Language	Math
Treatment effect	0.164	0.161	0.775*	0.811*	0.276*	0.215	0.288	-0.007
(SE)	(0.146)	(0.177)	(0.449)	(0.443)	(0.145)	(0.154)	(0.332)	(0.297)
2007 control	No	No	No	No	Yes	Yes	Yes	Yes
Nobs	41	41	33	33	40	40	33	33

NOTE — Results based on the following regression: $Y_{js} = \alpha + \beta T_{js} + \eta_s + \varepsilon_{js}$, where Y_{js} is the *Prova Brasil* test score for school j and strata s, T is the treatment indicator and η_s strata fixed effect. Outcome of interest is the schools' scores at *Prova Brasil* 2007 and 2009. In columns 6-9, the school's scores of 2007 is added as a regressor. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

IV. APPENDIX D

ALTERNATIVE DATASET [1] — USING ENEM AS PRIMARY SOURCE OF INFORMATION

Table 19: Main effects - Alternative dataset (ENEM) [1]

	Human Capital decisions					Administration	
	College	Labor	Only College	Only Labor	College-Labor	Public	Private
Control Mean	0.496	0.713	0.146	0.364	0.350	0.070	0.438
Treatment	0.034	0.009	0.009	-0.016	0.025	0.013**	0.023
(SE)	(0.026)	(0.008)	(0.009)	(0.021)	(0.019)	(0.007)	(0.022)
Lower bound	-0.026	-0.015	-0.077	-0.066	-0.044	-0.067	-0.038
(SE)	(0.024)	(0.017)	(0.02)	(0.037)	(0.023)	(0.014)	(0.021)
Upper bound	0.070	0.081	0.019	0.030	0.051	0.015	0.058
(SE)	(0.037)	(0.023)	(0.015)	(0.020)	(0.026)	(0.011)	(0.031)
Trimming	0.089	0.089	0.089	0.089	0.089	0.089	0.089
Nobs	19675	19675	19675	19675	19675	19675	19675
Ncluster	128	128	128	128	128	128	128

NOTE — First line presents the control mean for each variable, second line exhibits linear regression estimates, standard error estimation takes into account cluster at school level. Lower and upper bound estimations follow Lee (2009), trimming proportion is also displayed, standard errors are computed by bootstrap with 500 replications and cluster at school level.* significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table 20: Main effects - Alternative dataset (ENEM) [2]

	Selectiveness					Admission		
	Below Med	Above Med	Above 70th	Above 80th	Above 90th	Vestibular	ENEM	Quotas
Control Mean	0.102	0.182	0.100	0.055	0.024	0.333	0.152	0.037
Treatment	-0.007	0.027*	0.021**	0.013**	0.007**	0.006	0.025*	0.009**
(SE)	(0.005)	(0.016)	(0.010)	(0.006)	(0.003)	(0.017)	(0.015)	(0.004)
Lower bound	-0.087	-0.061	-0.066	-0.055	-0.024	-0.064	-0.058	-0.037
(SE)	(0.019)	(0.025)	(0.019)	(0.01)	(0.004)	(0.022)	(0.024)	(0.008)
Upper bound	0.004	0.035	0.024	0.014	0.008	0.032	0.037	0.010
(SE)	(0.009)	(0.02)	(0.014)	(0.009)	(0.005)	(0.024)	(0.018)	(0.006)
Trimming	0.089	0.089	0.089	0.089	0.089	0.089	0.089	0.089
Nobs	14216	14216	14216	14216	14216	14216	14216	14216
Ncluster	127	127	127	127	127	127	127	127

NOTE — First line presents the linear regression estimates, standard error estimation takes into account cluster at school level. Lower and Upper bound estimations follow Lee (2009), trimming proportion is also displayed, standard errors are computed by bootstrap with 500 replications and cluster at school level.* significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table 21: Main effects - Alternative dataset (ENEM) [3] - Heterogeneity by baseline proficiency

