# The Short- and Long-Run Impacts of Black Teachers: Evidence from Brazil*

Pedro Lopes†

**Abstract**

We analyze the short- and long-run effects of Black teachers on students' educational and labor-market outcomes in Brazil, tracking two million students up to 12 years. Our empirical strategy compares students within schools assigned to classrooms with varying shares of Black teachers, controlling for prior-year test scores. For Black students, a higher share of Black teachers increases high school graduation, college enrollment, college graduation, and earnings in adulthood. We find no effect for White students. Increasing the share of Black teachers from 0 to 50% would reduce within-school racial gaps in education by 30% and in earnings by 60%. (JEL I21, I24, I26, I28, J15)

Persistent Black-White income inequality is a critical challenge across many countries.[1] These enduring disparities are reinforced by racial gaps in intergenerational mobility, with Black children facing lower chances of moving up and higher risks of falling down the income ladder. While differences in family characteristics, including wealth, parental education,

*We are grateful to Fernanda Estevan, Vitor Possebom, Anjali Adukia, Bladimir Carrilo, Vladimir Ponckzek, Pedro Sant'Anna, Bruna Borges, Filipe Cavalcanti, Ursula Mello, and to seminar participants at the Sao Paulo School of Economics, FGV-CLEAR, the Summer School of Socioeconomic Opportunity and Inequality at the Institute of Applied Economic Research, and the REP LAC Conference 2026 for valuable comments and discussions. We gratefully acknowledge funding from the Sao Paulo Research Foundation (FAPESP – grant 2023/15016-1) and CAPES

†São Paulo School of Economics - FGV. E-mail: pedro.lopes.silva@fgv.br.

[1]For evidence from the United States, see Derenoncourt et al. (2023); for the United Kingdom, see Karagiannaki (2023); for France, see Govind et al. (2025); and for Latin America, see Telles et al. (2025).

family structure, and residential location, account for part of these mobility gaps, substantial disparities remain even among children from comparable backgrounds (Chetty et al., 2020; Britto et al., 2022). This persistence suggests that contextual factors within schools and communities may differentially shape opportunities and aspirations by race, thereby contributing to long-run inequalities in economic outcomes.

Among school-based factors, the racial composition of the teaching workforce may be particularly influential, as it determines the extent to which students are exposed to teachers who share their racial identity. Same-race teachers may serve as role models, raising students' aspirations and motivation to invest in education, provide mentoring and support that foster achievement, or exhibit less negative bias toward minority students. While studies find that teacher-student racial matching improves short-run educational outcomes for Black students, relatively little is known about its long-run impacts, particularly on labor-market outcomes in adulthood. This literature gap matters because, a priori, it is not clear whether these educational gains translate into labor market improvements, as effects may fade out over time, and economic returns depend on the specific educational pathways students pursue.

This paper examines the impacts of Black teachers on Black students' educational and labor-market outcomes using nationwide administrative data from Brazil.[2] We also examine impacts on White students, as these effects inform the potential implications of policies that shift the racial composition of the teacher workforce.

Our setting is particularly compelling for studying the long-run effects of teacher-student racial matching. Brazil has the largest Afro-descendant population outside Africa (about 56% of its citizens) and faces stark racial inequality, as Black individuals are 53% less likely to complete college and earn 41% less income than White individuals.[3] Crucially, Brazil also offers rich, linked administrative data covering the universe of students in basic education, college enrollees, and formal-sector workers. This enables us to track two cohorts of

[2]Throughout the paper, we consider as Black those individuals identified as either *preto* ("black") or *pardo* ("mixed-race"), as is standard in Brazilian official statistics and academic literature (Coelho et al., 2024).

[3]According to the 2022 Population Census, 112.8 million out of 203.1 million Brazilians are of Afro-descendant origin. According to the 2023 National Household Sample Survey (*Pesquisa Nacional por Amostra de Domicílios Contínua*), higher education attainment (individuals 25+) is 13.0% for Black individuals vs. 27.7% for White individuals. Average real monthly income from all jobs (individuals 14+ with positive earnings) is 2,264 BRL for Black individuals vs. 3,847 BRL for White individuals.

approximately one million students each— one beginning in 6$^{th}$ grade (middle school) and the other in 10$^{th}$ grade (high school)—over a 12-year period, following them through higher education and into the labor market. We estimate effects on school progression, test scores, college enrollment and completion, employment, and earnings.

The challenge in estimating the causal effect of Black teachers is that student assignment to classrooms might not be random. Our empirical strategy compares students within the same school who are assigned to classrooms with varying shares of Black teachers, while controlling for prior-year test scores and a rich set of student-, peer-, and teacher-level covariates. Our identification assumption is that, conditional on observables, assignment to Black teachers is as good as random. Although we cannot rule out selection on unobservables, evidence from Chetty et al. (2014a) lends credibility to our strategy, as it demonstrates that conditioning on prior test scores is the key to obtaining unbiased teacher value-added estimates. Moreover, we show that, within schools, predetermined student characteristics do not jointly predict assignment to Black teachers, further supporting our approach.

We find that a higher share of Black teachers in the classroom in a given year improves both short- and long-run outcomes for Black students, with no effect on their White peers. For students in the 10$^{th}$-grade cohort, increasing the share of Black teachers from zero to 50 percent raises the probability of graduating high school on time by 1.18 percentage points, corresponding to a 1.9% increase relative to the mean for Black students.[4] We also observe gains in higher education, as on-time college enrollment rises by 0.44 percentage points (3.9%) and college graduation by age 25 increases by 0.53 percentage points (5.2%), with gains in postsecondary education concentrated in high-quality private institutions. These effects are economically meaningful, reducing the within-school racial disparities in educational attainment, adjusted for covariates, by at least 37.3%.[5] Moreover, the educational improvements translate into labor market advantages by inducing Black students to enter jobs that are more intensive in human capital. Overall, this increased exposure raises average

[4]Effects are reported as a zero-to-50-percent contrast to quantify the total contribution of Black teachers to student achievement when their representation approximates their population share. Nonetheless, Black teachers are underrepresented relative to Black students. According to the School Census, among individuals with observed race, Black teachers constituted 38.5% of the K-12 teaching workforce in 2012, the year in which teacher-student matches were observed for this analysis. By 2024, this share had risen to 44.4%.

[5]This value corresponds to the lowest gap reduction among high school completion, college enrollment, and college graduation.

monthly earnings at age 27 by 32.77 BRL (2.3%) and closes the within-school racial earnings gap, adjusted for covariates, by 60.9%.

Our analysis of the 6$^{\text{th}}$-grade cohort confirms these patterns and enables us to examine learning outcomes using a compulsory national standardized exam. For these students, an increase in the share of Black teachers from zero to 50 percent raises 9$^{\text{th}}$ grade Portuguese test scores by 0.015 standard deviations, on-time high school graduation by 0.64 percentage points (1.4%), and on-time college enrollment by 0.42 percentage points (4.2%). These educational gains reduce the within-school racial gap in long-term educational attainment outcomes, adjusted for covariates, by at least 30.3%. As before, postsecondary effects concentrate in high-quality private institutions.

These average effects, however, mask substantial heterogeneity across student characteristics and school contexts. At the individual level, the positive gains in educational attainment and earnings are driven by Black boys across both cohorts. The results also vary by baseline test scores, with gains in earnings being concentrated among lower-achieving Black students. In contrast, gains in college enrollment and graduation are stronger for their higher-achieving peers, who are arguably closer to the college-going margin. At the school level, the effects are larger in low-socioeconomic-status schools, particularly for the 6$^{\text{th}}$-grade cohort.

To shed light on potential mechanisms, we explore whether the observed effects operate through direct instructional quality within specific subjects or through broader motivational and mentoring pathways. We conduct a cross-effects analysis, restricting the sample to classrooms where the subject teacher (in Math or Portuguese) is White, and examine whether exposure to Black teachers in other subjects still affects test scores in these subjects. We find that Black students exposed to more Black teachers in other subjects perform relatively better on Math and Portuguese tests compared to their White peers, even when the respective subject teacher is White. This cross-subject spillover suggests that mechanisms such as role modeling that enhance students' motivation, or mentoring and support that foster achievement, rather than subject-specific instructional quality, play an important role in explaining our findings.

Our research builds on a large literature on teacher effects, which documents meaningful impacts on a wide range of student outcomes, including test scores, non-cognitive skills,

and long-run earnings (Chetty et al., 2014a,b; Jackson, 2018). A central challenge in this literature, however, is identifying a priori the characteristics of effective teachers, which has led many studies to focus on observable attributes such as race. In the United States, research finds that same-race teachers have positive educational effects for Black students. Several studies document short-term effects, showing gains in test scores, course-passing rates, and student behavior (Dee, 2004; Fairlie et al., 2014; Egalite et al., 2015; Lindsay and Hart, 2017; Yarnell and Bohrnstedt, 2018; Harbatkin, 2021). In contrast, few studies examine whether effects persist many years after exposure to same-race teachers. Gershenson et al. (2022) leverages the random assignment of students to teachers in the Tennessee STAR experiment to show that exposure to same-race teachers in grades K-3 increases high school completion and two-year college enrollment for Black students. Delhommer (2022) similarly finds positive effects on college enrollment using quasi-experimental methods and administrative data from Texas, examining the effects of same-race teachers in $9^{\text{th}}$ grade.

We contribute to the literature in three ways. First, we provide novel evidence on the long-term consequences of student-teacher racial matching, documenting effects on college graduation and, most importantly, on labor-market outcomes. While the educational gains documented in prior work are valuable, their ultimate importance rests on the assumption that they translate into long-run economic success. This link, however, is not guaranteed. For instance, it is unclear whether students who enroll in college actually complete their degrees, given high dropout rates in postsecondary institutions.[6] Moreover, even among those who complete, economic returns to postsecondary education vary widely by field of study, degree type, and institution quality, with returns ranging from positive to negative (Andrews et al., 2016, 2024). We show that Black students' exposure to Black teachers increases both college enrollment and completion at high-quality institutions, and that these gains translate into higher earnings in early adulthood. Our findings suggest that increasing Black teacher representation could yield efficiency gains by mitigating the misallocation of talent resulting from student underinvestment in human capital.

[6]This is a widespread pattern. Among students who enroll in postsecondary education, only 55% in the United States (2014-2015 cohort) and 39% in Brazil (2015 cohort) complete their initial program of study within eight years; these rates do not account for students who transfer and complete degrees in different programs. Sources: National Center for Education Statistics (U.S.) and Higher Education Census (Brazil).

Second, we use nationwide administrative data to evaluate multiple outcomes across different school stages within a single, large-scale framework. This design offers advantages over prior work, as national scope enables broader generalization beyond the state- or district-specific settings common in prior research and provides greater statistical power for inference. In addition, applying a consistent empirical strategy across multiple outcomes enables direct comparison of effect magnitudes, revealing whether effects persist, fade, or amplify as students progress through the educational system and into adulthood.

Third, our results add to the growing body of evidence on teacher effects in developing countries (Araujo et al., 2016; Bau and Das, 2020), particularly regarding the characteristics associated with their effectiveness. Most research on the determinants of teacher effectiveness has focused on the United States. However, understanding what makes teachers effective may be equally important in developing countries. We provide evidence from Brazil showing that racial representation among teachers contributes to student outcomes, indicating that teacher–student racial matching is an important dimension of teacher effectiveness. Our findings can inform policies in similar contexts, particularly in Latin America and the Caribbean. This region is home to over 153 million Afro-descendants, and Black-White disparities in education and income are substantial (FAO and ECLAC, 2025; Telles et al., 2025).[7] Yet, Black teachers often remain underrepresented in the teaching workforce (Elacqua et al., 2024), positioning increased racial representation among teachers as a potential pathway for reducing racial disparities in both educational attainment and income levels.[8]

The rest of this paper is organized as follows. Section I. describes the institutional setting of Brazilian education. Section II. details the data and research design. Section III. presents the sample and summary statistics. Section IV. outlines our empirical strategy, which includes the econometric specification, identification assumptions, and validation tests. Section V. presents the main results. Section VI. provides heterogeneity analyses, Section VII. discusses potential mechanisms, and Section VIII. contains robustness checks. Section IX. concludes.

[7] Afro-descendants account for roughly 24% of the region's total population.

[8] Elacqua et al. (2024) analyzes the ethnic-racial composition of teachers in five countries, comparing it to the racial composition of the overall population from the National Censuses. Afro-descendant teachers are underrepresented in Brazil, Ecuador, Uruguay, and Peru, while they are overrepresented in Guatemala.

# I. Institutional Setting of Brazilian Education

In this section, we provide an overview of the Brazilian education system, which consists of basic education, equivalent to K-12 in the United States, and higher education, offered by universities and other postsecondary institutions.

## A. Basic Education

Brazil's basic education system mandates compulsory schooling for children aged 4 to 17. As of 2012, the year our analysis begins, this compulsory structure consists of four stages: preschool (ages 4-5), primary education (grades 1-5, ages 6-10), lower secondary education (middle school, grades 6-9, ages 11-14), and upper secondary education (high school, grades 10-12, ages 15-17).[9]

At that time, the curriculum followed a nationally standardized structure with no student choice. All subjects were mandatory. Regular academic schools did not offer elective courses, advanced tracks, or honors options. In primary education, a single generalist teacher typically covered all core subjects (Portuguese language, mathematics, science, history, geography, arts, and physical education). Lower secondary education introduced subject-specific teachers and English instruction. Upper secondary education expanded the mandatory curriculum to include biology, physics, chemistry, philosophy, and sociology. At all school levels, students were assigned to fixed classroom groups covering all subjects throughout the year. From lower secondary onward, within a given class, all students were taught by the same set of teachers, who rotated into the classroom to teach their specialized courses. Differences in teacher composition arose across different classes within the same grade and school.[10]

Administratively, the system is split between the public and private sectors. The public system, which enrolled 84.4% of students during our analysis period, operates on a federative model. In this model, the government divides responsibilities by education stage. Municipal

[9] According to the 2010 Population Census, the last available census data before our analysis period, enrollment was not universal, particularly among older students. In 2010, 3.3% of children aged 6 to 14 were not attending school, while 16.7% of those aged 15 to 17 were out of school.

[10] This structure is notably different from that in the United States, where schools generally provide a range of elective options and various academic tracks. Additionally, in the U.S., students have personalized class schedules, move between distinct classrooms throughout the day, and engage with different peer groups in each course.

governments hold primary responsibility for the initial levels and manage the majority of pre-primary and primary education. State governments, in contrast, manage the largest share of enrollments in lower and upper secondary schools. The federal government's role remains primarily regulatory and financial; its Ministry of Education sets national policies, while its direct management of regular schools is minimal. The private, tuition-based sector accounts for the remaining 15.6% of total enrollments. This sector primarily serves higher-income families and is perceived as offering, on average, a higher quality of education than the public sector.

## B. Higher Education

After completing upper secondary education, students may pursue higher education at the undergraduate level, which includes three types of degrees. The Bachelor's degree provides broad academic or professional training; the Teacher degree prepares graduates to teach in basic education, combining subject-specific and pedagogical studies; and the Technological degree offers shorter, career-oriented programs focused on applied professional skills.

Similar to basic education, the higher education system is split between public and private sectors. Over three-quarters of students attend private institutions. Public institutions are divided among federal, state, and municipal systems, with federal institutions enrolling the largest share of public sector students, state institutions maintaining a moderate share, and municipal institutions representing a small fraction of the total. Federal and state institutions do not charge tuition and often receive higher ratings in official government rankings that assess the quality of higher education institutions than their private counterparts, which charge tuition. This combination of higher quality and free access makes public universities more selective, resulting in a smaller share of incoming students who attended public basic education in comparison to private institutions.[11]

Admission processes for higher education vary across institutions. The gateway for federal universities is the ENEM (*Exame Nacional do Ensino Médio*), a national standardized exam administered by the Ministry of Education and available once a year. Students use their

[11] According to the 2018 Higher Education Census, about 64% of students entering public institutions and 78% of students entering private institutions had attended public high school.

exam scores to compete for vacancies at multiple universities nationwide under a centralized admission system. Although some state, municipal, and private institutions also use the ENEM as part of their selection process, most use their own *Vestibular* exams, which are institution-specific and serve as a competitive selection process. Furthermore, ENEM scores are also used to determine student eligibility for federal scholarships or financing at private institutions.

# II. Data and Research Design

This section outlines the data and research design used in our analysis. We begin by introducing the administrative datasets we use and describing the key information they contain. Next, we detail our research design—a longitudinal strategy that uses the data to follow students through three key stages: baseline, treatment, and outcomes—specifying the arrangement of datasets by calendar year. We conclude by describing the linkage procedure used to construct our longitudinal panels.

## A. Data Description

This paper relies on five main administrative data sources: the School Census (*Censo Escolar*), the Basic Education Assessment System (SAEB, *Sistema de Avaliação da Educação Básica*), the National High School Exam (ENEM, *Exame Nacional do Ensino Médio*), the Higher Education Census (*Censo da Educação Superior*), and the Annual Social Information Report (RAIS, *Relação Anual de Informações Sociais*). These datasets are accessed through the Secure Data Access Room (*Sala de Acesso a Dados Protegidos*) of INEP, the National Institute for Educational Studies and Research, which is the federal agency responsible for official education statistics and large-scale assessments in Brazil. In this secure environment, the datasets are available in identified form, allowing us to link records across different databases at the individual level and to follow students from early schooling through higher education and into the labor market.

The School Census provides information on the universe of students enrolled in the school system and their teachers. From this source, we obtain classroom-level allocations of stu-

dents and teachers, as well as rich covariates for both groups. For students, we observe demographic characteristics such as race, gender, date of birth, and schooling level. For teachers, we observe demographic and professional characteristics, including race, gender, age, academic qualifications (undergraduate, specialization, Master's, and Ph.D.), and employment characteristics such as statutory contracts and multiple school assignments.[12] Additionally, the School Census records academic flow indicators (e.g., end-of-year promotion/retention, dropout, and transfers), which we use to determine whether the student was approved at the end of the school year.

The SAEB is Brazil's nationwide standardized testing system, conducted biennially to monitor student achievement and educational quality. It applies exams in Mathematics and Portuguese to students in the $2^{nd}$, $5^{th}$, $9^{th}$, and $12^{th}$ grades. In public schools, the assessment is census-based in the $5^{th}$ and $9^{th}$ grades, while in the $2^{nd}$ and $12^{th}$ grades it is sample-based.[13] In private schools, participation is always sample-based. From SAEB, we obtain baseline test scores in Mathematics and Portuguese, and, when available, subsequent scores to measure medium-run performance. In addition, SAEB provides background variables such as parental education (high school or college completion), family structure (living with a male or female guardian), parental involvement (attendance at school meetings), and household assets (car and computer ownership).

The ENEM is a national high-stakes exam that serves as the main gateway to higher education. Participation is voluntary, and the primary test-takers are high school seniors and recent graduates seeking university admission. The exam is composed of five components: Mathematics, Portuguese Language, Natural Sciences, Human Sciences, and a Writing essay. It plays a central role in college admissions, as public universities rely heavily on ENEM scores in their selection process. In our analysis, we usually summarize student performance by computing the average across the five components, providing a single measure of overall achievement.

The Higher Education Census is a nationwide administrative census that covers all public

[12] A challenge in using the School Census is that a portion of the sample lacks race information for both students and teachers. We address this by imputing the modal race for each individual, identified across multiple census years and other linked data sources. We discuss this procedure in detail in Appendix O.1.

[13] The assessment for $12^{th}$ grade became census-based only from 2017 onwards. In the years we use SAEB data in this study, the assessment was sample-based.

and private higher education institutions and students in Brazil. It provides comprehensive information on student enrollment and institutional characteristics. From this dataset, we collect whether the student is enrolled, the type of institution (public or private), the modality of instruction (in-person or distance learning), the type of program (bachelor's or technical degree), whether the student receives financial aid, and whether they have completed the program. To measure institutional quality, we use the General Index of Courses (IGC, *Índice Geral de Cursos*), an official composite index produced annually by the Ministry of Education. The IGC combines data on undergraduate course evaluations and graduate program performance to create a summary score. It uses a scale ranging from 1 to 5, where higher scores indicate greater quality. We define a high-quality institution as one with a score of 4 or higher.

Finally, we use the RAIS, a matched employer–employee dataset that covers the universe of formal employment relationships in Brazil. It is an administrative registry collected annually by the Ministry of Labor, based on mandatory reports filed by all registered firms. The RAIS provides longitudinal information on every formal job contract in the country, including worker demographics and detailed labor market outcomes. At the worker–year level, we observe race, educational level, contract wages, average monthly earnings (which include additional benefits such as bonuses and allowances), and the number of months worked in each calendar year. The data also include workers' educational levels, occupations from the Brazilian Occupational Classification (CBO), and industries from the National Classification of Economic Activities (CNAE). These data allow us to track whether former students are employed in the formal sector and how much they earn. To better characterize the quality of job placements, we also construct a measure of a job's educational intensity. We build this measure by combining the industry and occupation codes with the average education of workers within each corresponding industry-occupation cell.

To make monetary values comparable over time, we deflate all wage measures using the national consumer price index (IPCA, *Índice Nacional de Preços ao Consumidor Amplo*) and express them in constant 2023 Brazilian reais.

## B. Research Design and Data Linkage

Our research design leverages the rich administrative data described above to construct two longitudinal panels, allowing us to follow students over more than a decade. We focus on two distinct cohorts who entered critical transition years—$10^{th}$ and $6^{th}$ grade—in 2012. For each cohort, our strategy is organized into three key stages: establishing a pre-treatment baseline, measuring students' exposure to Black teachers during a treatment year, and tracking a sequence of short- and long-term outcomes.

For the $10^{th}$ grade cohort, we establish a pre-treatment baseline using the 2011 SAEB, which captures students' $9^{th}$ grade test scores and socioeconomic background. The treatment year is 2012, when the School Census records students' $10^{th}$ grade classroom allocation—our key exposure variable. We then track this cohort through multiple outcome stages: in 2014, we measure high school completion using the School Census and college entrance exam performance using ENEM; from 2015 to 2021, we follow students' college enrollment and graduation through the Higher Education Census; and finally, in 2022–2023, we capture their early-career labor market outcomes using employment and earnings data from RAIS.

For the $6^{th}$ grade cohort, we establish the baseline in 2011 using SAEB data from students' $5^{th}$ grade at the end of elementary school. Treatment exposure occurs in 2012, when the School Census records $6^{th}$ grade classroom assignments. We then track a comprehensive set of outcomes across three educational stages. First, in 2015, we measure middle school outcomes using both the SAEB ($9^{th}$ grade test scores) and the School Census (on-time middle school completion). Second, in 2018, we capture high school outcomes through the School Census (on-time high school graduation) and ENEM (college entrance exam performance). Finally, from 2019 to 2021, we follow students' postsecondary transitions using the Higher Education Census.[14] See Figures O.1 and O.2 in the Online Appendix for a visual representation of the longitudinal data structure for the $10^{th}$- and $6^{th}$-grade cohorts.

We construct these two longitudinal panels by linking records across databases. A link is possible only when an individual has a common unique identifier and a record in both

[14] We do not include the most recently available years of the Higher Education Census (2022–2023) in our analyses. For these years, the linking variables necessary to merge these data with the School Census (a procedure we discuss below) are missing for a large portion of the Higher Education Census sample, precluding their use.

datasets. Our strategy relies on two unique, time-invariant identifiers: the INEP student ID and the National Individual Tax Registration Number (CPF, *Cadastro de Pessoa Física*). The School Census is our central database because it contains both identifiers, while the SAEB uses only the INEP ID, and the post-secondary outcome datasets—ENEM, the Higher Education Census, and RAIS—use only the CPF. The INEP ID is universally available across the in-school datasets (School Census and SAEB), and the CPF is likewise universal within the post-secondary datasets.

With these identifiers, we perform a two-step linkage. First, we use the INEP ID to create an exact match between the School Census and SAEB records. Second, we use the CPF to link students from the School Census to their records in the post-secondary outcome databases. The key linkage challenge is that the CPF is not universally available in the School Census, where a portion of our sample lacks this identifier. For this subset of students, we code any individual with a missing CPF as not having a record in the post-secondary outcome datasets (i.e., not having taken the ENEM, not having enrolled in higher education, and not having entered the formal labor market). We discuss this methodological choice in detail in the Online Appendix O.3 and demonstrate throughout the paper that our results are not affected by it.

# III. Sample and Summary Statistics

This section describes the construction of the analytical sample and reports key descriptive statistics. We first explain how we build the final sample of students from the raw administrative data, detailing the inclusion criteria and restrictions applied. We then present baseline characteristics of students by race and compare their academic and labor market outcomes. Finally, we summarize teacher characteristics by race and document Black teachers' underrepresentation.

## A. Sample

The final sample for our main analysis consists of 1,018,697 students in $6^{\text{th}}$ grade and 965,664 in $10^{\text{th}}$ grade. We construct this sample by applying a series of restrictions to an

initial dataset of all 6th and 10th grade students in Brazil with available baseline scores from the previous SAEB assessment. These restrictions, which parallel those in prior work, are detailed below.

First, we focus on students in classrooms that cover a core set of mandatory academic disciplines. For 6th grade, these subjects are Mathematics, Portuguese, History, Geography, and Sciences. For 10th grade, we require classroom coverage of at least eight of the nine following disciplines: Mathematics, Portuguese, History, Geography, Physics, Chemistry, Biology, Philosophy, and Sociology. We then impose plausible teacher-clasroom ratios to avoid potential data mislinking.[15] Next, we exclude students if the race of any of their classroom teachers is not reported. This restriction prevents measurement error in the treatment variable, which could bias our results. We then drop students with missing race, followed by observations missing other demographic or socioeconomic covariates, and retain only classrooms with five to fifty observed students. Because the analysis focuses on the Black-White achievement gap, we exclude Indigenous students.[16] In Table 1, we show the impact of each of these restrictions on the sample size.

Table 1: Sample construction for 10th and 6th grade

| | 10th Grade | | 6th Grade | |
|---|---|---|---|---|
| Sample | N | % | N | % |
| Public school students who took SAEB in the previous year | 1609137 | 100.00 | 1910804 | 100.00 |
| Minimum subject coverage ensured | 1489322 | 92.55 | 1835449 | 96.06 |
| Plausible teacher–subject ratios | 1485304 | 92.30 | 1804236 | 94.42 |
| No missing teacher race | 1109626 | 68.96 | 1347109 | 70.50 |
| No missing student race | 1030444 | 64.04 | 1167563 | 61.10 |
| No missing other covariates | 986915 | 61.33 | 1043939 | 54.63 |
| At least 5 students with full information per classroom | 974642 | 60.57 | 1028614 | 53.83 |
| Excluding Indigenous students | 965664 | 60.01 | 1018697 | 53.31 |

[15] We restrict the number of teachers per classroom to between 3-7 in 6th grade and 4-14 in 10th grade, which is approximately ±2 and ±5 around the expected number of core subjects, respectively. These bounds reflect common classroom structures, reducing the likelihood of student-teacher mislinking and excluding cases that likely reflect high teacher turnover or incomplete staffing records.

[16] There are five official racial categories from the Brazilian Institute of Geography and Statistics: *branco* (White), *preto* (Black), *pardo* (Mixed-Race), *amarelo* (Asian), and *indígena* (Indigenous). Consistent with our definition of Black (individuals identified as *preto* or *pardo*), we group individuals identified as *amarelo* (Asian) with White individuals. This group comprises approximately 1% of our final sample. We adopt this classification because it aligns with the structure of most affirmative action policies in Brazil, which typically target the PPI (*preto, pardo, indígenas*) group separately from White and Asian individuals. As our analysis focuses on the disparities these policies address, this grouping is the most relevant.

## B. Summary Statistics

Table 2 reports baseline means by race for the $10^{th}$ and $6^{th}$ grades, along with p-values from within-grade two-sided t-tests of mean differences by race.[17] We find large and statistically significant achievement gaps between Black and White students. In $10^{th}$ grade, Black students score $0.30\sigma$ lower in Math and $0.26\sigma$ lower in Portuguese than White students. In $6^{th}$ grade, these respective gaps are $0.36\sigma$ and $0.30\sigma$.[18] The results show other systematic differences. Black students are slightly older, their parents have less schooling, and their households hold fewer assets (car, computer). Exposure to Black teachers also differs, as Black students are taught by them more often than their White peers (38% versus 20% in $10^{th}$ grade; 45% versus 22% in $6^{th}$ grade). In contrast, the gender composition, family structure, and attendance of guardians at school meetings are very similar across groups. In both cohorts, Black students represent a slim majority of the sample: 505,289 of 965,664 students in $10^{th}$ grade and 543,174 of 1,018,697 students in $6^{th}$ grade.

Table 2: Baseline Means by Student Race

| | $10^{th}$ Grade | | | $6^{th}$ Grade | | |
|---|---|---|---|---|---|---|
| Variable | Black | White | p-value (Diff.) | Black | White | p-value (Diff.) |
| Male | 0.46 | 0.47 | < 0.001 | 0.50 | 0.49 | < 0.001 |
| Age | 15.75 | 15.51 | < 0.001 | 11.87 | 11.58 | < 0.001 |
| Math z-score 2011 | -0.14 | 0.16 | < 0.001 | -0.17 | 0.19 | < 0.001 |
| Portuguese z-score 2011 | -0.12 | 0.14 | < 0.001 | -0.14 | 0.16 | < 0.001 |
| Female guardian has college degree | 0.06 | 0.08 | < 0.001 | 0.10 | 0.12 | < 0.001 |
| Male guardian has college degree | 0.03 | 0.05 | < 0.001 | 0.08 | 0.10 | < 0.001 |
| Female guardian has high school | 0.21 | 0.24 | < 0.001 | 0.12 | 0.15 | < 0.001 |
| Male guardian has high school | 0.13 | 0.18 | < 0.001 | 0.08 | 0.10 | < 0.001 |
| Lives with male guardian | 0.76 | 0.79 | < 0.001 | 0.82 | 0.83 | < 0.001 |
| Lives with female guardian | 0.95 | 0.96 | < 0.001 | 0.97 | 0.97 | < 0.001 |
| Guardians attend school meetings | 0.94 | 0.93 | < 0.001 | 0.91 | 0.93 | < 0.001 |
| Household asset: car | 0.40 | 0.62 | < 0.001 | 0.39 | 0.61 | < 0.001 |
| Household asset: computer | 0.56 | 0.74 | < 0.001 | 0.47 | 0.65 | < 0.001 |
| Share Black Teachers | 0.38 | 0.20 | < 0.001 | 0.45 | 0.22 | < 0.001 |
| Observations | 505289 | 460375 | | 543174 | 475523 | |

Notes: The table reports mean values of students' characteristics. Reported p-values are from two-sided t-tests of mean differences between White and Black students. All variables are binary (0/1) except $ShareBlackTeachers$, $Age$, and baseline test scores.

[17] Online Appendix Table O.4 reports these means disaggregated for all five official racial categories from the Brazilian Institute of Geography and Statistics.

[18] Online Appendix Figure O.4 shows baseline z-score densities by race; in every grade–subject pair, the Black distribution is shifted left relative to the White distribution.

Table 3 reports the means of students by race across several key academic and labor market outcomes. For the $10^{th}$ grade cohort (Panel A), there is a three percentage point gap in grade approval in the treatment year. This disparity rises to a seven percentage point gap in on-time high school graduation. Black students are also 3 percentage points less likely to participate in the ENEM, and among test-takers, there is an achievement gap of $0.33\sigma$. These differences persist through higher education, with a six percentage point gap in both on-time college enrollment and college graduation seven years after expected high school completion. These attainment gaps are mirrored in early-career labor outcomes. Black students are three percentage points less likely to be formally employed, have average monthly earnings that are 320 BRL (22%) lower, and among those employed, face a monthly wage gap of 405 BRL (14%). This pay gap is associated with a disparity in job quality. Black students are more likely to work in industry-occupations where a smaller share of the workforce has some college education (24% vs. 30%).

The pattern of widening disparity is even more pronounced for the $6^{th}$ grade cohort (Panel B). An initial five percentage point gap in grade approval more than doubles to a substantial 12 percentage point gap in on-time high school graduation. On the $9^{th}$ grade SAEB exam, Black students are 9 percentage points less likely to participate and score $0.28\sigma$ lower in Math and $0.22\sigma$ lower in Portuguese. The gap in on-time high school graduation is a substantial 13 percentage points. Consistent with the findings for the $10^{th}$ grade cohort, Black students are 9 percentage points less likely to take the ENEM exam, exhibit an achievement gap of 0.31 standard deviations among test-takers, and are 7 percentage points less likely to enroll in college on time.

Table 4 presents the baseline characteristics of teachers in our sample, revealing significant observable differences between Black and White teachers.[19] On average, Black teachers are more likely to be male, slightly younger, and have lower levels of formal education. However, the most significant finding is the underrepresentation of Black teachers relative to the student body. While Black students form slightly above 50% of the student sample, Black teachers constitute a minority of the teacher sample, at 31% in $10^{th}$ grade (49,915 of 162,457

[19]Online Appendix Table O.5 reports these teachers' statistics disaggregated for all five official racial categories from the Brazilian Institute of Geography and Statistics.

Table 3: Outcome Means by Student Race

| Variable | Black | White | p-value (Diff.) |
|---|---|---|---|
| *Panel A: $10^{th}$ Grade* | | | |
| Grade approval 2012 | 0.79 | 0.82 | $< 0.001$ |
| High school graduation 2014 | 0.61 | 0.68 | $< 0.001$ |
| ENEM participation 2014 | 0.39 | 0.42 | $< 0.001$ |
| ENEM z-score 2014 | -0.16 | 0.17 | $< 0.001$ |
| College enrollment 2015 | 0.11 | 0.17 | $< 0.001$ |
| College graduation by 2021 | 0.11 | 0.17 | $< 0.001$ |
| Formal employment 2022–2023 | 0.65 | 0.68 | $< 0.001$ |
| Avg monthly earnings (BRL) 2022–2023 | 1430.59 | 1750.57 | $< 0.001$ |
| Avg monthly wage (BRL) 2022–2023 | 2827.97 | 3232.98 | $< 0.001$ |
| % Some college (industry–occupation) 2022–2023 | 0.24 | 0.30 | $< 0.001$ |
| *Panel B: $6^{th}$ Grade* | | | |
| Grade approval 2012 | 0.85 | 0.90 | $< 0.001$ |
| SAEB participation 2015 | 0.56 | 0.65 | $< 0.001$ |
| Math z-score 2015 | -0.14 | 0.14 | $< 0.001$ |
| Portuguese z-score 2015 | -0.11 | 0.11 | $< 0.001$ |
| Middle school graduation 2015 | 0.63 | 0.75 | $< 0.001$ |
| High school graduation 2018 | 0.45 | 0.58 | $< 0.001$ |
| ENEM participation 2018 | 0.30 | 0.39 | $< 0.001$ |
| ENEM z-score 2018 | -0.16 | 0.15 | $< 0.001$ |
| College enrollment 2019 | 0.10 | 0.17 | $< 0.001$ |

Notes: The table reports mean values of students' characteristics. Reported p-values are from two-sided t-tests of mean differences between White and Black students.

teachers) and 37% in 6$^{\text{th}}$ grade (56,212 of 150,721 teachers). This imbalance is a widespread phenomenon within schools, as shown in Appendix Figures O.5 and O.6. These figures show the joint distribution of the share of Black teachers and Black students across schools. They illustrate a clear pattern where the proportion of Black teachers is systematically lower than that of Black students. This underrepresentation at the school-grade level is observed in nearly 81% of schools in the 10$^{\text{th}}$ grade sample and in over 73% of schools in the 6$^{\text{th}}$ grade sample.

# IV. Empirical Strategy

This section describes the empirical framework we use to estimate the effects of exposure to Black teachers on student outcomes. We first present the econometric specification and define the key parameters of interest. We then discuss how our setting makes a causal interpretation of the estimates plausible and state the identification hypothesis. Finally, we describe the validation test used to assess whether this hypothesis is consistent with the

Table 4: Baseline Means by Teacher Race

| | 10h Grade | | | 6h Grade | | |
|---|---|---|---|---|---|---|
| Variable | Black | White | p-value (Diff.) | Black | White | p-value (Diff.) |
| Male | 0.42 | 0.35 | $< 0.001$ | 0.29 | 0.24 | $< 0.001$ |
| Age | 39.62 | 40.53 | $< 0.001$ | 39.43 | 40.49 | $< 0.001$ |
| Undergraduate degree | 0.93 | 0.95 | $< 0.001$ | 0.87 | 0.94 | $< 0.001$ |
| Specialization degree | 0.31 | 0.33 | $< 0.001$ | 0.31 | 0.38 | $< 0.001$ |
| Master degree | 0.02 | 0.03 | $< 0.001$ | 0.01 | 0.02 | $< 0.001$ |
| Ph.D. degree | 0.00 | 0.00 | $< 0.001$ | 0.00 | 0.00 | $< 0.001$ |
| Any statutory work regime | 0.70 | 0.70 | 0.130 | 0.73 | 0.74 | $< 0.001$ |
| Work in multiple schools | 0.45 | 0.45 | 0.315 | 0.43 | 0.45 | $< 0.001$ |
| Observations | 49915 | 112542 | | 56212 | 94509 | |

Notes: The table reports mean values of teachers' characteristics. Reported p-values are from two-sided t-tests of mean differences between White and Black students. All variables are binary (0/1) except *Age*.

data.

## A. Econometric Specification

We conduct the analysis separately by grade. Our natural experiment compares students in the same school who differ in their exposure to Black teachers. The identifying variation arises from schools with multiple teachers of different races, so that students within the same cohort are assigned to classrooms with varying shares of Black teachers. For each grade, we estimate:

$$\begin{aligned} Y_{ics} = \beta_0 &+ \beta_1 \,\text{ShareBlackTeachers}_{cs} + \beta_2 \,\text{Black}_{ics} \\ &+ \beta_3 \,\text{ShareBlackTeachers}_{cs} \times \text{Black}_{ics} \\ &+ \Phi f(\mathbf{A}_{ics}) + \Omega f(\bar{\mathbf{A}}_{-i,cs}) + \gamma \mathbf{X}_{ics} + \delta \mathbf{W}_{ics} + \alpha_s + \varepsilon_{ics} \end{aligned} \tag{1}$$

where $Y_{ics}$ is the outcome of student $i$ in classroom $c$ and school $s$. $\text{ShareBlackTeachers}_{ics}$ is the fraction of Black teachers in class $c$, a value ranging from 0 to 1.[20] $\text{Black}_{ics}$ is a binary variable equal to one if the student $i$ is Black, and zero otherwise. $\mathbf{A}_{ics}$ denotes the student's own lagged test scores in Mathematics and Portuguese, included as a cubic

[20] In the Brazilian educational system, students are assigned to a fixed classroom covering all grade-specific subjects throughout the year. All students in each classroom $c$ are thus exposed to the same set of teachers. Still, students in distinct classes within the same grade and school face different racial compositions of their teachers.

polynomial, and $\bar{\mathbf{A}}_{-i,cs}$ is the leave-out average of classmates' lagged test scores in the same subjects, also included as a cubic polynomial. $\mathbf{X}_{ics}$ is a vector of other student-level controls and includes indicators for gender and age; parental education (high school and college completion); family structure (living with a male or female guardian); parental involvement (attendance at school meetings); and household assets (ownership of a car and a computer). $\mathbf{W}_{ics}$ is a vector of other classroom-level composition controls: the share of Black peers, and teacher shares by age (30+), statutory work regime, multiple-school assignment, academic qualification (undergraduate, specialization, master's, Ph.D.), and gender. $\alpha_s$ represents school fixed effects. The term of error $\varepsilon_{ics}$ captures unobserved determinants of the outcome variable. Standard errors are clustered at the school level.

Our specification yields several key parameters of interest. First, $\beta_1$ measures the impact of increasing the share of Black teachers on White students; this is policy-relevant as it shows whether a more diverse teacher workforce has any spillover effects on the White student population. Second, the coefficient $\beta_2$ captures the within-school achievement gap between Black and White students, adjusted for covariates, when the share of Black teachers is zero. The coefficient $\beta_3$ is the differential effect for Black students. This parameter measures whether racial matching narrows or widens the racial achievement gap, thereby showing how teacher diversity affects racial inequality in outcomes. Finally, the effect of Black teachers on Black students is given by the sum of the effect on White students with the differential effect for Black students, $\beta_1 + \beta_3$.[21]

## B. Identification

In the ideal setting, students and teachers would be randomly assigned to classrooms within schools. This would ensure that exposure to Black teachers is independent of all other determinants of student outcomes, allowing for clean identification of causal effects. In

[21] Let $B$ be an indicator for a Black student and $S$ be the share of Black teachers. The effect on a Black student of moving from a classroom with no Black teachers ($S = 0$) to a classroom where the share is $s$ is given by:

$$E[Y \mid B{=}1, S{=}s] - E[Y \mid B{=}1, S{=}0] = \ (\beta_0 + s\beta_1 + \beta_2 + s\beta_3) \ - \ (\beta_0 + \beta_2) = \boxed{s(\beta_1 + \beta_3)}.$$

reality, however, assignment might not be random: schools might track students into specific classes, and teachers could be allocated strategically. Such mechanisms would raise concerns about sorting and the possibility of bias if students' treatment assignment correlates with unobserved determinants of student achievement.

Two features of our institutional setting help mitigate these concerns about sorting. First, we focus on public schools, where principals are appointed through bureaucratic or community-based processes, including some tenured civil service positions, rather than hired by school owners or managers responding to market demand. This bureaucratic framework is largely insulated from the parental pressure for specific placements that is common in market-oriented private schools. Second, we examine key transition years: the first year of middle school ($6^{\text{th}}$ grade) and the first year of high school ($10^{\text{th}}$ grade), both characterized by a significant influx of new students.[22] At this stage, information asymmetry is high—administrators lack personal knowledge of students, while new parents and students lack both information about teacher reputations and the social networks needed to influence classroom assignments— weakening channels for strategic tracking and influence-driven placements.

To address any sorting that remains unmitigated by these institutional features, our identification strategy relies on a rich set of controls, including, crucially, baseline test scores. As discussed in Chetty et al. (2014a), conditioning on prior student performance is the key to obtaining unbiased estimates of teacher value-added. Following this approach, we flexibly control for students' baseline achievement by including their own lagged test scores in Math and Portuguese, entered as cubic polynomials (linear, squared, and cubic terms). In addition, we include a set of student, peer, and teacher-level covariates, along with school fixed effects, which mitigate remaining sorting concerns and thereby strengthen our identification strategy. Our identification hypothesis is that, conditional on control variables, student assignment across classrooms with different shares of Black teachers within the same school is as good as random.

[22] This influx occurs because in Brazil's public school system, municipalities generally manage schools that serve grades 1-5, while states manage schools that serve grades 6-12. Additionally, state education systems often separate middle schools (grades 6-9) from high schools (grades 10-12). Thus, students are often required to switch institutions when moving from $5^{\text{th}}$ to $6^{\text{th}}$ grade and from $9^{\text{th}}$ to $10^{\text{th}}$ grade.