	Human capital decisions					Administration		Selectiveness				
	College	Labor	Only College	Only Labor	College-Labor	Public	Private	Below Med	Above Med	Above 70th	Above 80th	Above 90th
PANEL A - Below median												
Control Mean	0.318	0.469	0.173	0.325	0.145	0.040	0.281	0.122	0.122	0.052	0.031	0.014
Treatment Effect	0.033	0.021	0.000	-0.013	0.034	0.015	0.021	-0.003	0.027	0.020**	0.011	0.003
(SE)	(0.034)	(0.027)	(0.019)	(0.014)	(0.028)	(0.014)	(0.026)	(0.018)	(0.019)	(0.010)	(0.010)	(0.006)
Nobs	1055	1055	1055	1055	1055	1055	1055	1055	1055	1055	1055	1055
Ncluster	30	30	30	30	30	30	30	30	30	30	30	30
PANEL B - 3rd quartile												
Control Mean	0.473	0.495	0.237	0.258	0.237	0.092	0.385	0.134	0.219	0.095	0.06	0.025
Treatment Effect	0.061	0.031	0.035*	0.004	0.027	0.036**	0.042	0.049	0.027	0.035*	0.030***	0.003
(SE)	(0.039)	(0.03)	(0.018)	(0.035)	(0.028)	(0.018)	(0.038)	(0.038)	(0.034)	(0.02)	(0.012)	(0.008)
Nobs	527	527	527	527	527	527	527	527	527	527	527	527
Ncluster	30	30	30	30	30	30	30	30	30	30	30	30
PANEL C - 4th quartile												
Control Mean	0.567	0.483	0.303	0.218	0.264	0.146	0.441	0.134	0.303	0.172	0.100	0.054
Treatment Effect	0.037	-0.059	0.060	-0.035	-0.024	0.062	-0.013	-0.038**	0.056	0.032	0.045*	0.039**
(SE)	(0.038)	(0.040)	(0.038)	(0.033)	(0.035)	(0.041)	(0.036)	(0.018)	(0.034)	(0.028)	(0.027)	(0.017)
Nobs	527	527	527	527	527	527	527	527	527	527	527	527
Ncluster	27	27	27	27	27	27	27	27	27	27	27	27

NOTE — In each panel, first line presents the control mean for each variable, second line exhibits linear regression estimates, standard error estimation takes into account cluster at school level. Panel A presents the effects for the below median students in the *Prova Brasil* distribution, while panel B is restricted to students in th 3rd quartile and panel C for students in the 4th quartile. This baseline distribution is derived by the scores in *Prova Brasil* in 2009 for the RJ area. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.

V. APPENDIX E

ALTERNATIVE DATASET [2] — INCLUDING SP AND VP AREAS

Table 22: Main effects - Alternative dataset (SP) [2]

	Human Capital decisions					Administration	
	College	Labor	Only College	Only Labor	College-Labor	Public	Private
Control Mean	0.462	0.695	0.146	0.379	0.316	0.055	0.415
Treatment Effect	0.032*	-0.015*	0.019**	-0.027*	0.013	0.015***	0.020
(SE)	(0.018)	(0.008)	(0.008)	(0.016)	(0.013)	(0.005)	(0.016)
Lower bound	-0.011	-0.036	-0.045	-0.076	-0.043	-0.031	-0.026
(SE)	(0.024)	(0.017)	(0.019)	(0.030)	(0.022)	(0.012)	(0.022)
Upper bound	0.073	0.048	0.034	0.009	0.042	0.020	0.059
(SE)	(0.030)	(0.025)	(0.017)	(0.024)	(0.022)	(0.010)	(0.025)
Trimming	0.086	0.086	0.086	0.086	0.086	0.086	0.086
Nobs	19569	19569	19569	19569	19569	19569	19569
Ncluster	208	208	208	208	208	208	208

NOTE —First line presents the control mean for each variable, second line exhibits linear regression estimates, standard error estimation takes into account cluster at school level. Lower and upper bound estimations follow Lee (2009), trimming proportion is also displayed, standard errors are computed by bootstrap with 500 replications and cluster at school level.* significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table 23: Main effects - Alternative dataset (SP) [2]

	Selectiveness					Admission		
	Below Med	Above Med	Above 70th	Above 80th	Above 90th	Vestibular	ENEM	Quotas
Control Mean	0.191	0.245	0.123	0.066	0.027	0.358	0.165	0.032
Treatment Effect	0.003	0.032**	0.029***	0.020***	0.009**	0.015	0.026**	0.011***
(SE)	(0.006)	(0.016)	(0.011)	(0.006)	(0.004)	(0.015)	(0.011)	(0.003)
Lower bound	-0.061	-0.020	-0.022	-0.027	-0.019	-0.035	-0.037	-0.018
(SE)	(0.024)	(0.024)	(0.022)	(0.016)	(0.007)	(0.022)	(0.023)	(0.008)
Upper bound	0.024	0.053	0.041	0.026	0.014	0.049	0.048	0.015
(SE)	(0.015)	(0.025)	(0.017)	(0.012)	(0.006)	(0.025)	(0.018)	(0.006)
Trimming	0.086	0.086	0.086	0.086	0.086	0.086	0.086	0.086
Nobs	14216	14216	14216	14216	14216	14216	14216	14216
Ncluster	127	127	127	127	127	127	127	127