## C. Validation Test

We do not take our identification assumption for granted. Instead, we test directly for sorting on predetermined observables. Specifically, we regress the share of Black teachers in a class on student characteristics:

$$ShareBlackTeachers_{cs} = \psi_0 + \psi_1 \mathbf{T}_{ics} + \alpha_s + \varepsilon_{ics}, \tag{2}$$

where $\mathbf{T}_{ics}$ is a vector of student-level covariates and $\alpha_s$ are school fixed effects.

We then conduct a Wald test of whether student characteristics (excluding school fixed effects) jointly predict the share of Black teachers in the classroom. Failure to reject the null indicates that observable characteristics do not predict exposure to Black teachers, supporting the interpretation that, conditional on our controls, student–teacher assignment approximates randomization.[23]

To account for potential heterogeneity, we also allow the allocation mechanism to vary flexibly by race. In practice, we split the sample between Black and White students and re-estimate the regression in Equation 2, allowing the coefficients on school and other controls to differ by group. This ensures that if sorting patterns operate differently across racial groups, we can detect them.

Table 5 reports the student-allocation tests for $10^{\text{th}}$ and $6^{\text{th}}$ grades, using the full sample and subsamples by race (Black and White students). Overall, across grades and samples, the estimated coefficients on student characteristics are very close to zero and, for the vast majority, statistically insignificant. This suggests little to no correlation between observable individual characteristics and the share of Black teachers in students' classrooms after controlling for school fixed effects. Wald tests of joint significance corroborate this interpretation: in all six specifications, p-values are high (0.18–0.97) and exceed conventional thresholds. Taken together, these results support the as-if random assignment assumption within schools.

[23] This test is a standard method for assessing non-random student assignment in the student-teacher matching literature. See Kofoed and mcGovney (2017) and Delhommer (2022) for examples.

Table 5: $10^{th}$ and $6^{th}$ Grade - Sorting Regressions

| | $10^{th}$ Grade | | | $6^{th}$ Grade | | |
|---|---|---|---|---|---|---|
| | All | Black | White | All | Black | White |
| Dependent Var. | Share Black Teachers | Share Black Teachers | Share Black Teachers | Share Black Share | Share Black Teachers | Share Black Teachers |
| Male | 0.0001 | 0.0000 | 0.0003 | 0.0000 | 0.0000 | -0.0001 |
| | (0.0002) | (0.0003) | (0.0002) | (0.0002) | (0.0003) | (0.0002) |
| Age | 0.0001 | 0.0000 | 0.0002 | -0.0001 | -0.0002 | 0.0000 |
| | (0.0002) | (0.0003) | (0.0002) | (0.0002) | (0.0003) | (0.0002) |
| Portuguese z-score 2011 | -0.0024 | -0.0078 | 0.0025 | 0.0024 | 0.0033 | 0.0011 |
| | (0.0036) | (0.0052) | (0.0047) | (0.0031) | (0.0044) | (0.0044) |
| Math z-score 2011 | 0.0021 | 0.0028 | 0.0029 | -0.0005 | -0.0012 | 0.0016 |
| | (0.0037) | (0.0054) | (0.0048) | (0.0036) | (0.0051) | (0.0053) |
| Female guardian has college degree | -0.0005 | 0.0000 | -0.0009** | -0.0001 | -0.0007 | 0.0003 |
| | (0.0003) | (0.0005) | (0.0004) | (0.0003) | (0.0005) | (0.0004) |
| Male guardian has college degree | -0.0002 | -0.0004 | -0.0001 | 0.0002 | 0.0004 | 0.0003 |
| | (0.0004) | (0.0007) | (0.0004) | (0.0003) | (0.0005) | (0.0004) |
| Female guardian has high school | -0.0003 | 0.0003 | -0.0007*** | 0.0002 | 0.0004 | 0.0000 |
| | (0.0002) | (0.0003) | (0.0002) | (0.0003) | (0.0004) | (0.0003) |
| Male guardian has high school | 0.0001 | 0.0001 | 0.0001 | 0.0004 | 0.0003 | 0.0005 |
| | (0.0002) | (0.0004) | (0.0003) | (0.0003) | (0.0005) | (0.0004) |
| Lives with male guardian | 0.0000 | -0.0001 | 0.0001 | -0.0003 | -0.0001 | -0.0005 |
| | (0.0002) | (0.0003) | (0.0003) | (0.0002) | (0.0003) | (0.0003) |
| Lives with female guardian | 0.0002 | 0.0001 | 0.0000 | 0.0001 | -0.0004 | 0.0006 |
| | (0.0004) | (0.0006) | (0.0005) | (0.0005) | (0.0007) | (0.0008) |
| Guardians attend school meetings | 0.0005 | 0.0005 | 0.0006 | -0.0002 | 0.0005 | -0.0012*** |
| | (0.0003) | (0.0005) | (0.0004) | (0.0003) | (0.0005) | (0.0005) |
| Household asset: car | 0.0001 | 0.0002 | 0.0001 | -0.0001 | -0.0001 | 0.0001 |
| | (0.0002) | (0.0003) | (0.0002) | (0.0002) | (0.0003) | (0.0003) |
| Household asset: computer | -0.0001 | -0.0003 | 0.0002 | -0.0001 | -0.0001 | -0.0001 |
| | (0.0002) | (0.0003) | (0.0003) | (0.0002) | (0.0003) | (0.0003) |
| Polynomials z-scores 2011 | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Dependent Mean | 0.2925 | 0.3778 | 0.1988 | 0.3420 | 0.4454 | 0.2238 |
| Wald (joint nullity), p-value | 0.4343 | 0.3700 | 0.1819 | 0.9691 | 0.8882 | 0.4860 |
| Observations | 965,664 | 505,289 | 460,375 | 1,018,697 | 543,174 | 475,523 |

*Notes:* Wald tests jointly whether any student-level coefficient is different from zero. Polynomials z-scores 2011 include individuals' own Math and Portuguese z-scores entered flexibly as a cubic polynomial (squared and cubic terms). Standard errors clustered at the school level are in parentheses. P-values: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

# V. Main Results

This section presents the main results. We begin with the estimates for the $10^{th}$-grade cohort, following the chronological order of outcomes from short- to long-run measures. We then turn to the $6^{th}$ grade cohort and replicate the same sequence of analyses. The treatment variable equals the share of Black teachers in a student's classroom (ranging from 0 to 1). We discuss the coefficients' magnitudes as the effect of a 0.5-unit increase in this share, corresponding to a shift from no representation to a level more consistent with the share of

Black students in the population.

## A. $10^{th}$ Grade

### 1. Educational Outcomes

Table 6 reports estimates of the effects of exposure to Black teachers on $10^{th}$-grade upper secondary education outcomes. Column 1 considers grade approval in the treatment year. For White students, the exposure coefficient is not statistically different from zero. Black students, conditional on controls and fixed effects, are 1.65 p.p. less likely to be approved than White peers. Raising exposure from 0 to 50% reduces this difference by 0.90 p.p. (about 54%), a statistically significant gap-closing effect ($p < 0.01$). However, the effect for Black students is imprecisely estimated and not statistically different from zero ($p = 0.152$).

For the longer-run upper secondary education outcomes—on-time high school graduation, ENEM participation, and ENEM performance (Columns 2–5)—we do not detect effects for White students; the estimates are near zero. By contrast, the effects for Black students are positive, reducing the within-school racial gap by 40–57%, with statistically significant gains. Column 2 shows that increasing exposure from 0 to 50% raises the probability of graduating by 1.18 p.p. ($p = 0.006$). Column 3 presents evidence of greater college intent, with ENEM participation rising by 0.79 p.p. ($p = 0.039$). To estimate the effects on test scores, we account for treatment-induced differential selection into the exam in two ways. In Column 4, we impute the minimum observed score for all non-participants. In this scenario, exposure from 0 to 50% raises the ENEM z-score by 0.0179 SD ($p = 0.018$). In Column 5, we assume all non-participants would score below the median among test takers and define the outcome as an indicator for scoring above that median. Using this conservative approach, the same change in exposure raises the probability by 0.62 p.p. ($p = 0.02$).

We examine college outcomes in Table 7. For White students, effects on college enrollment and graduation are not statistically distinguishable from zero. Consistent with the increase in college intent, Column 1 shows that raising exposure from 0 to 50% increases on-time college enrollment by 0.44 p.p. ($p = 0.082$) for Black students. The impact on educational attainment persists through to completion. We measure graduation seven years

Table 6: $10^{th}$ Grade - Upper Secondary Education Outcomes

| Outcomes | (1) Grade approval 2012 | (2) High school graduation 2014 | (3) ENEM participation 2014 | (4) ENEM z-score (imputation min.) 2014 | (5) ENEM > 50th pct (among participants) 2014 |
|---|---|---|---|---|---|
| Share Black Teachers | -0.0061 | -0.0054 | -0.0003 | 0.0011 | 0.0024 |
| | (0.0084) | (0.0087) | (0.0079) | (0.0156) | (0.0057) |
| Black | -0.0165*** | -0.0254*** | -0.0185*** | -0.0406*** | -0.0141*** |
| | (0.0013) | (0.0015) | (0.0015) | (0.0029) | (0.0011) |
| Share Black Teachers x Black | 0.0180*** | 0.0291*** | 0.0162*** | 0.0346*** | 0.0099*** |
| | (0.0033) | (0.0039) | (0.0040) | (0.0079) | (0.0032) |
| Student-level covariates | Yes | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes |
| Effect (Black) | 0.0119 | 0.0236 | 0.0158 | 0.0358 | 0.0124 |
| p-value | 0.1521 | 0.0055 | 0.0391 | 0.0176 | 0.0221 |
| Mean Black | 0.7871 | 0.6139 | 0.3922 | -0.0488 | 0.1648 |
| Mean White | 0.8179 | 0.6849 | 0.4237 | 0.0536 | 0.2360 |
| SD Black | 0.4094 | 0.4868 | 0.4882 | 0.9695 | 0.3710 |
| SD White | 0.3859 | 0.4646 | 0.4942 | 1.0298 | 0.4246 |
| Observations | 965,664 | 965,664 | 965,664 | 965,664 | 965,664 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments, by academic qualification (undergraduate, specialization, master's, Ph.D.), and male. Standard errors clustered at the school level in parentheses. Share Black Teachers ranges from 0 to 1. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

after expected high school completion to capture students who may not enroll in college immediately. Column 2 shows that exposure to Black teachers increases a Black student's probability of graduating from college by 0.53 p.p. ($p = 0.029$), a 5% increase relative to the mean, closing nearly 40% of the within-school racial gap.

Columns 3 and 4 disaggregate this graduation effect by institution type and show that it is driven by the private sector, which enrolls the majority of students. In Column 3, we estimate a precise null effect on Black students graduating from a public university. In Column 4, we estimate a positive and significant effect for graduation from a private one. While private universities in Brazil are, on average, less selective, the gains are not concentrated in lower-tier institutions. Instead, Column 5 indicates that an increase from 0 to 0.5 in exposure yields a 0.35 p.p. rise in graduation from IGC-measured high-quality colleges, accounting for 67% of the total graduation effect. We show corroborating evidence that Black students shift into higher-value-added pathways in Online Appendix Table 1..

The estimated effects are largest for long-duration degrees ($\geq$ 2,400 hours) and for in-person programs, which are associated with significantly higher value-added than online alternatives (Barahona et al., 2025).[24]

Table 7: 10th Grade - College Outcomes

| Outcomes | (1) College enrollment 2015 | (2) College graduation by 2021 | (3) Public college graduation by 2021 | (4) Private college graduation by 2021 | (5) High quality college graduation by 2021 |
|---|---|---|---|---|---|
| Share Black Teachers | -0.0036 | -0.0082 | 0.0025 | -0.0102** | -0.0008 |
| | (0.0054) | (0.0051) | (0.0023) | (0.0047) | (0.0030) |
| Black | -0.0165*** | -0.0239*** | -0.0018*** | -0.0221*** | -0.0070*** |
| | (0.0011) | (0.0011) | (0.0005) | (0.0010) | (0.0006) |
| Share Black Teachers x Black | 0.0123*** | 0.0187*** | 0.0007 | 0.0179*** | 0.0078*** |
| | (0.0030) | (0.0029) | (0.0014) | (0.0028) | (0.0014) |
| Student-level covariates | Yes | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes |
| Effect (Black) | 0.0087 | 0.0105 | 0.0032 | 0.0077 | 0.0070 |
| p-value | 0.0816 | 0.0290 | 0.1357 | 0.0822 | 0.0137 |
| Mean Black | 0.1125 | 0.1055 | 0.0187 | 0.0871 | 0.0248 |
| Mean White | 0.1749 | 0.1709 | 0.0285 | 0.1429 | 0.0457 |
| SD Black | 0.3160 | 0.3072 | 0.1356 | 0.2820 | 0.1554 |
| SD White | 0.3799 | 0.3764 | 0.1665 | 0.3500 | 0.2089 |
| Observations | 965,664 | 965,664 | 965,664 | 965,664 | 965,664 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments,by academic qualification (undergraduate, specialization, master's, Ph.D.), and male. Standard errors clustered at the school level in parentheses. Share Black Teachers ranges from 0 to 1. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## 2. Labor Market Outcomes

Table 8 examines whether previously observed educational gains translate into improved early-career labor outcomes, measured between 2022 and 2023. Column 1 examines formal employment at the extensive margin. While Black teachers' effects for White students are indistinguishable from zero, we find a significant and negative differential effect for Black students. This pattern is consistent with the treatment encouraging Black students, who

[24]Degrees with at least 2,400 class hours typically correspond to four-year bachelor's degrees in Brazil. The sector, modality, and quality patterns discussed here remain when we focus on on-time enrollment in 2015 rather than graduation by 2021, although estimates are somewhat less precise (see Online Appendix Table O.8).

are, within-school, more likely to enter the labor force early and less likely to get in college, to postpone their workforce entry to invest in further education. Column 2 examines the number of months worked and shows that this delay does not impact it. The share of Black teachers does not have a statistically significant effect for any student group, nor is there a significant differential effect.

Columns 3 and 4 report the effects on average monthly earnings and wages, respectively. Although the wage analysis of the employed subsample requires awareness of the selection effects identified in Column 1, we find remarkably similar patterns for both outcomes. In both cases, we detect no significant effects for White students but find significant positive differential effects for Black students ($p < 0.05$). Moving the share of Black teachers from 0 to 50% has a substantial impact on racial disparities, closing 60.9% of the adjusted within-school earnings gap (Column 3) and 31.5% of the wage gap (Column 4). This translates into positive, though imprecisely measured, total gains for Black students: a 32.77 BRL increase in monthly earnings ($p = 0.128$) and a 56.62 BRL increase in monthly wages ($p = 0.124$).

We next examine whether higher education serves as the mechanism driving labor market gains, a pathway supported by a substantial literature documenting positive returns across diverse contexts (Zimmerman, 2014; Ost et al., 2018; Estevan et al., 2024). To test this, we construct a measure of job skill intensity: for each student's industry-occupation cell, we calculate the share of all workers in that cell who have at least some college education. This measure captures whether students sort into more skill-intensive occupations, that is, jobs typically requiring higher education. Column 5 presents these results. For White students, we find no significant effect. For Black students, however, increasing exposure to Black teachers from 0 to 50% increases the college-educated share in their occupation by 0.73 percentage points ($p = 0.0122$). This finding provides direct evidence for the mechanism linking our results: exposure to Black teachers increases educational attainment, which enables Black students to enter higher-skill occupations, thereby generating higher earnings and wages. Importantly, these gains occur without negative effects on White students, suggesting that increasing teacher diversity can reduce racial economic inequality without coming at the expense of White students.

Table 8: $10^{th}$ Grade - Labor Market Outcomes

| | Full Sample | | | Employed only | |
|---|---|---|---|---|---|
| Outcomes | (1) Formal employment 2022-2023 | (2) Months worked 2022-2023 | (3) Avg monthly earnings (BRL) 2022-2023 | (4) Avg monthly wage (BRL) 2022-2023 | (5) % Some college (industry-occupation) 2022-2023 |
| Share Black Teachers | 0.0075 | 0.0462 | 12.4141 | 14.4761 | -0.0024 |
| | (0.0081) | (0.1639) | (45.4073) | (77.5815) | (0.0061) |
| Black | 0.0160*** | 0.2388*** | -43.5784*** | -156.7459*** | -0.0310*** |
| | (0.0017) | (0.0342) | (10.1771) | (16.2884) | (0.0012) |
| Share Black Teachers x Black | -0.0101** | -0.0815 | 53.1212** | 98.7664** | 0.0169*** |
| | (0.0043) | (0.0881) | (25.0803) | (43.8174) | (0.0034) |
| Student-level covariates | Yes | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes |
| Effect (Black) | -0.0026 | -0.0353 | 65.5352 | 113.2424 | 0.0145 |
| p-value | 0.7377 | 0.8235 | 0.1282 | 0.1236 | 0.0122 |
| Mean Black | 0.5142 | 9.2836 | 1,430.5880 | 2,827.9706 | 0.2389 |
| Mean White | 0.5492 | 10.1251 | 1,750.5685 | 3,232.9828 | 0.3039 |
| SD Black | 0.4998 | 10.2253 | 2,810.0946 | 3,414.4153 | 0.2813 |
| SD White | 0.4976 | 10.3875 | 3,342.3793 | 3,979.8410 | 0.3136 |
| Observations | 965,664 | 965,664 | 965,664 | 504,891 | 512,655 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments,by academic qualification (undergraduate, specialization, master's, Ph.D.),and male. Standard errors clustered at the school level in parentheses. Net Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## B. $6^{th}$ Grade

Table 9 reports effects of exposure to Black teachers on $6^{th}$ grade lower secondary education outcomes. Column 1 shows no effect on grade approval in the treatment year for either race; coefficients are near zero and not statistically different from zero. The Black–White differential is precisely estimated as zero, in contrast to the significant differential we estimate in $10^{th}$ grade. This divergence is plausibly driven by a ceiling effect, given higher baseline approval rates in the $6^{th}$ grade. Similarly, Column 2 shows no effect on $9^{th}$ grade SAEB participation for either race, nor a differential effect.

Conditional on participation, test-score estimates differ by subject. Column 3 reports a negative, statistically significant effect on Math for White students and no effect for Black students, whereas Column 4 shows a positive effect on Portuguese for Black students and no effect for White students.[25] Column 5 examines on-time middle school graduation. We

[25]The negative Math estimate for White students is specific to that single outcome within the 6th-

find a modest and marginally significant positive differential effect for Black students ($p < 0.10$). The estimate indicates that increasing the share of Black teachers from 0 to 50% is associated with a 0.28 p.p. reduction in the graduation gap between Black and White students. However, the effects for both groups individually are not statistically significant.

Table 9: 6th Grade - Lower Secondary Education Outcomes

| Outcomes | (1) Grade approval 2012 | (2) SAEB participation 2015 | (3) Math z-score 2015 | (4) Portuguese z-score 2015 | (5) Middle school graduation 2015 |
|---|---|---|---|---|---|
| Share Black Teachers | 0.0018 | 0.0075 | -0.0296** | 0.0073 | -0.0019 |
| | (0.0049) | (0.0064) | (0.0137) | (0.0136) | (0.0059) |
| Black | -0.0049*** | -0.0088*** | -0.0515*** | -0.0365*** | -0.0171*** |
| | (0.0010) | (0.0015) | (0.0030) | (0.0030) | (0.0013) |
| Share Black Teachers x Black | 0.0000 | -0.0012 | 0.0396*** | 0.0304*** | 0.0056* |
| | (0.0024) | (0.0033) | (0.0071) | (0.0073) | (0.0031) |
| Student-level covariates | Yes | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes |
| Effect (Black) | 0.0018 | 0.0063 | 0.0100 | 0.0377 | 0.0037 |
| p-value | 0.7131 | 0.3134 | 0.4549 | 0.0050 | 0.5241 |
| Mean Black | 0.8507 | 0.5624 | -0.1400 | -0.1147 | 0.6341 |
| Mean White | 0.8964 | 0.6452 | 0.1394 | 0.1142 | 0.7467 |
| SD Black | 0.3564 | 0.4961 | 0.9652 | 0.9881 | 0.4817 |
| SD White | 0.3048 | 0.4785 | 1.0145 | 0.9987 | 0.4349 |
| Observations | 1,018,697 | 1,018,697 | 612,268 | 612,268 | 1,018,697 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out averag eof peers' prior Math and Portuguese scores,also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged30+, under statutory contracts, with multiple-school assignments,by academic qualification (undergraduate, specialization, master's, Ph.D.),and male. Standard errors clustered at the school level in parentheses. Share Black Teachers ranges from 0 to 1. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

We next turn to the long-run upper secondary school education in Table 10. Column 1 examines on-time high school graduation. We replicate the findings from our previous analysis of the 10th grade cohort. For White students, the effect of exposure to Black teachers is statistically indistinguishable from zero. In contrast, increasing a Black student's

grade cohort and is not observed across other outcomes in the same cohort. Given the large number of comparisons, we view it as consistent with multiple-hypothesis testing. Section VIII. discusses this in detail; Online Appendix Table O.14 shows that the estimate does not survive multiple-testing corrections based on a learning index that aggregates SAEB scores. Under these corrections, effects for Black students remain positive and statistically significant, whereas those for White students are indistinguishable from zero.

exposure to Black teachers from 0 to 50% raises their probability of graduating high school by a statistically significant 0.64 percentage points ($p = 0.0361$). This effect is particularly meaningful in closing the racial achievement gap, which widens over the school years, closing about 30.5% of the within-school gap of 2.78 p.p.