NOTE — First line presents the control mean for each variable, second line exhibits linear regression estimates, standard error estimation takes into account cluster at school level. Lower and upper bound estimations follow Lee (2009), trimming proportion is also displayed, standard errors are computed by bootstrap with 500 replications and cluster at school level.* significant at 10% level, ** significant at 5% level and *** significant at 1% level.

Table 24: Main effects - Alternative dataset (SP) [3] - Heterogeneity by baseline proficiency

	Human capital decisions					Administration		Selectiveness				
	College	Labor	Only College	Only Labor	College-Labor	Public	Private	Below Med	Above Med	Above 70th	Above 80th	Above 90th
PANEL A - 1sr quartile												
Control Mean	0.276	0.565	0.113	0.402	0.163	0.020	0.256	0.186	0.085	0.033	0.02	0.005
Treatment Effect	0.001	-0.008	-0.010	-0.019	0.011	-0.013**	0.014	0.034	-0.032*	-0.008	-0.013*	-0.002
(SE)	(0.024)	(0.003)	(0.017)	(0.030)	(0.019)	(0.006)	(0.023)	(0.026)	(0.018)	(0.010)	(0.008)	(0.001)
Nobs	795	795	795	795	795	795	795	795	795	795	795	795
Ncluster	105	105	105	105	105	105	105	105	105	105	105	105
PANEL B - 2nd quartile												
Control Mean	0.350	0.578	0.137	0.365	0.213	0.022	0.329	0.234	0.105	0.031	0.012	0.004
Treatment Effect	0.013	-0.003	0.000	-0.016	0.013	-0.008	0.018	0.015	-0.002	-0.002	0.001	0.003
(SE)	(0.020)	(0.022)	(0.014)	(0.022)	(0.020)	(0.006)	(0.019)	(0.019)	(0.012)	(0.006)	(0.005)	(0.003)
Nobs	1369	1369	1369	1369	1369	1369	1369	1369	1369	1369	1369	1369
Ncluster	110	110	110	110	110	110	110	110	110	110	110	110
PANEL C - 3rd quartile												
Control Mean	0.403	0.531	0.182	0.309	0.221	0.031	0.375	0.252	0.137	0.053	0.032	0.010
Treatment Effect	0.012	0.013	0.008	0.010	0.003	0.013*	0.000	-0.016	0.022**	0.017**	0.012*	0.006*
(SE)	(0.019)	(0.020)	(0.017)	(0.016)	(0.017)	(0.008)	(0.020)	(0.018)	(0.011)	(0.008)	(0.006)	(0.003)
Nobs	1859	1859	1859	1859	1859	1859	1859	1859	1859	1859	1859	1859
Ncluster	110	110	110	110	110	110	110	110	110	110	110	110
PANEL D - 4th quartile												
Control Mean	0.577	0.507	0.282	0.212	0.296	0.099	0.485	0.241	0.322	0.158	0.106	0.036
Treatment Effect	-0.004	-0.022	0.024	0.006	-0.028**	0.021	-0.016	-0.008	0.007	0.025	0.007	0.011**
(SE)	(0.016)	(0.017)	(0.016)	(0.015)	(0.011)	(0.014)	(0.015)	(0.013)	(0.019)	(0.015)	(0.011)	(0.005)
Nobs	2682	2682	2682	2682	2682	2682	2682	2682	2682	2682	2682	2682
Ncluster	111	111	111	111	111	111	111	111	111	111	111	111

NOTE — In each panel, First line presents the control mean for each variable, second line exhibits linear regression estimates, standard error estimation takes into account cluster at school level. Panel A presents the effects for the students in the 1st quartile in the *Prova Brasil* distribution, while panel B, C and D is restricted to students, respectively in th 2nd, 3rd and 4th quartiles. This baseline distribution is derived by the scores in *Prova Brasil* in 2009 for the RJ, SP and VP areas. * significant at 10% level, ** significant at 5% level and *** significant at 1% level.