Columns 2–4 analyze ENEM participation and scores. We estimate significant differential positive effects for Black students but no significant effects for either race; this pattern is consistent with limited power to estimate group-specific effects precisely. In Column 2, we examine participation and increasing exposure from 0 to 0.5, which narrows the gap by 0.37 p.p. ($p < 0.05$). In Column 3, we address differential effects on exam taking by imputing the minimum observed score to all non-participants. Increasing the fraction of Black teachers from 0 to 50% yields a 0.0069 SD relative gain for Black students ($p < 0.05$). In Column 4, we instead code the outcome as an indicator for scoring above the median among test takers and assign zero to non-participants; the change raises Black students' relative probability by 0.42 p.p. ($p < 0.10$).

Table 11 focuses on college outcomes. We analyze on-time enrollment in 2019 and cumulative enrollment by 2021, since many students do not enroll immediately after the expected high-school graduation year. The 10th-grade pattern persists on this longer horizon: for White students, we cannot reject that effects are zero, while for Black students, higher exposure to Black teachers yields positive and significant gains in enrollment. Column 1 shows that increasing exposure from 0 to 0.5 raises on-time enrollment by 0.42 p.p. for Black students ($p = 0.032$). Column 2 indicates a 0.53 p.p. increase in enrollment by 2021 ($p = 0.032$). These magnitudes equal 4.2% and 2.9% of the Black mean college enrollment in 2019 and 2021.

Columns 3 and 4 disaggregate enrollment by institution type. Column 3 shows no change in Black students' public-college enrollment, while Column 4 indicates that private colleges drive the increase. Column 5 shows that the enrollments are mostly in high-quality institutions: shifting the share of Black teachers from 0 to 50% increases enrollment in high-quality colleges by 0.30 p.p. ($p = 0.044$), which accounts for 55% of the estimated effect on college enrollment by 2021. We show other evidence consistent with higher value-added degrees in Online Appendix Table O.9: effects are higher for long-duration ($\geq$ 2,400 hours) and for

Table 10: 6th Grade - Upper Secondary Education Outcomes

| Outcomes | (1) High school graduation 2018 | (2) ENEM participation 2018 | (3) ENEM z-score (imputation min) 2018 | (4) ENEM > 50th pct (among participants) 2018 |
|---|---|---|---|---|
| Share Black Teachers | -0.0041 | -0.0028 | -0.0029 | -0.0046 |
| | (0.0063) | (0.0059) | (0.0123) | (0.0041) |
| Black | -0.0278*** | -0.0247*** | -0.0527*** | -0.0.182*** |
| | (0.0014) | (0.0014) | (0.0029) | (0.0022) |
| Share Black Teachers x Black | 0.0169*** | 0.0073** | 0.0138** | 0.0083* |
| | (0.0033) | (0.0032) | (0.0065) | (0.0022) |
| Student-level covariates | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes |
| Effect (Black) | 0.0128 | 0.0045 | 0.0109 | 0.0037 |
| p-value | 0.0361 | 0.4290 | 0.3532 | 0.3477 |
| Mean Black | 0.4545 | 0.3042 | -0.0855 | 0.1058 |
| Mean White | 0.5806 | 0.3903 | 0.0977 | 0.1775 |
| SD Black | 0.4979 | 0.4601 | 0.9371 | 0.3076 |
| SD White | 0.4935 | 0.4878 | 1.0589 | 0.3821 |
| Observations | 1,018,697 | 1,018,697 | 1,018,697 | 1,018,697 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments, by academic qualification (undergraduate, specialization, master's, Ph.D.), and male. Standard errors clustered at the school level in parentheses. Share Black Teachers ranges from 0 to 1. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

in-person programs.

# VI. Heterogeneity Analyses

In this section, we investigate for which individuals and in which contexts the intervention is most effective by analyzing heterogeneous effects across student and school characteristics. We begin by testing for individual-level heterogeneity by baseline math achievement, augmenting our main model with interaction terms in a structure akin to a triple-difference

Table 11: 6th Grade - College Outcomes

| Outcomes | (1) College enrollment 2019 | (2) College enrollment by 2021 | (3) Public college enrollment by 2021 | (4) Private college enrollment by 2021 | (5) High quality college enrollment by 2021 |
|---|---|---|---|---|---|
| Share Black Teachers | -0.0050 | -0.0033 | 0.0041 | -0.0066 | -0.0027 |
| | (0.0041) | (0.0052) | (0.0030) | (0.0048) | (0.0032) |
| Black | -0.0187*** | -0.0278*** | -0.0010 | -0.0280*** | -0.0095*** |
| | (0.0011) | (0.0013) | (0.0007) | (0.0012) | (0.0008) |
| Share Black Teachers x Black | 0.0134*** | 0.0139*** | -0.0036** | 0.0183*** | 0.0087*** |
| | (0.0023) | (0.0028) | (0.0017) | (0.0026) | (0.0016) |
| Student-level covariates | Yes | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes |
| Effect (Black) | 0.0084 | 0.0106 | 0.0005 | 0.0118 | 0.0059 |
| p-value | 0.0321 | 0.0319 | 0.8695 | 0.0096 | 0.0436 |
| Mean Black | 0.0991 | 0.1844 | 0.0486 | 0.1441 | 0.0469 |
| Mean White | 0.1651 | 0.2873 | 0.0688 | 0.2320 | 0.0893 |
| SD Black | 0.2988 | 0.3878 | 0.2151 | 0.3512 | 0.2114 |
| SD White | 0.3713 | 0.4525 | 0.2530 | 0.4221 | 0.2851 |
| Observations | 1,018,697 | 1,018,697 | 1,018,697 | 1,018,697 | 1,018,697 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out averag eof peers' prior Math and Portuguese scores,also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged30+, under statutory contracts, with multiple-school assignments,by academic qualification (undergraduate, specialization, master's, Ph.D.),and male. Standard errors clustered at the school level in parentheses. Share Black Teachers ranges from 0 to 1. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

design. We estimate:

$$
\begin{aligned}
Y_{ics} &= \beta_0 + \beta_1 \,\text{ShareBlackTeachers}_{cs} + \beta_2 \,\text{Black}_{ics} \\
&+ \beta_3 \,\text{BelowMedianMath}_{ics} + \beta_4 \,\text{ShareBlackTeachers}_{cs} \times \text{Black}_{ics} \\
&+ \beta_5 \,\text{ShareBlackTeachers}_{cs} \times \text{BelowMedianMath}_{ics} + \beta_6 \,\text{Black}_{ics} \times \text{BelowMedianMath}_{ics} \\
&+ \beta_7 \,\text{ShareBlackTeachers}_{cs} \times \text{Black}_{ics} \times \text{BelowMedianMath}_{ics} \\
&+ \Phi \, f(\mathbf{A}_{ics}) + \Omega \, f(\bar{\mathbf{A}}_{-i,cs}) + \gamma \, \mathbf{X}_{ics} + \delta \, \mathbf{W}_{ics} + \alpha_s + \varepsilon_{ics}.
\end{aligned}
\tag{3}
$$

where $\text{BelowMedianMath}_{ics}$ is an indicator for students who scored below the median on the baseline math test, and all other notation follows previous definitions. The estimated coefficients allow us to calculate the effects of interest for different subgroups. For Black students, the effect of increasing the share of Black teachers from 0 to 1 is $\beta_1 + \beta_4 + \beta_5 + \beta_7$ for

those with below-median scores, and $\beta_1 + \beta_4$ for their high-achieving peers. The differential effect within Black students is therefore $\beta_5 + \beta_7$. We construct the same parameters for White students: the effect for low-achievers is $\beta_1 + \beta_5$, the effect for high-achievers is simply $\beta_1$, and the resulting differential is $\beta_5$. Appendix O.7 provides the detailed algebra for obtaining these parameters.

Table 12 reports the results for the $10^{\text{th}}$ and $6^{\text{th}}$ grades. For the $10^{\text{th}}$-grade cohort, the positive effects of teacher-student racial matching on high school graduation are almost entirely driven by Black students with below-median baseline achievement, as their higher-achieving Black peers were likely already on track to succeed. However, the impact on the longer-term outcome of college graduation is more distributed. While our estimates suggest an impact on both Black student groups, the effect is larger and more statistically significant for their higher-achieving peers. This may be due to the fact that high-achieving Black students are closer to the college-going margin. A different pattern emerges, however, when examining long-term labor market outcomes. There is a significant and positive effect on earnings only among Black students from the lower half of the achievement distribution. This is likely because the disadvantaged students may realize the greatest economic gains from educational interventions that fundamentally alter their life-course trajectories, unlocking outcomes they would not have achieved otherwise (Zimmerman, 2014).

For the $6^{\text{th}}$-grade cohort, the benefits of having a Black teacher are concentrated among high-achieving Black students, who see significant positive effects on middle school graduation, high school graduation, and college enrollment. For their low-achieving peers, the results suggest a positive long-term impact on their likelihood of on-time high school graduation. This effect is weaker than that seen for low-achieving $10^{\text{th}}$ graders. Nonetheless, cross-grade contrasts are challenging to interpret because the $10^{\text{th}}$-grade sample is conditioned on enrollment in high school, making it difficult to distinguish between selection and other mechanisms.

We next examine heterogeneity by gender. The methodology is analogous to the previous analysis, substituting the BelowMedianMath indicator with a Male dummy in Equation 3. The results, presented in Table 13, reveal a consistent pattern across both cohorts: the positive effects of teacher-student racial matching are driven by the impact on Black boys.

Table 12: $10^{th}$ and $6^{th}$ Grade - Heterogeneity Analysis Across Baseline Math Score

| | $10^{th}$ Grade | | | $6^{th}$ Grade | | |
|---|---|---|---|---|---|---|
| Outcomes | (1) High school graduation 2014 | (2) College graduation by 2021 | (3) Avg monthly earnings (BRL) 2022-2023 | (4) Middle school graduation 2015 | (5) High school graduation 2018 | (6) College enrollment 2019 |
| Share Black Teachers | -0.0314*** | -0.0135** | 6.5259 | 0.0135** | -0.0060 | -0.0050 |
| | (0.0092) | (0.0060) | (51.7376) | (0.0064) | (0.0068) | (0.0049) |
| Black | -0.0247*** | -0.0364*** | -66.4934*** | -0.0169*** | -0.0339*** | -0.0302*** |
| | (0.0019) | (0.0017) | (15.3554) | (0.0016) | (0.0019) | (0.0016) |
| Below median math | -0.0161*** | -0.0142*** | -50.4386*** | 0.0053** | -0.0102*** | -0.0125*** |
| | (0.0025) | (0.0019) | (17.0460) | (0.0023) | (0.0025) | (0.0019) |
| Share Black Teachers x Black | 0.0194*** | 0.0251*** | 17.2606 | 0.0036 | 0.0206*** | 0.0205*** |
| | (0.0053) | (0.0047) | (40.3181) | (0.0042) | (0.0048) | (0.0039) |
| Share Black Teachers x Below median math | 0.0517*** | 0.0137*** | 18.9891 | -0.0310*** | 0.0064 | 0.0048 |
| | (0.0064) | (0.0047) | (40.3608) | (0.0051) | (0.0053) | (0.0039) |
| Black x Below median math | 0.0012 | 0.0266*** | 52.1306*** | -0.0018 | 0.0138*** | 0.0257*** |
| | (0.0028) | (0.0020) | (18.5889) | (0.0025) | (0.0027) | (0.0019) |
| Share Black Teachers x Black x Below median math | 0.0113 | -0.0176*** | 48.6964 | 0.0066 | -0.0108* | -0.0199*** |
| | (0.0076) | (0.0055) | (47.9093) | (0.0059) | (0.0063) | (0.0045) |
| Student-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Effect (Black below median math) | 0.0510 | 0.0078 | 91.4720 | -0.0074 | 0.0102 | 0.0005 |
| p-value | 0.0000 | 0.1045 | 0.0352 | 0.2282 | 0.1069 | 0.8992 |
| Effect (Black above median math) | -0.0120 | 0.0117 | 23.7865 | 0.0171 | 0.0146 | 0.0155 |
| p-value | 0.1765 | 0.0295 | 0.6159 | 0.0060 | 0.0256 | 0.0004 |
| Differential effect (Black below - Black above) | 0.0630 | -0.0039 | 67.6856 | -0.0244 | -0.0044 | -0.0150 |
| p-value | 0.0000 | 0.2219 | 0.0138 | 0.0000 | 0.2780 | 0.0000 |
| Effect (White below median math) | 0.0204 | 0.0002 | 25.5150 | -0.0176 | 0.0004 | -0.0001 |
| p-value | 0.0298 | 0.9622 | 0.5914 | 0.0069 | 0.9526 | 0.9724 |
| Mean Black | 0.6139 | 0.1055 | 1,430.5880 | 0.6341 | 0.4545 | 0.0991 |
| Mean White | 0.6849 | 0.1709 | 1,750.5685 | 0.7467 | 0.5806 | 0.1651 |
| SD Black | 0.4868 | 0.3072 | 2,810.0946 | 0.4817 | 0.4979 | 0.2988 |
| SD White | 0.4646 | 0.3764 | 3,342.3793 | 0.4349 | 0.4935 | 0.3713 |
| Observations | 965,664 | 965,664 | 965,664 | 1,018,697 | 1,018,697 | 1,018,697 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments, by academic qualification (undergraduate, specialization, master's, Ph.D.), and male. Standard errors clustered at the school level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

For the $10^{th}$-grade cohort, this intervention significantly boosts their rates of high school and college graduation, as well as their long-term earnings. This gender-specific effect holds for the younger $6^{th}$-grade cohort, where having a Black teacher leads to significant gains in their future high school graduation and college enrollment. In contrast, Black girls in both cohorts experience little to no statistically significant benefit, making the differential impact between Black boys and girls both large and consistently significant.

Finally, we examine whether the effects of teacher-student racial matching vary by school context. We split the sample into high and low socioeconomic status schools and estimate our main specification (Equation 1) separately for each group.[26]

[26]We classify schools based on SAEB's school socioeconomic index, which is constructed from student-reported socioeconomic variables and divided into seven levels, from lowest to highest SES. We define low SES schools as those at level 4 or below, and high SES schools as those at level 5 or above. We use the SES

Table 13: $10^{th}$ and $6^{th}$ Grade - Heterogeneity Analysis Across Gender

| | $10^{th}$ Grade | | | $6^{th}$ Grade | | |
|---|---|---|---|---|---|---|
| Outcomes | (1) High school graduation 2014 | (2) College graduation by 2021 | (3) Avg monthly earnings (BRL) 2022-2023 | (4) Middle school graduation 2015 | (5) High school graduation 2018 | (6) College enrollment 2019 |
| Share Black Teachers | -0.0084 | -0.0270*** | 13.3352 | 0.0058 | -0.0110 | -0.0222*** |
| | (0.0092) | (0.0057) | (48.4511) | (0.0063) | (0.0068) | (0.0046) |
| Black | -0.0211*** | -0.0417*** | -58.7523*** | -0.0094*** | -0.0236*** | -0.0274*** |
| | (0.0019) | (0.0015) | (12.6591) | (0.0017) | (0.0020) | (0.0016) |
| Male | -0.0842*** | -0.0858*** | 337.6043*** | -0.0662*** | -0.1138*** | -0.0706*** |
| | (0.0018) | (0.0015) | (14.2239) | (0.0016) | (0.0018) | (0.0015) |
| Share Black Teachers x Black | 0.0207*** | 0.0349*** | 7.3614 | 0.0022 | 0.0089** | 0.0165*** |
| | (0.0051) | (0.0040) | (30.5631) | (0.0041) | (0.0045) | (0.0034) |
| Share Black Teachers x Male | 0.0065 | 0.0390*** | -7.0396 | -0.0156*** | 0.0142*** | 0.0347*** |
| | (0.0060) | (0.0047) | (41.0937) | (0.0047) | (0.0050) | (0.0037) |
| Black x Male | -0.0091*** | 0.0374*** | 31.0806* | -0.0156*** | -0.0081*** | 0.0178*** |
| | (0.0027) | (0.0019) | (18.5354) | (0.0024) | (0.0026) | (0.0020) |
| Share Black Teachers x Black x Male | 0.0179** | -0.0324*** | 108.4671** | 0.0066 | 0.0155** | -0.0065 |
| | (0.0073) | (0.0052) | (48.0102) | (0.0058) | (0.0061) | (0.0043) |
| Student-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Effect (Black boy) | 0.0367 | 0.0144 | 122.1241 | -0.0008 | 0.0276 | 0.0226 |
| p-value | 0.0000 | 0.0035 | 0.0085 | 0.8912 | 0.0000 | 0.0000 |
| Effect (Black girl) | 0.0123 | 0.0079 | 20.6966 | 0.0081 | -0.0021 | -0.0057 |
| p-value | 0.1586 | 0.1194 | 0.6391 | 0.1828 | 0.7427 | 0.1710 |
| Differential effect (Black boy - Black Girl) | 0.0244 | 0.0066 | 101.4275 | -0.0089 | 0.0297 | 0.0283 |
| p-value | 0.0000 | 0.0219 | 0.0002 | 0.0185 | 0.0000 | 0.0000 |
| Effect (White boy) | -0.0019 | 0.0120 | 6.2956 | -0.0097 | 0.0032 | 0.0126 |
| p-value | 0.8369 | 0.0292 | 0.9025 | 0.1287 | 0.6385 | 0.0046 |
| Mean Black | 0.6139 | 0.1055 | 1,430.5880 | 0.6341 | 0.4545 | 0.0991 |
| Mean White | 0.6849 | 0.1709 | 1,750.5685 | 0.7467 | 0.5806 | 0.1651 |
| SD Black | 0.4868 | 0.3072 | 2,810.0946 | 0.4817 | 0.4979 | 0.2988 |
| SD White | 0.4646 | 0.3764 | 3,342.3793 | 0.4349 | 0.4935 | 0.3713 |
| Observations | 965,664 | 965,664 | 965,664 | 1,018,697 | 1,018,697 | 1,018,697 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments, by academic qualification (undergraduate, specialization, master's, Ph.D.), and male. Standard errors clustered at the school level in parentheses. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Tables 14 and 15 present the results. Among the $10^{th}$-grade cohort, the effects of Black teachers on reducing racial inequalities in educational attainment, as measured by high school and college graduation, are observed in both high- and low-SES schools. However, the evidence of benefits for Black students is more pronounced in low-SES schools. Estimates for high-SES schools are less precise, likely due to smaller sample sizes. Regarding earnings, the effects are concentrated in low-SES schools, where Black students exposed to more Black teachers experience significant economic gains. In contrast, the point estimate for high-SES schools is near zero and statistically insignificant. For the $6^{th}$-grade cohort, the long-term educational benefits are concentrated almost exclusively in low-SES schools. Black

level closest to our 2012 baseline, drawing on SAEB editions from 2013, 2015, 2017, and 2019

students in these schools experience significant increases in both high school graduation and college enrollment. In contrast, the effects in high-SES schools are economically small and statistically indistinguishable from zero.

Table 14: 10th Grade - Heterogeneity Analysis Across Schools' Socioeconomic Status

| | Low Socioeconomic Status Schools | | | High Socioeconomic Status Schools | | |
|---|---|---|---|---|---|---|
| Outcomes | (1) High school graduation 2014 | (2) College graduation by 2021 | (3) Avg monthly earnings (BRL) 2022-2023 | (4) High school graduation 2014 | (5) College graduation by 2021 | (6) Avg monthly earnings (BRL) 2022-2023 |
| Share Black Teachers | -0.0011 | -0.0072 | 65.3441 | -0.0038 | -0.0088 | -75.2929 |
| | (0.0110) | (0.0060) | (51.9463) | (0.0146) | (0.0099) | (92.3786) |
| Black | -0.0232*** | -0.0239*** | -37.2893** | -0.0236*** | -0.0231*** | -40.8197** |
| | (0.0022) | (0.0016) | (13.9545) | (0.0021) | (0.0016) | (15.8653) |
| Share Black Teachers x Black | 0.0225*** | 0.0180*** | 36.3058 | 0.0304*** | 0.0200*** | 59.0866 |
| | (0.0048) | (0.0035) | (29.7805) | (0.0083) | (0.0067) | (58.8251) |
| Student-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Effect (Black) | 0.0214 | 0.0108 | 101.6499 | 0.0266 | 0.0112 | -16.2063 |
| p-value | 0.0409 | 0.0519 | 0.0329 | 0.0709 | 0.2411 | 0.8602 |
| Mean Black | 0.6126 | 0.0984 | 1290.0688 | 0.6172 | 0.1215 | 1738.5613 |
| Mean White | 0.6770 | 0.1568 | 1554.6908 | 0.6923 | 0.1837 | 1938.6623 |
| Observations | 560,846 | 560,846 | 560,846 | 400,221 | 400,221 | 400,221 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments, by academic qualification (undergraduate, specialization, master's, Ph.D.), and male. Standard errors clustered at the school level in parentheses. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Taken together, our heterogeneity analysis results align with prior work. The benefits of teacher-student racial matching are larger for Black boys, for students with lower baseline achievement, and for those attending low socioeconomic status schools. Stronger responses for Black boys and disadvantaged Black students are consistent with evidence on the long-run impacts of same-race teachers on educational outcomes (Gershenson et al., 2022; Delhommer, 2022), and the higher earnings for these groups align with high returns at the college margin for relatively disadvantaged men (Zimmerman, 2014). The concentration of these gains is plausible if these groups tend to underinvest in human capital, perhaps due to lower expectations or difficulty focusing on school (Fortin et al., 2015). While such underinvestment can depress overall college enrollment, it also creates the potential for high marginal returns from interventions that shift student beliefs through new information or guidance. Black teachers can play this role, raising student expectations and thereby improving long-run

Table 15: 6th Grade - Heterogeneity Analysis Across Schools' Socioeconomic Status

| | Low Socioeconomic Status Schools | | | High Socioeconomic Status Schools | | |
|---|---|---|---|---|---|---|
| Outcomes | (1) Middle school graduation 2015 | (2) High School graduation 2018 | (3) College enrollment 2019 | (4) Middle school graduation 2015 | (5) High School graduation 2018 | (6) College enrollment 2019 |
| Share Black Teachers | 0.0005 | 0.0022 | -0.0015 | -0.0029 | -0.0118 | -0.0055 |
| | (0.0080) | (0.0084) | (0.0050) | (0.0087) | (0.0098) | (0.0071) |
| Black | -0.0173*** | -0.0300*** | -0.0172*** | -0.0163*** | -0.0247*** | -0.0175*** |
| | (0.0021) | (0.0023) | (0.0016) | (0.0017) | (0.0019) | (0.0014) |
| Share Black Teachers x Black | 0.0072* | 0.0191*** | 0.0129*** | -0.0003 | 0.0122** | 0.0050 |
| | (0.0041) | (0.0044) | (0.0029) | (0.0054) | (0.0060) | (0.0046) |
| Student-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Effect (Black) | 0.0078 | 0.0214 | 0.0114 | -0.0033 | 0.0003 | -0.0005 |
| p-value | 0.3109 | 0.0071 | 0.0152 | 0.7233 | 0.9770 | 0.9397 |
| Mean Black | 0.5974 | 0.4206 | 0.0817 | 0.6999 | 0.5055 | 0.1303 |
| Mean White | 0.7008 | 0.5299 | 0.1281 | 0.7781 | 0.6164 | 0.1902 |
| Observations | 536,858 | 536,858 | 536,858 | 477,794 | 477,794 | 477,794 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments, by academic qualification (undergraduate, specialization, master's, Ph.D.), and male. Standard errors clustered at the school level in parentheses. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

outcomes (Gershenson et al., 2016; Papageorge et al., 2020).

# VII. Mechanisms

The literature identifies that teachers could impact students by two channels: passive and active teacher effects (Dee, 2004; Gershenson et al., 2022). The first posits that the mere presence of a teacher from a similar demographic background can influence student outcomes. Same-race teachers may act as role models, shaping students' aspirations and motivation to invest in their own education, even if their behavior is identical to that of other teachers. In contrast, the second channel suggests that these effects may arise from differential teacher behaviors, which include employing culturally relevant pedagogy methods, providing enhanced mentoring and support, and reducing subjective biases (Ladson-Billings, 1995; Dee, 2005; Botelho et al., 2015; Gershenson et al., 2016). Empirically separating these channels in our setting is difficult, as both may operate simultaneously.

Recent research highlights the importance of specific active channels related to teacher preparation. For instance, Edmonds (2024) finds that the positive test-score effects of Black teachers on Black students in North Carolina are largely driven by teachers who graduated from Historically Black Colleges and Universities (HBCUs).[27] This result suggests that a key mechanism is the set of distinct pedagogical skills imparted by such specialized institutional training. This specific active channel is unlikely to be a primary driver of the findings in our context. Unlike the United States, Brazil has no historical system of institutions analogous to HBCUs. Furthermore, specialized teacher trainings that could, in principle, improve outcomes for Black students—such as training on ethnic-racial relations and Afro-Brazilian and African history and culture—are rare.[28]

Another potential active channel is teacher subjective bias, whereby teachers' grading or promotion decisions differ systematically by student race. This may be relevant in our context; for instance, Botelho et al. (2015) provides evidence of racial discrimination through biased student assessments in Brazilian schools. Such findings raise the possibility that differential grading standards among Black teachers could contribute to the observed matching effects. However, if grading bias were the primary mechanism at play, we would expect to see immediate differences in short-run outcomes—such as approval rates during the treatment year—followed by longer-term impacts on educational attainment that stem mechanically from this initial distortion. In our data, however, we find no significant effects on $6^{\text{th}}$-grade approval in the treatment year, while positive effects on on-time high school completion only emerge several years later. This timing pattern is inconsistent with grading-based bias as the main channel, although we cannot entirely rule out other, unobserved forms of teacher bias.

While training-based active channels and evaluation bias seem unlikely to explain our results, we find evidence consistent with broader motivational or mentoring mechanisms. To investigate this, Table 16 reports a "cross-effects" exercise for $6^{\text{th}}$-grade students that re-

[27] Several other studies in the racial-matching literature also use data from states where HBCUs are widespread, including North Carolina. (Lindsay and Hart, 2017; Harbatkin, 2021; Gershenson et al., 2022) and Texas, Tennessee, and Florida (Dee, 2004; Egalite et al., 2015; Delhommer, 2022; Gershenson et al., 2022).

[28] In our sample, fewer than 1% of teachers report having completed any training related to ethnic-racial relations and Afro-Brazilian and African history and culture.

stricts the sample to classrooms where the subject teacher (in Math or Portuguese) is White, and examines whether exposure to Black teachers in other subjects still affects students' outcomes.[29] The logic of this test is that any positive impact observed in this setting cannot arise from a direct instructional channel within the subject itself. The interaction between the share of Black teachers and the Black student indicator remains positive and statistically significant across all specifications examining test scores, including when the subject teacher is White. These cross-subject effects indicate that the benefits of exposure to Black teachers extend beyond the disciplines they directly teach, implying that these teachers may influence students through broader mechanisms that affect engagement and academic effort. This pattern aligns with both passive role-model effects, where the presence of same-race teachers strengthens students' sense of belonging and aspirations, and more active forms of mentoring and support that enhance motivation and persistence across subjects.

# VIII. Robustness Checks

We conduct several robustness checks to validate our main findings. Our first check addresses the sensitivity of our results to the handling of missing CPF identifiers. Our main analysis imputes non-participation in the ENEM, higher education, and the formal labor market for all students without a CPF. We re-run our analysis on the subsample of students with a non-missing CPF. Online Appendix Tables O.10 and O.11 demonstrate that using this alternative sample yields similar results, with a slight decrease in statistical precision from the smaller sample size.

Second, to address the concern that our set of controls may not fully account for non-random sorting on unobservable characteristics, we apply the methodology of Oster (2019). We calculate bias-adjusted treatment effects for 13 main outcomes from the 10th and 6th-grade samples, examining the effects for both White and Black students. As detailed in

[29] We focus on $6^{th}$-grade students for two main reasons. First, at this stage subjects are more weakly linked than in high school, where teachers of disciplines such as Physics or Chemistry may indirectly influence performance in Math. Focusing on $6^{th}$ grade, therefore, provides a cleaner test of cross-subject effects that are less likely to reflect overlapping content or skill transfer. Second, in the $6^{th}$-grade we find no evidence of selection into SAEB exam, which allows us to estimate these relationships without relying on imputations to correct for sample selection.

Table 16: 6th Grade - Mechanisms - Cross-Effects

| | White Portuguese Teacher | | | White Math Teacher | | |
|---|---|---|---|---|---|---|
| | (1) SAEB participation 2015 | (2) Math z-score 2015 | (3) Portuguese z-score 2015 | (4) SAEB participation 2015 | (5) Math z-score 2015 | (6) Portuguese z-score 2015 |
| Share Black Teachers | 0.0090 | -0.0319 | 0.0011 | 0.0111 | -0.0199 | 0.0141 |
| | (0.0103) | (0.0215) | (0.0217) | (0.0101) | (0.0212) | (0.0218) |
| Black | -0.0096*** | -0.0516*** | -0.0355*** | -0.0093*** | -0.0488*** | -0.0348*** |
| | (0.0016) | (0.0033) | (0.0034) | (0.0016) | (0.0033) | (0.0033) |
| Share Black Teachers x Black | 0.0072 | 0.0524*** | 0.0321** | -0.0048 | 0.0326*** | 0.0261** |
| | (0.0059) | (0.0128) | (0.0129) | (0.0059) | (0.0124) | (0.0127) |
| Student-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Effect (Black) | 0.0162 | 0.0205 | 0.0332 | 0.0064 | 0.0127 | 0.0402 |
| p-value | 0.1122 | 0.3421 | 0.1281 | 0.5342 | 0.5576 | 0.0662 |
| Mean Black | 0.5816 | -0.1584 | -0.1328 | 0.5822 | -0.1593 | -0.1334 |
| Mean White | 0.6557 | 0.1138 | 0.0954 | 0.6569 | 0.1137 | 0.0952 |
| Observations | 660,142 | 410,966 | 410,966 | 666,611 | 415,689 | 415,689 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments, by academic qualification (undergraduate, specialization, master's, Ph.D.), and male. Standard errors clustered at the school level in parentheses. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendix Tables O.12 and O.13, our results are highly robust: in all cases, the bias-adjusted coefficient falls within the 99.5% confidence interval of the original estimate.

Third, to address concerns related to multiple hypothesis testing, i.e., the risk of finding significant effects by chance when testing numerous outcomes, we follow the methodology of Anderson (2008). We construct aggregate indices for families of related outcomes and re-estimate our main specification (Equation 1) using these indices as the dependent variables.

For the $10^{\text{th}}$ grade cohort, we construct an educational attainment index and a labor market index. The first index aggregates grade approval, high school graduation, on-time college enrollment, and college graduation. The second index aggregates, for those who work, the number of months worked, the share of workers with some college in the industry-occupation of employment, and the average monthly wage. Similarly, for the $6^{\text{th}}$ grade cohort, we construct a learning index that aggregates student scores on the SAEB exam, and an educational attainment index that aggregates grade approval, middle school completion, on-

time high school graduation, and on-time college enrollment. Online Appendix Table O.14 shows that across all four indices, we find a consistent pattern: the differential effects and the effects for Black students are positive and statistically significant, whereas the effects for White students are indistinguishable from zero. This result confirms that our main findings are not due to multiple hypothesis testing.

# IX. Conclusion

This paper investigates the short- and long-run impacts of Black teachers on student outcomes in Brazil. We leverage linked, nationwide administrative data to track two cohorts of approximately one million students each over 12 years, from middle and high school into higher education and the formal labor market. Our empirical strategy exploits within-school variation by comparing students assigned to classrooms with different shares of Black teachers while controlling for prior test scores. We find that Black teachers have a positive impact on test scores, high school completion rates, college enrollment, and college graduation rates for Black students. These gains lead to higher earnings and increased entry into more human capital–intensive jobs. For White students, the effects are null. Overall, our results suggest that increasing the share of Black teachers to 50% would reduce the covariate-adjusted within-school racial gap in education by at least 30% and in earnings by about 60%, compared to a scenario with no Black teachers.

These findings highlight the importance of racial representation in the teaching workforce for reducing educational and labor market inequalities. In the Brazilian context, increasing Black students' exposure to Black teachers by addressing the underrepresentation of Black educators may serve as an effective policy tool for narrowing racial gaps. According to the 2024 School Census, Black students represent 56.8% of the student population, whereas Black teachers account for only 44.4%, resulting in a 12.4 percentage point disparity. Addressing this disparity may be facilitated by Brazil's affirmative action policies in higher education, implemented in 2014, which have substantially increased Black enrollment in public universities (Mello, 2022), including teacher training programs. While these policies have expanded the pool of prospective Black teachers, achieving greater representation in schools will also

require targeted hiring and retention initiatives.

Given the urgency of addressing racial inequality, a complementary short-term policy could involve redistributing existing Black teachers across classrooms to provide more equitable exposure for Black students. However, our results do not justify segregating schools or classrooms by race, as such policies would have harmful effects on students. Furthermore, although we do not find impacts of Black teachers on White students' academic or labor-market outcomes, such exposure could still positively influence White students' racial attitudes, potentially contributing to the reduction of racial disparities in the longer run.

# References


M. L. Anderson. 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, 2008. doi: 10.1198/016214508000000841. URL https://doi.org/10.1198/016214508000000841.

R. J. Andrews, J. Li, and M. F. Lovenheim. Quantile Treatment Effects of College Quality on Earnings. *Journal of Human Resources*, 51(1):200–238, 2016. ISSN 0022-166X. doi: 10.3368/jhr.51.1.200. URL https://jhr.uwpress.org/content/51/1/200.

R. J. Andrews, S. A. Imberman, M. F. Lovenheim, and K. Stange. The Returns to College Major Choice: Average and Distributional Effects, Career Trajectories, and Earnings Variability. *The Review of Economics and Statistics*, pages 1–45, 09 2024. ISSN 0034-6535. doi: 10.1162/rest_a_01503. URL https://doi.org/10.1162/rest_a_01503.

M. C. Araujo, P. Carneiro, Y. Cruz-Aguayo, and N. Schady. Teacher Quality and Learning Outcomes in Kindergarten. *The Quarterly Journal of Economics*, 131(3):1415–1454, 2016. ISSN 00335533, 15314650. URL https://www.jstor.org/stable/26372667.

N. Barahona, C. Dobbin, J. Fuenzalida, and S. Otero. The Effects of Widespread Online Education on Market Structure and Enrollment. *Working Paper*, March 2025. URL https://hbaraho.github.io/papers/onlineeduc.pdf.

N. Bau and J. Das. Teacher Value Added in a Low-Income Country. *American Economic Journal: Economic Policy*, 12(1):62–96, February 2020. doi: 10.1257/pol.20170243. URL https://www.aeaweb.org/articles?id=10.1257/pol.20170243.

F. Botelho, R. A. Madeira, and M. A. Rangel. Racial Discrimination in Grading: Evidence from Brazil. *American Economic Journal: Applied Economics*, 7(4):37–52, October 2015. doi: 10.1257/app.20140352. URL https://www.aeaweb.org/articles?id=10.1257/app.20140352.

D. G. Britto, A. d. A. Fonseca, P. Pinotti, B. Sampaio, and L. Warwar. Intergenerational Mobility in the Land of Inequality. *IZA Discussion Paper*, September 2022. URL https://ssrn.com/abstract=4241590.

R. Chetty, J. N. Friedman, and J. E. Rockoff. Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review*, 104(9): 2593–2632, September 2014a. doi: 10.1257/aer.104.9.2593. URL https://www.aeaweb.org/articles?id=10.1257/aer.104.9.2593.

R. Chetty, J. N. Friedman, and J. E. Rockoff. Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, 104(9): 2633–79, September 2014b. doi: 10.1257/aer.104.9.2633. URL https://www.aeaweb.org/articles?id=10.1257/aer.104.9.2633.

R. Chetty, N. Hendren, M. R. Jones, and S. R. Porter. Race and Economic Opportunity in the United States: an Intergenerational Perspective. *The Quarterly Journal of Economics*, 135(2):711–783, 5 2020. ISSN 0033-5533. doi: 10.1093/qje/qjz042. URL https://doi.org/10.1093/qje/qjz042.

T. d. L. Coelho, F. Estevan, M. Nakaguma, and R. Rabelo, Alexandre. Do Black Politicians Matter? Political Leadership and Racial Composition in Top Public Sector Positions. *Working Paper*, December 2024. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5183898.

T. S. Dee. Teachers, Race, and Student Achievement in a Randomized Experiment. *The Review of Economics and Statistics*, 86(1):195–210, 02 2004. ISSN 0034-6535. doi: 10.1162/003465304323023750. URL https://doi.org/10.1162/003465304323023750.

T. S. Dee. A Teacher Like Me: Does Race, Ethnicity, or Gender Matter? *American Economic Review*, 95(2):158–165, May 2005. doi: 10.1257/000282805774670446. URL https://www.aeaweb.org/articles?id=10.1257/000282805774670446.

S. Delhommer. High School Role Models and Minority College Achievement. *Economics of Education Review*, 87:102222, 2022. ISSN 0272-7757. doi: https://doi.org/10.1016/

j.econedurev.2021.102222. URL https://www.sciencedirect.com/science/article/pii/S0272775721001333.

E. Derenoncourt, C. H. Kim, M. Kuhn, and M. Schularick. Wealth of Two Nations: The U.S. Racial Wealth Gap, 1860–2020. *The Quarterly Journal of Economics*, 139(2):693–750, 09 2023. ISSN 0033-5533. doi: 10.1093/qje/qjad044. URL https://doi.org/10.1093/qje/qjad044.

L. Edmonds. Which Match Matters? HBCUs, Same-Race Teacher Effects, and Black Student Achievement. *Working Paper*, July 2024. URL https://www.lavaredmonds.com/uploads/1/4/2/8/142800166/hbcus_draft_edmonds_july_2024.pdf.

A. J. Egalite, B. Kisida, and M. A. Winters. Representation in the Classroom: The Effect of Own-Race Teachers on Student Achievement. *Economics of Education Review*, 45:44–52, 2015. ISSN 0272-7757. doi: https://doi.org/10.1016/j.econedurev.2015.01.007. URL https://www.sciencedirect.com/science/article/pii/S0272775715000084.

G. Elacqua, G. Pérez-Nuñez, E. Abarca, and L. Etcheverry. Diversity and Representation Gaps in the Latin American and Caribbean School Workforce, 2024. URL https://doi.org/10.18235/0013225.

F. Estevan, V. Renato, P. Teixeira, M. Rodrigues, and C. Azzoni. For-Profit Higher Education Wage Returns: Evidence from Brazil. *Working Paper*, October 2024. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5043169.

R. W. Fairlie, F. Hoffmann, and P. Oreopoulos. A Community College Instructor Like Me: Race and Ethnicity Interactions in the Classroom. *American Economic Review*, 104(8): 2567–91, August 2014. doi: 10.1257/aer.104.8.2567. URL https://www.aeaweb.org/articles?id=10.1257/aer.104.8.2567.

FAO and ECLAC. People of African Descent in Latin America and the Caribbean: An Exploration of Social and Territorial Realities in the Rural World. Technical report, FAO, 2025. URL https://openknowledge.fao.org/items/2bbef261-790d-499e-82b0-eafca8d36fd2.

N. M. Fortin, P. Oreopoulos, and S. Phipps. Leaving Boys Behind. *Journal of Human Resources*, 50(3):549–579, 2015. ISSN 0022-166X. doi: 10.3368/jhr.50.3.549. URL https://jhr.uwpress.org/content/50/3/549.

F. Gerard, L. Lagos, E. Severnini, and D. Card. Assortative Matching or Exclusionary Hiring? The Impact of Employment and Pay Policies on Racial Wage Differences in Brazil.

*American Economic Review*, 111(10):3418–57, October 2021. doi: 10.1257/aer.20181596. URL https://www.aeaweb.org/articles?id=10.1257/aer.20181596.

S. Gershenson, S. B. Holt, and N. W. Papageorge. Who Believes in Me? The Effect of Student–Teacher Demographic Match on Teacher Expectations. *Economics of Education Review*, 52:209–224, 2016. ISSN 0272-7757. doi: https://doi.org/10.1016/j.econedurev.2016.03.002. URL https://www.sciencedirect.com/science/article/pii/S0272775715300959.

S. Gershenson, C. M. D. Hart, J. Hyman, C. A. Lindsay, and N. W. Papageorge. The Long-Run Impacts of Same-Race Teachers. *American Economic Journal: Economic Policy*, 14 (4):300–342, November 2022. doi: 10.1257/pol.20190573. URL https://www.aeaweb.org/articles?id=10.1257/pol.20190573.

Y. Govind, P. Santini, and E. Derenoncourt. Racial Inequality in France. NBER Working Paper 34013, National Bureau of Economic Research, 7 2025. URL https://www.nber.org/papers/w34013.

E. Harbatkin. Does Student-Teacher Race Match Affect Course Grades? *Economics of Education Review*, 81:102081, 2021. ISSN 0272-7757. doi: https://doi.org/10.1016/j.econedurev.2021.102081. URL https://www.sciencedirect.com/science/article/pii/S0272775721000042.

C. K. Jackson. What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes. *Journal of Political Economy*, 126(5):2072–2107, 2018. doi: 10.1086/699018. URL https://doi.org/10.1086/699018.

E. Karagiannaki. The Scale and Drivers of Ethnic Wealth Gaps across the Wealth Distribution in the UK: Evidence from Understanding Society. LSE Research Online Documents on Economics 119885, London School of Economics and Political Science, LSE Library, Jul 2023. URL https://ideas.repec.org/p/ehl/lserod/119885.html.

M. S. Kofoed and E. mcGovney. The Effect of Same-Gender and Same-Race Role Models on Occupation Choice: Evidence from Randomly Assigned Mentors at West Point. *Journal of Human Resources*, 2017. ISSN 0022-166X. doi: 10.3368/jhr.54.2.0416.7838R1. URL https://jhr.uwpress.org/content/early/2017/10/02/jhr.54.2.0416.7838R1.

G. Ladson-Billings. Toward a Theory of Culturally Relevant Pedagogy. *American Educational Research Journal*, 32(3):465–491, 1995. doi: 10.3102/00028312032003465. URL https://doi.org/10.3102/00028312032003465.

C. A. Lindsay and C. M. D. Hart. Exposure to Same-Race Teachers and Student Disciplinary Outcomes for Black Students in North Carolina. *Educational Evaluation and Policy Analysis*, 39(3):485–510, 2017. doi: 10.3102/0162373717693109. URL https://doi.org/10.3102/0162373717693109.

U. Mello. Centralized Admissions, Affirmative Action, and Access of Low-Income Students to Higher Education. *American Economic Journal: Economic Policy*, 14(3):166–97, August 2022. doi: 10.1257/pol.20190639. URL https://www.aeaweb.org/articles?id=10.1257/pol.20190639.

B. Ost, W. Pan, and D. Webber. The Returns to College Persistence for Marginal Students: Regression Discontinuity Evidence from University Dismissal Policies. *Journal of Labor Economics*, 36(3):779–805, 2018. doi: 10.1086/696204. URL https://doi.org/10.1086/696204.

E. Oster. Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*, 37(2):187–204, 2019. doi: 10.1080/07350015.2016.1227711. URL https://doi.org/10.1080/07350015.2016.1227711.

N. W. Papageorge, S. Gershenson, and K. M. Kang. Teacher Expectations Matter. *The Review of Economics and Statistics*, 102(2):234–251, 05 2020. ISSN 0034-6535. doi: 10.1162/rest_a_00838. URL https://doi.org/10.1162/rest_a_00838.

E. E. Telles, S. R. Bailey, S. Davoudpour, and N. C. Freeman. Racial Inequality in Latin America. *Oxford Open Economics*, 4, 03 2025. ISSN 2752-5074. doi: 10.1093/ooec/odae022. URL https://doi.org/10.1093/ooec/odae022.

L. M. Yarnell and G. W. Bohrnstedt. Student-Teacher Racial Match and Its Association With Black Student Achievement: An Exploration Using Multilevel Structural Equation Modeling. *American Educational Research Journal*, 55(2):287–324, 2018. doi: 10.3102/0002831217734804. URL https://doi.org/10.3102/0002831217734804.

S. D. Zimmerman. The Returns to College Admission for Academically Marginal Students. *Journal of Labor Economics*, 32(4):711–754, 2014. doi: 10.1086/676661. URL https://doi.org/10.1086/676661.

# O Supplemental Appendix

## O.1 Race Variable

A key variable in our analysis is the race/color of students and teachers. However, the raw administrative data for this variable suffers from high non-response rates. To address this, we construct a modal race variable for both students and teachers, assigning each individual their most frequently observed non-missing race/color category across multiple records. We adopt this procedure to increase precision and reduce missing data, following the precedent of other studies that have examined racial issues in Brazil and faced similar data challenges (Gerard et al., 2021). In this section, we detail the construction of this variable for both groups.

For students, we draw race information from two key sources: the School Census (2007-2012) and the Basic Education Assessment System (SAEB) of 2011. Crucially, our calculation of the modal race is limited to these baseline and treatment-year records. This restriction is necessary to prevent potential endogeneity, as the treatment itself may influence who persists in school. The School Census uses the five official race/color categories (*Branco*, *Pardo*, *Preto*, *Amarelo*, and *Indígena*) of the Brazilian Institute of Geography and Statistics (IBGE), while SAEB includes these five plus an "I don't know" option. The data collection methods differ across these sources. In the School Census, race is hetero-reported for students aged 15 or younger, meaning a parent or guardian provides the information. For students aged 16 and over, it is primarily self-reported. In both age groups, the field is filled by the school staff. In contrast, the SAEB race data is self-reported and self-filled by students as part of a socioeconomic questionnaire.

For teachers, we construct a modal race variable using all available years for which we had access to the race variable from the School Census (2007 to 2023) and the Annual Report of Social Information (2008 to 2019). Both datasets use the five official IBGE categories. In the School Census, the race of school professionals is based on self-reporting, though this information is filled in by administrative staff. In RAIS, the race field is completed by the employer, and data-collection processes were not uniform across establishments during our analysis period. While some employers collected this information through employee self-reporting, others may have used hetero-identification.

In Table O.1, Panel A shows that missing race information among students declines substantially from 40.6% to 7.5% in the $10^{\text{th}}$ grade and from 34.7% to 14.0% in the $6^{\text{th}}$ grade when moving from the 2012 School Census classification to the modal-race assignment. Panel B reports a similar reduction for teachers, with missing rates falling from 27.7% to 5.6% in the $10^{\text{th}}$ grade and from 29.9% to 8.5% in the $6^{\text{th}}$ grade.

Table O.1: Distribution of Students' and Teachers' Race by Grade

| | $10^{th}$ Grade | | $6^{th}$ Grade | |
|---|---|---|---|---|
| Race | 2012 School Census (%) | Modal Race (%) | 2012 School Census (%) | Modal Race (%) |
| *Panel A: Students* | | | | |
| Branco | 27.56 | 38.48 | 29.05 | 34.96 |
| Amarelo | 0.47 | 1.67 | 0.37 | 0.95 |
| Indígena | 0.22 | 1.01 | 0.26 | 0.97 |
| Pardo | 28.53 | 44.36 | 32.43 | 42.81 |
| Preto | 2.62 | 6.97 | 3.21 | 6.30 |
| Missing | 40.60 | 7.51 | 34.68 | 14.01 |
| *Panel B: Teachers* | | | | |
| Branco | 48.15 | 60.78 | 41.04 | 52.31 |
| Pardo | 19.46 | 27.56 | 24.38 | 33.04 |
| Preto | 3.65 | 4.72 | 3.68 | 4.72 |
| Amarelo | 0.78 | 0.99 | 0.66 | 0.96 |
| Indígena | 0.25 | 0.31 | 0.37 | 0.45 |
| Missing | 27.72 | 5.63 | 29.86 | 8.52 |

*Notes:* Shares use individual-level raw data (pre-cleaning, pre-filtering). We pool 6th and 10th grades and restrict the universe to students and teachers in these cohorts. The student sample comprises public-school students with prior SAEB records. We code SAEB's "I don't know" as missing. Modal race for students is constructed from the School Census (2007–2012) and SAEB (2011). Modal race for teachers is constructed from the School Census (2007–2023) and RAIS (2008–2019).

## O.2 Research Design

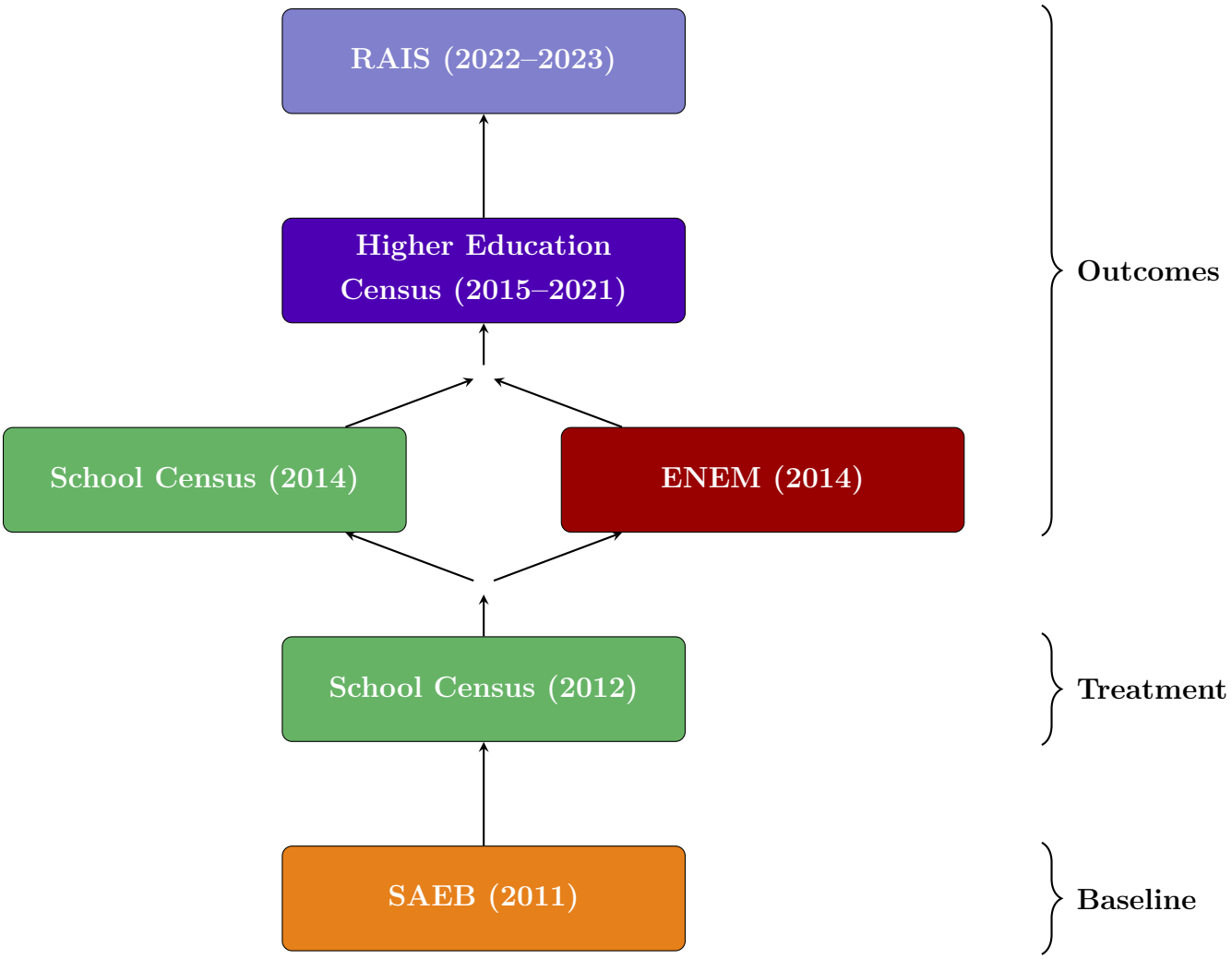



Figure O.1: Longitudinal data structure for the $10^{\text{th}}$ grade cohort analysis

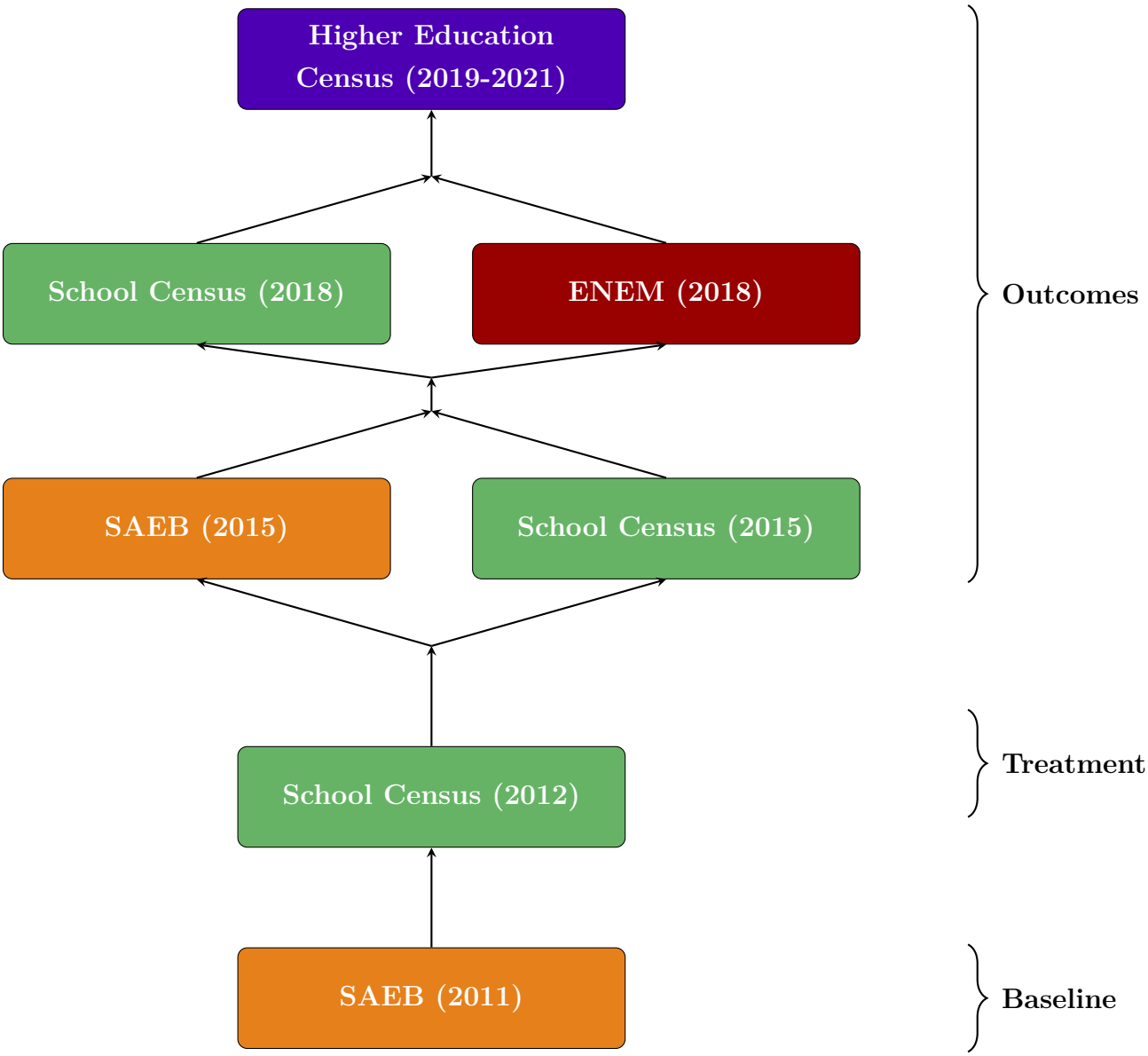



Figure O.2: Longitudinal data structure for the 6$^{th}$ grade cohort analysis.

## O.3 Data Linkage

Our linkage strategy relies on the National Individual Tax Registration Number (Cadastro de Pessoa Física, or CPF). The CPF is Brazil's primary individual taxpayer ID, essential for formal employment and financial activities. As part of the annual School Census, schools collect the CPF along with other personal identification documents, such as the identity card and birth certificate. Although the School Census has included a field for the CPF since 2007, its completion was not mandatory for students during our period of analysis. This non-mandatory collection in schools reflects Brazil's national policy at the time; the

universal CPF requirement for all tax dependents was not established until 2019. This policy context explains why a portion of our 2011 sample lacks a recorded CPF, as is the case for 20.2% of 10th-graders and 24.8% of 6th-graders.

Table O.2 reveals that these missing records are concentrated among observably disadvantaged students, a pattern consistent with lower-income households having less interaction with the tax system. Students without a CPF are more likely to be Black, are slightly older, and have substantially lower baseline test scores. In 10th grade, students with a CPF score $0.16\sigma$ higher in math and $0.14\sigma$ higher in Portuguese; in 6th grade, these gaps are even larger, at $0.35\sigma$ and $0.33\sigma$. These students also face socioeconomic disadvantages, as their guardians have less schooling and their households are less likely to own a car or computer. This pattern extends to academic outcomes observable for the entire sample. For the 10th-grade cohort, students without a CPF have a 6.1 percentage point lower rate of grade approval and a 14.1 percentage point lower on-time high school graduation rate. These differences are even more pronounced for the 6th-grade cohort, which exhibits an 8.3 percentage point gap in grade approval and a 24.8 percentage point gap in high school graduation rates.

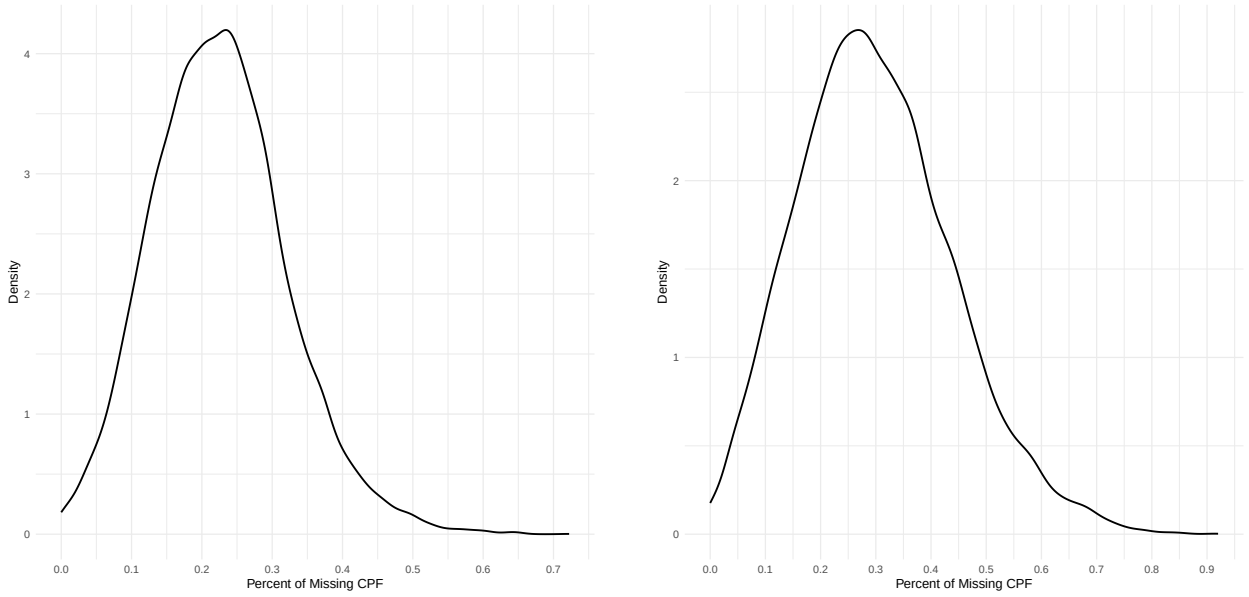
Table O.2: Baseline and Outcome Means by Missing CPF

| | 10th Grade | | | 6th Grade | | |
|---|---|---|---|---|---|---|
| Variable | CPF | No CPF | p-value (Diff.) | CPF | No CPF | p-value (Diff.) |
| Black student | 0.519 | 0.539 | $< 0.001$ | 0.512 | 0.596 | $< 0.001$ |
| Male | 0.460 | 0.464 | $< 0.001$ | 0.491 | 0.504 | $< 0.001$ |
| Age | 15.633 | 15.644 | $< 0.001$ | 11.680 | 11.902 | $< 0.001$ |
| Math z-score 2011 | 0.032 | -0.126 | $< 0.001$ | 0.088 | -0.266 | $< 0.001$ |
| Portuguese z-score 2011 | 0.028 | -0.109 | $< 0.001$ | 0.083 | -0.251 | $< 0.001$ |
| Female guardian has college degree | 0.071 | 0.062 | $< 0.001$ | 0.111 | 0.087 | $< 0.001$ |
| Male guardian has college degree | 0.042 | 0.040 | $< 0.001$ | 0.089 | 0.077 | $< 0.001$ |
| Female guardian has high school | 0.227 | 0.214 | $< 0.001$ | 0.142 | 0.104 | $< 0.001$ |
| Male guardian has high school | 0.159 | 0.144 | $< 0.001$ | 0.094 | 0.069 | $< 0.001$ |
| Lives with male guardian | 0.772 | 0.765 | $< 0.001$ | 0.826 | 0.818 | $< 0.001$ |
| Lives with female guardian | 0.957 | 0.954 | $< 0.001$ | 0.971 | 0.963 | $< 0.001$ |
| Guardians attend school meetings | 0.937 | 0.936 | 0.141 | 0.927 | 0.906 | $< 0.001$ |
| Household asset: car | 0.512 | 0.490 | $< 0.001$ | 0.519 | 0.417 | $< 0.001$ |
| Household asset: computer | 0.648 | 0.620 | $< 0.001$ | 0.583 | 0.461 | $< 0.001$ |
| Share Black Teachers | 0.289 | 0.308 | $< 0.001$ | 0.329 | 0.382 | $< 0.001$ |
| Grade approval 2012 | 0.814 | 0.753 | $< 0.001$ | 0.892 | 0.809 | $< 0.001$ |
| High school graduation 2014 | 0.676 | 0.535 | $< 0.001$ | - | - | - |
| High school graduation 2018 | - | - | - | 0.574 | 0.326 | $< 0.001$ |
| Observations | 777389 | 197253 | | 773508 | 255106 | |

Notes: The table reports mean values of students' characteristics. Reported p-values are from two-sided t-tests of mean differences between students *with* and *without* CPF within each grade.

Figure O.3 shows that the missingness of CPFs is a widespread phenomenon across schools rather than concentrated in a few institutions. For both cohorts, the distribution of the within-school share of students with a missing CPF is unimodal and right-skewed. In the

10th-grade cohort, very few schools have a missing-CPF share of zero, and the distribution peaks with a mode of approximately 22%. A similar pattern holds for the 6th-grade cohort, although the distribution is more dispersed, with a slightly higher mode around 28%. The fact that nearly all schools have a share of students without a CPF reinforces that the missing data are a systemic feature reflecting the broader policy context rather than differences in schools' data collection practices.




10th Grade 6th Grade

Figure O.3: Distribution of the share of students without CPF across schools

This pattern, and our empirical strategy, discussed in detail in Section IV., motivates us to test for systematic differences between students with and without a CPF data within the same school. To conduct this test, we estimate the following regression:

$$MissingCPF_{ics} = \psi_0 + \psi_1 \mathbf{T}_{ics} + \alpha_s + \varepsilon_{ics}, \tag{4}$$

where we regress an indicator for whether student $i$ in classroom $c$ of school $s$ has a missing CPF on $\mathbf{T}_{ics}$, a vector of student-level covariates, and $\alpha_s$, school fixed effects. We then conduct a Wald test of the null hypothesis that student characteristics jointly fail to predict the probability of having a missing CPF.

Table O.3 reports the results. The estimates confirm that, within schools, students without a CPF are a negatively selected group based on characteristics that are positively correlated with future academic success. The Wald tests reject the null hypothesis of no joint predictive power in all specifications ($p < 0.001$). The coefficients on individual character-

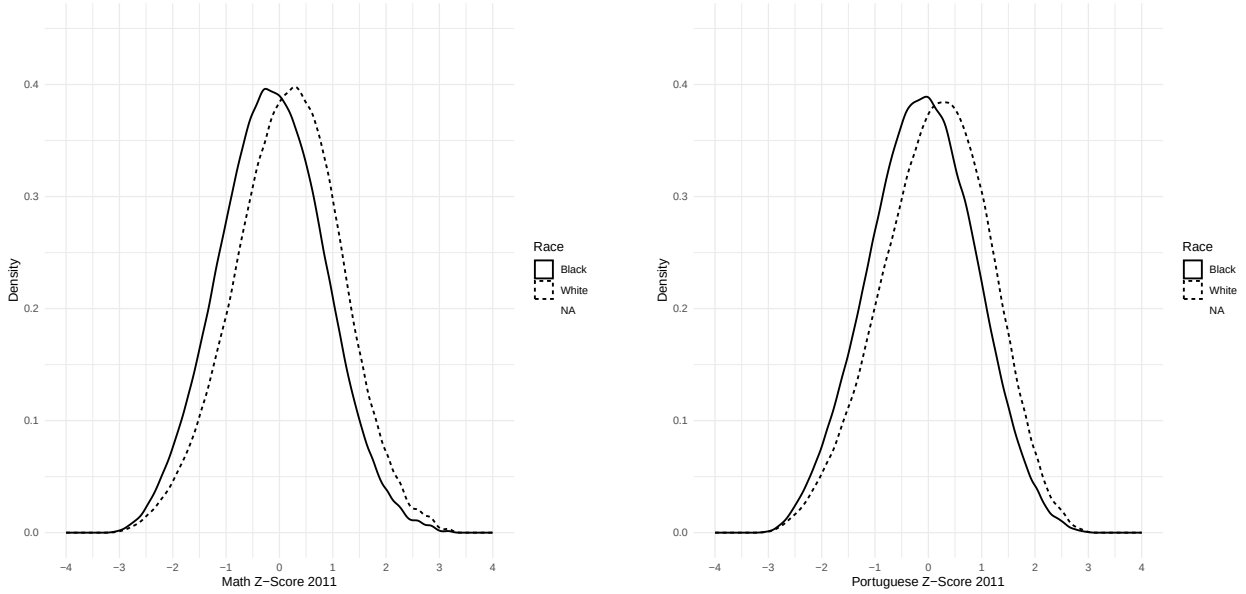
istics reveal the nature of this selection. Lower baseline math and Portuguese scores, fewer household assets, and decreased maternal education are all associated with a significantly higher likelihood of a student having a missing CPF. This evidence of systematic negative selection motivates our decision to impute non-participation in post-secondary outcomes for this group in our main analyses.

Table O.3: 10th and 6th Grade - Sorting on CPF Regressions

| | 10th Grade | | | 6th Grade | | |
|---|---|---|---|---|---|---|
| | All | Black | White | All | Black | White |
| Dependent Var. | Missing CPF | Missing CPF | Missing CPF | Missing CPF | Missing CPF | Missing CPF |
| Male | 0.0053*** | 0.0069*** | 0.0035*** | 0.0025*** | 0.0006 | 0.0047*** |
| | (0.0009) | (0.0012) | (0.0012) | (0.0009) | (0.0012) | (0.0012) |
| Age | -0.0075*** | -0.0070*** | -0.0083*** | 0.0277*** | 0.0260*** | 0.0306*** |
| | (0.0005) | (0.0007) | (0.0008) | (0.0006) | (0.0007) | (0.0009) |
| Portuguese z-score 2011 | -0.0072*** | -0.0077*** | -0.0063*** | -0.0201*** | -0.0208*** | -0.0190*** |
| | (0.0006) | (0.0008) | (0.0008) | (0.0006) | (0.0008) | (0.0008) |
| Math z-score 2011 | -0.0139*** | -0.0134*** | -0.0140*** | -0.0300*** | -0.0313*** | -0.0278*** |
| | (0.0006) | (0.0008) | (0.0008) | (0.0006) | (0.0008) | (0.0008) |
| Female guardian has college degree | -0.0071*** | -0.0071*** | -0.0061*** | -0.0152*** | -0.0180*** | -0.0125*** |
| | (0.0017) | (0.0026) | (0.0023) | (0.0015) | (0.0022) | (0.0020) |
| Male guardian has college degree | 0.0045** | 0.0030 | 0.0053* | 0.0031* | 0.0039 | 0.0020 |
| | (0.0021) | (0.0033) | (0.0028) | (0.0016) | (0.0024) | (0.0022) |
| Female guardian has high school | -0.0055*** | -0.0092*** | -0.0022 | -0.0171*** | -0.0219*** | -0.0119*** |
| | (0.0011) | (0.0015) | (0.0015) | (0.0013) | (0.0019) | (0.0017) |
| Male guardian has high school | -0.0046*** | -0.0032* | -0.0058*** | -0.0058*** | -0.0078*** | -0.0044** |
| | (0.0012) | (0.0018) | (0.0016) | (0.0015) | (0.0023) | (0.0020) |
| Lives with male guardian | -0.0036*** | -0.0018 | -0.0055*** | -0.0031*** | -0.0023 | -0.0039** |
| | (0.0010) | (0.0014) | (0.0016) | (0.0012) | (0.0016) | (0.0017) |
| Lives with female guardian | -0.0081*** | -0.0058** | -0.0109*** | -0.0242*** | -0.0248*** | -0.0236*** |
| | (0.0020) | (0.0028) | (0.0031) | (0.0026) | (0.0035) | (0.0039) |
| Guardians attend school meetings | -0.0075*** | -0.0055** | -0.0097*** | -0.0217*** | -0.0204*** | -0.0235*** |
| | (0.0017) | (0.0025) | (0.0024) | (0.0017) | (0.0023) | (0.0025) |
| Household asset: car | 0.0002 | 0.0019 | -0.0016 | -0.0163*** | -0.0155*** | -0.0171*** |
| | (0.0010) | (0.0013) | (0.0014) | (0.0010) | (0.0014) | (0.0014) |
| Household asset: computer | -0.0158*** | -0.0168*** | -0.0142*** | -0.0327*** | -0.0330*** | -0.0326*** |
| | (0.0011) | (0.0014) | (0.0016) | (0.0010) | (0.0014) | (0.0015) |
| School fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Dependent Mean | 0.2021 | 0.2081 | 0.1955 | 0.2476 | 0.2768 | 0.2143 |
| Wald (joint nullity), p-value | $< 0.001$ | $< 0.001$ | $< 0.001$ | $< 0.001$ | $< 0.001$ | $< 0.001$ |
| Observations | 965,664 | 505,289 | 460,375 | 1,018,697 | 543,174 | 475,523 |

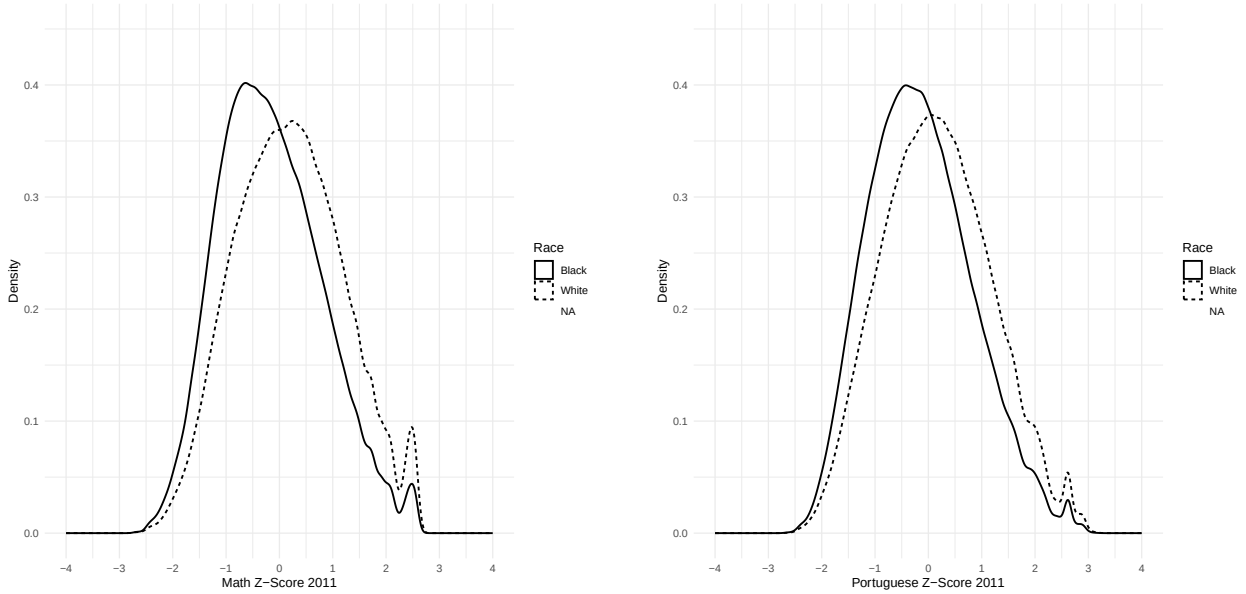
*Notes:* Wald tests jointly whether any student-level coefficient is different from zero. Standard errors clustered at the school level are in parentheses. P-values: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## O.4 Sample and Summary Statistics




(a) Math — $10^{\text{th}}$ grade

(b) Portuguese — $10^{\text{th}}$ grade




(c) Math — $6^{\text{th}}$ grade

(d) Portuguese — $6^{\text{th}}$ grade

Figure O.4: Baseline score distributions by subject and grade. Dashed lines represent White students and solid lines represent Black students.

## Table O.4: Baseline Means by Student Race

| | 6th Grade | | | | | 10th Grade | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Branco | Amarelo | Indígena | Pardo | Preto | Branco | Amarelo | Indígena | Pardo | Preto |
| Missing CPF | 0.213 | 0.269 | 0.291 | 0.275 | 0.291 | 0.195 | 0.224 | 0.233 | 0.206 | 0.224 |
| Male | 0.492 | 0.480 | 0.460 | 0.493 | 0.536 | 0.468 | 0.409 | 0.433 | 0.448 | 0.510 |
| Age | 11.572 | 11.860 | 12.025 | 11.848 | 12.010 | 15.506 | 15.680 | 15.808 | 15.731 | 15.844 |
| Math z-score 2011 | 0.200 | -0.067 | -0.267 | -0.148 | -0.295 | 0.167 | -0.018 | -0.253 | -0.127 | -0.247 |
| Portuguese z-score 2011 | 0.170 | -0.077 | -0.197 | -0.124 | -0.265 | 0.146 | -0.019 | -0.235 | -0.107 | -0.235 |
| Female guardian has college degree | 0.118 | 0.105 | 0.093 | 0.095 | 0.098 | 0.082 | 0.071 | 0.050 | 0.058 | 0.060 |
| Male guardian has college degree | 0.096 | 0.088 | 0.082 | 0.078 | 0.079 | 0.051 | 0.042 | 0.037 | 0.033 | 0.034 |
| Female guardian has high school | 0.148 | 0.124 | 0.116 | 0.121 | 0.110 | 0.245 | 0.221 | 0.193 | 0.207 | 0.206 |
| Male guardian has high school | 0.102 | 0.083 | 0.082 | 0.078 | 0.067 | 0.182 | 0.148 | 0.133 | 0.134 | 0.129 |
| Lives with male guardian | 0.835 | 0.812 | 0.818 | 0.819 | 0.789 | 0.786 | 0.747 | 0.773 | 0.763 | 0.726 |
| Lives with female guardian | 0.973 | 0.959 | 0.961 | 0.966 | 0.964 | 0.959 | 0.946 | 0.954 | 0.954 | 0.954 |
| Guardians attend school meetings | 0.930 | 0.912 | 0.909 | 0.917 | 0.902 | 0.935 | 0.917 | 0.927 | 0.940 | 0.927 |
| Household asset: car | 0.615 | 0.465 | 0.377 | 0.396 | 0.371 | 0.627 | 0.476 | 0.399 | 0.406 | 0.393 |
| Household asset: computer | 0.656 | 0.519 | 0.437 | 0.467 | 0.475 | 0.742 | 0.611 | 0.530 | 0.553 | 0.587 |
| Share Black Teachers | 0.220 | 0.398 | 0.371 | 0.453 | 0.392 | 0.193 | 0.355 | 0.326 | 0.385 | 0.329 |
| Observations | 465337 | 10186 | 9917 | 479541 | 63633 | 444838 | 15537 | 8978 | 441113 | 64176 |

Notes: The table reports mean values of students' characteristics.

## Table O.5: Baseline Means by Teacher Race

| | 6th Grade | | | | | 10th Grade | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variable | Branco | Amarelo | Indígena | Pardo | Preto | Branco | Amarelo | Indígena | Pardo | Preto |
| Male | 0.235 | 0.271 | 0.388 | 0.290 | 0.296 | 0.347 | 0.396 | 0.508 | 0.415 | 0.423 |
| Birthyear | 1971.513 | 1970.709 | 1973.390 | 1972.715 | 1971.657 | 1971.495 | 1969.773 | 1972.010 | 1972.586 | 1971.289 |
| Undergraduate degree | 0.940 | 0.940 | 0.705 | 0.866 | 0.903 | 0.946 | 0.959 | 0.808 | 0.934 | 0.939 |
| Specialization degree | 0.385 | 0.352 | 0.189 | 0.309 | 0.322 | 0.334 | 0.327 | 0.246 | 0.316 | 0.281 |
| Master degree | 0.020 | 0.026 | 0.015 | 0.010 | 0.013 | 0.028 | 0.031 | 0.017 | 0.020 | 0.021 |
| Ph.D. degree | 0.002 | 0.003 | 0.000 | 0.001 | 0.002 | 0.003 | 0.006 | 0.003 | 0.002 | 0.002 |
| Any statutory work regime | 0.741 | 0.817 | 0.525 | 0.730 | 0.751 | 0.697 | 0.798 | 0.534 | 0.698 | 0.724 |
| Work in multiple schools | 0.455 | 0.473 | 0.315 | 0.429 | 0.472 | 0.448 | 0.440 | 0.343 | 0.446 | 0.470 |
| Observations | 92526 | 1596 | 387 | 48422 | 7790 | 110555 | 1690 | 297 | 41878 | 8037 |

Notes: The table reports mean values of teachers' characteristics.

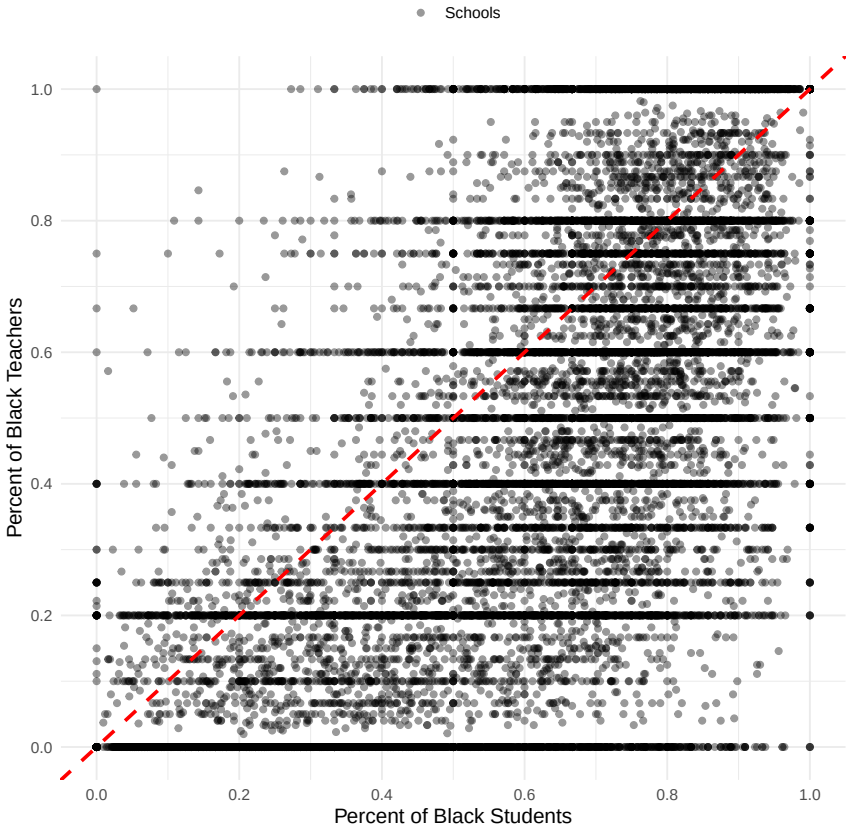



Figure O.5: 6$^{\text{th}}$ Grade: Scatter plot of the share of Black teachers against the share of Black students across schools. Each dot represents a school. The red dashed line indicates parity (45-degree line) between teacher and student racial composition.

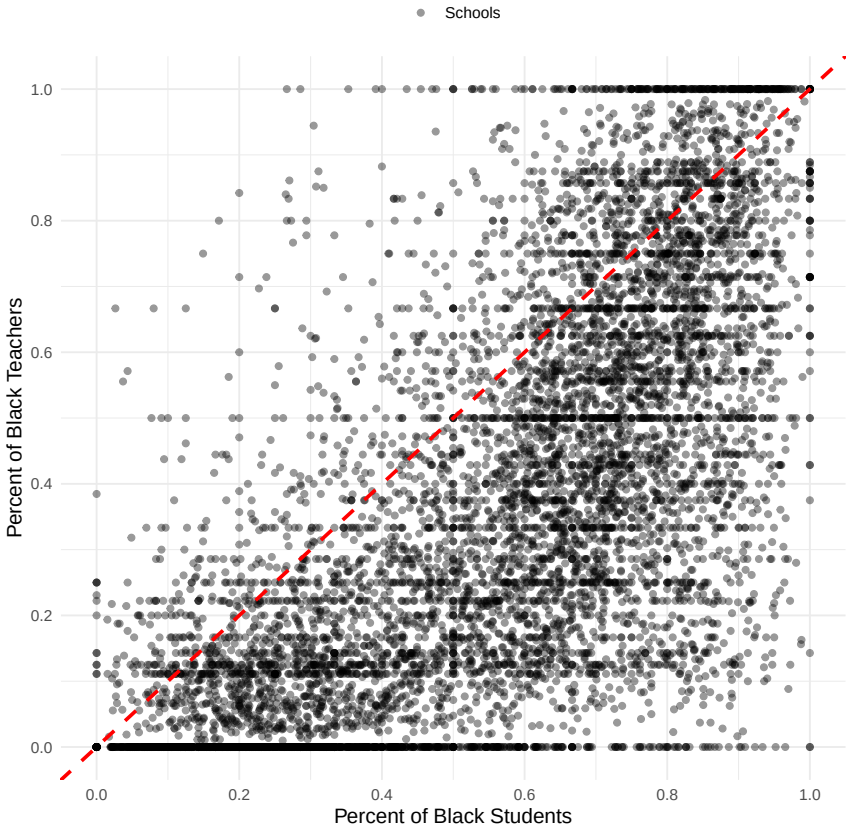



Figure O.6: $10^{\text{th}}$ Grade: Scatter plot of the share of Black teachers against the share of Black students across schools. Each dot represents a school. The red dashed line indicates parity (45-degree line) between teacher and student racial composition.

## O.5 Empirical Stategy

Table O.6: 10th and 6th Grade - Sorting Regressions - CPF Sample

| | 10th Grade | | | 6th Grade | | |
|---|---|---|---|---|---|---|
| | All | Black | White | All | Black | White |
| Dependent Var. | Share Black Teachers | Share Black Teachers | Share Black Teachers | Share Black Share | Share Black Teachers | Share Black Teachers |
| Male | 0.0001 | 0.0000 | 0.0002 | 0.0000 | 0.0000 | -0.0001 |
| | (0.0002) | (0.0003) | (0.0002) | (0.0002) | (0.0003) | (0.0003) |
| Age | 0.0001 | 0.0000 | 0.0003 | -0.0002 | -0.0003 | -0.0001 |
| | (0.0002) | (0.0003) | (0.0002) | (0.0002) | (0.0003) | (0.0003) |
| Portuguese z-score 2011 | -0.0021 | -0.0066 | 0.0023 | 0.0032 | 0.0055 | 0.0009 |
| | (0.0040) | (0.0060) | (0.0052) | (0.0035) | (0.0051) | (0.0050) |
| Math z-score 2011 | 0.0040 | 0.0044 | 0.0059 | -0.0031 | -0.0082 | 0.0046 |
| | (0.0041) | (0.0061) | (0.0053) | (0.0042) | (0.0060) | (0.0061) |
| Female guardian has college degree | -0.0006* | -0.0002 | -0.0010** | 0.0000 | -0.0008 | 0.0006 |
| | (0.0004) | (0.0006) | (0.0004) | (0.0003) | (0.0005) | (0.0004) |
| Male guardian has college degree | 0.0002 | 0.0002 | 0.0002 | 0.0002 | 0.0004 | 0.0003 |
| | (0.0004) | (0.0007) | (0.0005) | (0.0004) | (0.0006) | (0.0005) |
| Female guardian has high school | -0.0002 | 0.0003 | -0.0007*** | 0.0000 | 0.0002 | -0.0002 |
| | (0.0002) | (0.0003) | (0.0003) | (0.0003) | (0.0004) | (0.0004) |
| Male guardian has high school | 0.0001 | 0.0004 | -0.0001 | 0.0003 | 0.0002 | 0.0004 |
| | (0.0002) | (0.0004) | (0.0003) | (0.0003) | (0.0005) | (0.0004) |
| Lives with male guardian | -0.0001 | -0.0002 | 0.0000 | -0.0005* | -0.0003 | -0.0007* |
| | (0.0002) | (0.0003) | (0.0003) | (0.0003) | (0.0004) | (0.0004) |
| Lives with female guardian | 0.0005 | 0.0003 | 0.0006 | 0.0000 | -0.0006 | 0.0005 |
| | (0.0004) | (0.0006) | (0.0006) | (0.0006) | (0.0008) | (0.0009) |
| Guardians attend school meetings | 0.0005 | 0.0007 | 0.0004 | -0.0003 | 0.0005 | -0.0014*** |
| | (0.0004) | (0.0006) | (0.0004) | (0.0004) | (0.0006) | (0.0005) |
| Has car | 0.0002 | 0.0002 | 0.0002 | -0.0001 | -0.0002 | 0.0002 |
| | (0.0002) | (0.0003) | (0.0003) | (0.0002) | (0.0003) | (0.0003) |
| Has computer | -0.0001 | -0.0004 | 0.0003 | 0.0001 | 0.0001 | 0.0002 |
| | (0.0002) | (0.0003) | (0.0003) | (0.0002) | (0.0004) | (0.0003) |
| Polynomials z-scores 2011 | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Dependent Mean | 0.2886 | 0.3747 | 0.1955 | 0.3289 | 0.4360 | 0.2162 |
| Wald (joint nullity), p-value | 0.6261 | 0.3636 | 0.2387 | 0.9531 | 0.8575 | 0.4162 |
| Observations | 770,505 | 400,132 | 370,373 | 766,473 | 392,845 | 373,628 |

*Notes:* Wald tests jointly whether any student-level coefficient is different from zero. Polynomials z-scores 2011 include individuals' own Math and Portuguese z-scores entered flexibly as a cubic polynomial (squared and cubic terms). Standard errors clustered at the school level are in parentheses. P-values: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## O.6 Main Results

### 1. Main Results - 10th Grade

Table O.7: 10th Grade - College Graduation Outcomes

| Outcomes | In-person college graduation 2021 | Online college graduation 2021 | Long course ($\geq$ 2400h) graduation 2021 | Short course ($<$ 2400h) graduation 2021 |
|---|---|---|---|---|
| Share Black Teachers | -0.0045 | -0.0037* | -0.0057 | -0.0031 |
| | (0.0049) | (0.0020) | (0.0048) | (0.0021) |
| Black | -0.0205*** | -0.0037*** | -0.0207*** | -0.0034*** |
| | (0.0010) | (0.0004) | (0.0010) | (0.0005) |
| Share Black Teachers x Black | 0.0147*** | 0.0042*** | 0.0136*** | 0.0054*** |
| | (0.0027) | (0.0012) | (0.0027) | (0.0011) |
| Student-level covariates | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes |
| Effect (Black) | 0.0102 | 0.0006 | 0.0079 | 0.0023 |
| p-value | 0.0256 | 0.7537 | 0.0751 | 0.2407 |
| Mean Black | 0.0906 | 0.0157 | 0.0887 | 0.0174 |
| Mean White | 0.1502 | 0.0221 | 0.1473 | 0.0246 |
| Observations | 965,664 | 965,664 | 965,664 | 965,664 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments,by academic qualification (undergraduate, specialization, master's, Ph.D.),and male. Standard errors clustered at the school level in parentheses. Share Black Teachers ranges from 0 to 1. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table O.8: 10th Grade - College Enrollment Outcomes

| Outcomes | Public college enrollment 2015 | Private college enrollment 2015 | High quality college enrollment 2015 | In-person college enrollment 2015 | Online college enrollment 2015 | Long course ($\geq$ 2400h) enrollment 2015 | Short course ($<$ 2400h) enrollment 2015 |
|---|---|---|---|---|---|---|---|
| Share Black Teachers | 0.0059** | -0.0093* | -0.0002 | -0.0032 | -0.0010 | -0.0043 | 0.0004 |
| | (0.0027) | (0.0049) | (0.0030) | (0.0053) | (0.0015) | (0.0052) | (0.0015) |
| Black | 0.0008 | -0.0174*** | -0.0047*** | -0.0158*** | -0.0008** | -0.0155*** | -0.0012*** |
| | (0.0005) | (0.0011) | (0.0006) | (0.0011) | (0.0003) | (0.0011) | (0.0004) |
| Share Black Teachers x Black | -0.0035** | 0.0153*** | 0.0055*** | 0.0111*** | 0.0011 | 0.0112*** | 0.0012 |
| | (0.0016) | (0.0028) | (0.0015) | (0.0029) | (0.0009) | (0.0029) | (0.0009) |
| Student-level covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Effect (Black) | 0.0024 | 0.0060 | 0.0054 | 0.0078 | 0.0001 | 0.0070 | 0.0016 |
| p-value | 0.3320 | 0.1789 | 0.0558 | 0.1078 | 0.9489 | 0.1463 | 0.2700 |
| Mean Black | 0.0253 | 0.0889 | 0.0252 | 0.1040 | 0.0092 | 0.1038 | 0.0093 |
| Mean White | 0.0316 | 0.1459 | 0.0478 | 0.1642 | 0.0119 | 0.1622 | 0.0140 |
| Observations | 965,664 | 965,664 | 965,664 | 965,664 | 965,664 | 965,664 | 965,664 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments,by academic qualification (undergraduate, specialization, master's, Ph.D.),and male. Standard errors clustered at the school level in parentheses. Share Black Teachers ranges from 0 to 1. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

### 2. Main Results - 6th Grade

Table O.9: 6th Grade - College Outcomes

| Outcomes | In-person college enrollment by 2021 | Online college enrollment by 2021 | Long course ($\geq$ 2400h) enrollment by 2021 | Short course ($<$ 2400h) enrollment by 2021 |
|---|---|---|---|---|
| Share Black Teachers | -0.0051 | 0.0018 | -0.0043 | 0.0010 |
| | (0.0047) | (0.0033) | (0.0050) | (0.0025) |
| Black | -0.0256*** | -0.0044*** | -0.0253*** | -0.0039*** |
| | (0.0012) | (0.0008) | (0.0012) | (0.0006) |
| Share Black Teachers x Black | 0.0151*** | -0.0004 | 0.0117*** | 0.0029** |
| | (0.0026) | (0.0019) | (0.0027) | (0.0013) |
| Student-level covariates | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes |
| Effect (Black) | 0.0099 | 0.0014 | 0.0074 | 0.0039 |
| p-value | 0.0236 | 0.6455 | 0.1166 | 0.0896 |
| Mean Black | 0.1366 | 0.0608 | 0.1628 | 0.0302 |
| Mean White | 0.2270 | 0.0818 | 0.2546 | 0.0472 |
| Observations | 1,018,697 | 1,018,697 | 1,018,697 | 1,018,697 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments, by academic qualification (undergraduate, specialization, master's, Ph.D.), and male. Standard errors clustered at the school level in parentheses. Share Black Teachers ranges from 0 to 1. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## O.7 Heterogeneity Analyses

We index students by $i$, classrooms by $c$, and schools by $s$. The outcome variable $Y_{ics}$ measures student $i$'s result (e.g., test score, graduation, wages) in classroom $c$ at school $s$. We are interested in three variables. First, $B \equiv \text{Black}_{ics} \in \{0,1\}$ is an indicator that the student is Black. Second, $L \equiv \text{BelowMedianMath}_{ics} \in \{0,1\}$ indicates whether the student's baseline math score is below the median (so $L = 1$ marks lower prior achievement). Third, $S \equiv \text{ShareBlackTeachers}_{cs} \in [0,1]$ is the fraction of Black teachers in classroom $c$ (e.g., $S = 0$ means 0% and $S = 1$ means 100%). $f(\mathbf{A}_{ics})$ is a flexible function of student attributes (e.g., baseline scores); $f(\bar{\mathbf{A}}_{-i,cs})$ is a function of peer attributes in the same classroom excluding

$i$; $\mathbf{X}_{ics}$ is a vector of student-level characteristics and $\mathbf{W}_{ics}$ is a vector of classroom-level characteristics. $\alpha_s$ are school fixed effects. Our equation is:

$$\begin{aligned} Y_{ics} &= \beta_0 + \beta_1 S + \beta_2 B + \beta_3 L + \beta_4 SB + \beta_5 SL + \beta_6 BL + \beta_7 SBL \\ &\quad + \Phi f(\mathbf{A}_{ics}) + \Omega f(\bar{\mathbf{A}}_{-i,cs}) + \gamma \mathbf{X}_{ics} + \delta \mathbf{W}_{ics} + \alpha_s + \varepsilon_{ics}. \end{aligned} \tag{5}$$

We present the Conditional Expectation Function (CEF) of the Equation 5 in Equation 6. We omit the full set of control variables and fixed effects from our main regression model for notational clarity. The CEF for individuals with $B = b$, $L = \ell$ and $S = s$ is:

$$E[Y \mid B{=}b, L{=}\ell, S{=}s] = \beta_0 + \beta_1 s + \beta_2 b + \beta_3 \ell + \beta_4 sb + \beta_5 s\ell + \beta_6 b\ell + \beta_7 sb\ell. \tag{6}$$

In all counterfactual exercises, we hold these omitted terms constant. This procedure ensures that any calculated difference in the expected outcome is due solely to changes in $(B, L, S)$, our main variables of interest. For example, a contrast might involve changing S from 0 to 1 for a student with the same observed characteristics within the same school. Using this framework, our parameters of interest are the effects of increasing the share of Black teachers ($S$) from 0 to 1 for different subgroups. Specifically, we are interested in the following parameters:

- **Effect on Black students (Below = 1):**

$$\begin{aligned} &E[Y \mid B{=}1, L{=}1, S{=}1] - E[Y \mid B{=}1, L{=}1, S{=}0] \\ &= (\beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 + \beta_7) \; - \; (\beta_0 + \beta_2 + \beta_3 + \beta_6) \\ &= \boxed{\beta_1 + \beta_4 + \beta_5 + \beta_7}. \end{aligned} \tag{7}$$

- **Effect on Black students (Below = 0):**

$$\begin{aligned} &E[Y \mid B{=}1, L{=}0, S{=}1] - E[Y \mid B{=}1, L{=}0, S{=}0] \\ &= (\beta_0 + \beta_1 + \beta_2 + \beta_4) \; - \; (\beta_0 + \beta_2) \\ &= \boxed{\beta_1 + \beta_4}. \end{aligned} \tag{8}$$

- **Differential effect on Black students: Below = 1 x Below = 0**

$$\begin{aligned} &(E[Y \mid B{=}1, L{=}1, S{=}1] - E[Y \mid B{=}1, L{=}1, S{=}0]) \\ &-(E[Y \mid B{=}1, L{=}0, S{=}1] - E[Y \mid B{=}1, L{=}0, S{=}0]) \\ &= (\beta_1 + \beta_4 + \beta_5 + \beta_7) \; - \; (\beta_1 + \beta_4) \\ &= \boxed{\beta_5 + \beta_7}. \end{aligned} \tag{9}$$

- **Effect on White students (Below = 1):**

$$\begin{aligned} & E[Y \mid B{=}0, L{=}1, S{=}1] - E[Y \mid B{=}0, L{=}1, S{=}0] \\ &= (\beta_0 + \beta_1 + \beta_3 + \beta_5) \; - \; (\beta_0 + \beta_3) \\ &= \boxed{\beta_1 + \beta_5}. \end{aligned} \tag{10}$$

- **Effect on White students (Below = 0):**

$$\begin{aligned} & E[Y \mid B{=}0, L{=}0, S{=}1] - E[Y \mid B{=}0, L{=}0, S{=}0]) - \\ &= (\beta_0 + \beta_1) \; - \; \beta_0 \\ &= \boxed{\beta_1}. \end{aligned} \tag{11}$$

- **Differential effect on White students: Below = 1 x Below = 0**

$$\begin{aligned} & (E[Y \mid B{=}0, L{=}1, S{=}1] - E[Y \mid B{=}0, L{=}1, S{=}0]) \\ & - (E[Y \mid B{=}0, L{=}0, S{=}1] - E[Y \mid B{=}0, L{=}0, S{=}0]) \\ &= (\beta_1 + \beta_5) \; - \; \beta_1 \\ &= \boxed{\beta_5}. \end{aligned} \tag{12}$$

## O.8 Robustness Checks

Table O.10: 10th Grade - Robustness Check (CPF Sample Analyses)

| Outcomes | ENEM participation 2014 | College enrollment 2015 | College graduation by 2021 | Formal employment 2022-2023 | Avg monthly earnings (R$) 2022-2023 |
|---|---|---|---|---|---|
| Share Black Teachers | -0.0015 | -0.0046 | -0.0094 | 0.0100 | 25.6441 |
| | (0.0090) | (0.0065) | (0.0063) | (0.0086) | (54.9058) |
| Black | -0.0223*** | -0.0198*** | -0.0289*** | 0.0215*** | -46.9006*** |
| | (0.0016) | (0.0013) | (0.0013) | (0.0017) | (12.3619) |
| Share Black Teachers x Black | 0.0171*** | 0.0128*** | 0.0206*** | -0.0140*** | 50.8150* |
| | (0.0044) | (0.0037) | (0.0036) | (0.0047) | (30.8713) |
| Student-level covariates | Yes | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes |
| Effect (Black) | 0.0156 | 0.0081 | 0.0112 | -0.0040 | 76.4591 |
| p-value | 0.0715 | 0.1800 | 0.0573 | 0.6327 | 0.1433 |
| Mean Black | 0.4953 | 0.1420 | 0.1332 | 0.6493 | 1,806.5548 |
| Mean White | 0.5267 | 0.2175 | 0.2124 | 0.6827 | 2,175.9630 |
| Observations | 770,505 | 770,505 | 770,505 | 770,505 | 770,505 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments,by academic qualification (undergraduate, specialization, master's, Ph.D.),and male. Standard errors clustered at the school level in parentheses. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table O.11: 6th Grade - Robustness Check (CPF Sample Analyses)

| Outcomes | ENEM participation 2018 | College enrollment 2019 | College enrollment by 2021 |
|---|---|---|---|
| Share Black Teachers | -0.0064 | -0.0048 | -0.0029 |
| | (0.0071) | (0.0054) | (0.0066) |
| Black | -0.0273*** | -0.0218*** | -0.0317*** |
| | (0.0017) | (0.0013) | (0.0016) |
| Share Black Teachers x Black | 0.0123*** | 0.0159*** | 0.0159*** |
| | (0.0038) | (0.0030) | (0.0035) |
| Student-level covariates | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes |
| Effect (Black) | 0.0059 | 0.0111 | 0.0130 |
| p-value | 0.3944 | 0.0305 | 0.0404 |
| Mean Black | 0.4206 | 0.1371 | 0.2550 |
| Mean White | 0.4968 | 0.2102 | 0.3656 |
| Observations | 766,473 | 766,473 | 766,473 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores,also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments, by academic qualification (undergraduate, specialization, master's, Ph.D.), and male. Standard errors clustered at the school level in parentheses. Share Black Teachers ranges from 0 to 1. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table O.12: 10th Grade - Regression Table with Oster's Robustness Test

| | Grade approval 2012 | High school graduation 2014 | ENEM participation 2014 | College enrollment 2015 | College graduation by 2021 | Avg monthly earnings (R$) 2022-2023 | Avg monthly wage (R$) 2022-2023 | % Some college (industry-occupation) 2022-2023 |
|---|---|---|---|---|---|---|---|---|
| Share Black Teachers | -0.0061 | -0.0054 | -0.0003 | -0.0036 | -0.0082 | 12.4141 | 14.4761 | -0.0024 |
| | (0.0084) | (0.0087) | (0.0079) | (0.0054) | (0.0051) | (45.4073) | (77.5815) | (0.0061) |
| Black | -0.0165*** | -0.0254*** | -0.0185*** | -0.0165*** | -0.0239*** | -43.5784*** | -156.7459*** | -0.0310*** |
| | (0.0013) | (0.0015) | (0.0015) | (0.0011) | (0.0011) | (10.1771) | (16.2884) | (0.0012) |
| Share Black Teachers x Black | 0.0180*** | 0.0291*** | 0.0162*** | 0.0123*** | 0.0187*** | 53.1212** | 98.7664** | 0.0169*** |
| | (0.0033) | (0.0039) | (0.0040) | (0.0030) | (0.0029) | (25.0803) | (43.8174) | (0.0034) |
| Student-level covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Effect (Black) | 0.0119 | 0.0236 | 0.0158 | 0.0087 | 0.0105 | 65.5352 | 113.2424 | 0.0145 |
| | (0.0083) | (0.0085) | (0.0077) | (0.0050) | (0.0048) | (43.0802) | (73.5418) | (0.0058) |
| p-value | 0.1521 | 0.0055 | 0.0391 | 0.0816 | 0.0290 | 0.1282 | 0.1236 | 0.0122 |
| Bias-Adjusted: Share Black Teachers | 0.0062 | 0.0100 | 0.0195 | 0.0082 | 0.0030 | 68.0982 | 81.4824 | 0.0072 |
| Robust to 99.5% CI | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Set Excludes Zero | No | No | No | No | No | Yes | Yes | No |
| Bias-Adjusted: Share Black Teachers x Black | 0.0088 | 0.0183 | 0.0016 | 0.0028 | 0.0099 | 5.1669 | 27.3559 | 0.0085 |
| Robust to 99.5% CI | Yes | Yes | No | No | No | Yes | Yes | Yes |
| Set Excludes Zero | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Bias-Adjusted: Effect (Black) | 0.0149 | 0.0280 | 0.0209 | 0.0110 | 0.0130 | 73.2649 | 108.8382 | 0.0156 |
| Robust to 99.5% CI | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Set Excludes Zero | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| R-squared | 0.1524 | 0.2006 | 0.2078 | 0.1528 | 0.1458 | 0.0480 | 0.0612 | 0.2462 |
| Observations | 965664 | 965664 | 965664 | 965664 | 965664 | 965664 | 504891 | 512655 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments,by academic qualification (undergraduate, specialization, master's, Ph.D.),and male. Standard errors clustered at the school level in parentheses. Net Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table O.13: 6th Grade - Regression Table with Oster's Robustness Test

| | Grade approval 2012 | Middle school graduation 2015 | High school graduation 2018 | ENEM participation 2018 | College enrollment 2019 |
|---|---|---|---|---|---|
| Share Black Teachers | 0.0018 | -0.0019 | -0.0041 | -0.0028 | -0.0050 |
| | (0.0049) | (0.0059) | (0.0063) | (0.0059) | (0.0041) |
| Black | -0.0049*** | -0.0171*** | -0.0278*** | -0.0247*** | -0.0187*** |
| | (0.0010) | (0.0013) | (0.0014) | (0.0014) | (0.0011) |
| Share Black Teachers x Black | 0.0000 | 0.0056* | 0.0169*** | 0.0073** | 0.0134*** |
| | (0.0024) | (0.0031) | (0.0033) | (0.0032) | (0.0023) |
| Student-level covariates | Yes | Yes | Yes | Yes | Yes |
| Peer-level covariates | Yes | Yes | Yes | Yes | Yes |
| Teacher-level covariates | Yes | Yes | Yes | Yes | Yes |
| School fixed effects | Yes | Yes | Yes | Yes | Yes |
| Effect (Black) | 0.0018 | 0.0037 | 0.0128 | 0.0045 | 0.0084 |
| | (0.0049) | (0.0058) | (0.0061) | (0.0057) | (0.0039) |
| p-value | 0.7131 | 0.5241 | 0.0361 | 0.4290 | 0.0321 |
| Bias-Adjusted: Share Black Teachers | 0.0073 | 0.0080 | 0.0070 | 0.0070 | 0.0021 |
| Robust to 99.5% CI | Yes | Yes | Yes | Yes | Yes |
| Set Excludes Zero | Yes | No | No | No | No |
| Bias-Adjusted: Share Black Teachers x Black | -0.0087 | -0.0093 | 0.0001 | -0.0079 | 0.0027 |
| Robust to 99.5% CI | No | No | No | No | No |
| Set Excludes Zero | No | No | Yes | No | Yes |
| Bias-Adjusted: Effect (Black) | -0.0013 | -0.0013 | 0.0071 | -0.0009 | 0.0048 |
| Robust to 99.5% CI | Yes | Yes | Yes | Yes | Yes |
| Set Excludes Zero | No | No | Yes | No | Yes |
| R-squared | 0.1750 | 0.2790 | 0.2647 | 0.2269 | 0.1518 |
| Observations | 1018697 | 1018697 | 1018697 | 1018697 | 1018697 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments,by academic qualification (undergraduate, specialization, master's, Ph.D.),and male. Standard errors clustered at the school level in parentheses. Net Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table O.14: Robustness Check - Multiple Hypothesis Testing

| | $10^{th}$ Grade | | $6^{th}$ Grade | |
|---|---|---|---|---|
| | Educational Attainment Index | Labor Market Index | Learning Index | Educational Attainment Index |
| Share Black Teachers | -0.0154 | -0.0125 | -0.0115 | -0.0065 |
| | (0.0116) | (0.0142) | (0.0112) | (0.0088) |
| Black | -0.0514*** | -0.525*** | -0.0440*** | -0.0414*** |
| | (0.0021) | (0.0028) | (0.0024) | (0.0020) |
| Share Black Teachers x Black | 0.0461*** | 0.0372*** | 0.0350*** | 0.0238*** |
| | (0.0054) | (0.0079) | (0.0058) | (0.0045) |
| Effect (Black) | 0.0307 | 0.0246 | 0.0235 | 0.0172 |
| p-value (Effect on Black students) | 0.0058 | 0.0700 | 0.0321 | 0.0418 |
| p-value (Effect on White students) | 0.1861 | 0.3771 | 0.3033 | 0.4554 |
| p-value (Differential effect on Black students) | $< 0.0001$ | $< 0.0001$ | $< 0.0001$ | $< 0.0001$ |
| Observations | 965,664 | 504,891 | 612,268 | 1,018,697 |

*Notes:* All regressions include controls for student characteristics, including gender, age, parental education, family structure, parental school involvement, and household assets, as well as baseline achievement in Math and Portuguese entered flexibly as a cubic polynomial (linear, squared, and cubic terms). Peer composition includes the class-level fraction of Black students and the leave-out average of peers' prior Math and Portuguese scores, also entered as cubic polynomials. Teacher composition is measured at the class level by fractions of teachers aged 30+, under statutory contracts, with multiple-school assignments, by academic qualification (undergraduate, specialization, master's, Ph.D.), and male. Standard errors clustered at the school level in parentheses. Share Black Teachers ranges from 0 to 1. Effect (Black) is computed as the sum of the coefficient on Share Black Teachers and the interaction term with Black Student. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.