Teacher Quality and Learning Inequality*

Calidad Docente y Desigualdades de Aprendizaje

Macarena kutscher**
Catalina morales***
Cristina riquelme****
Sergio urzúa*****

Abstract

This paper explores the contribution of teachers to student performance in Chile's college admission test (PSU). Our analysis is based on a unique teacher-student matched dataset and decomposition methods. The findings suggest that teachers' performance on the PSU and the characteristics of their educational degrees are significant predictors of students' success. When controlling for students' and predetermined school characteristics, the gapbetween voucher and public schools is reduced. Productivity differences emerge as key factors driving the disparities across school types. The analysis underscores the crucial role of teacher-student interactions in shaping student outcomes.

Key words: *Student performance, teacher characteristics, sorting, education inequalities.*

JEL Classification: *I2, I24, J24*

---

* This paper has been prepared for the conference "Aniversario N°50 Revista Estudios de Economía," organized by the School of Economics and Business of the University of Chile. We thank Valentina Paredes for her insightful comments and suggestions. This research used the SIMCE databases of the Chilean Ministry of Education as a source of information. The authors thank the Ministry of Education for access to the information. All the study results are the authors' responsibility and do not compromise this Institution in any way". We thank the Department of Evaluation, Measurement, and Educational Registration (DEMRE) of the University of Chile for providing the University Higher Education Admission System databases for developing this research. The authors presented this paper in seminars at the Universidad de Chile, The State University of New York at Buffalo, and Duke University. We thank the participants for their helpful comments.
** Inter-American Development Bank. e-mail: macarenak@iadb.org.
*** St. Mary's College of Maryland. E-mail: cbmoralesleman@smcm.edu.
**** Trinity University. E-mail: mriquelm@trinity.edu.
***** University of Maryland and NBER. E-mail: surzua@umd.edu.

Received: August, 2023 Accepted: August, 2024

Resumen

Este documento explora la contribución de los profesores al desempeño estudiantil en la prueba de admisión universitaria de Chile (PSU). Nuestro análisis utiliza un conjunto de datos único de profesor-estudiantes y métodos de descomposición. Los resultados sugieren que el desempeño de los profesores en la PSU y sus títulos educativos son predictores significativos del éxito estudiantil. Al controlar por las características predeterminadas de los estudiantes y la escuela, la brecha entre colegios subvencionados y públicos se reduce. Las diferencias de productividad surgen como factores clave que impulsan las disparidades entre los tipos de colegios. El análisis subraya el papel crucial de las interacciones entre profesores y estudiantes en los resultados educacionales.

Palabras clave: *Rendimiento estudiantil, características del profesor, selección, desigualdades educativas.*

Clasificación JEL: *JEL: I2, I24, J24.*

"The art of teaching is the art of assisting discovery",
Mark Van Doren

1. Introduction

Education plays a vital role as a determinant of personal and societal development (Heckman 2000), with teachers identified as a critical input in this process (Rockoff 2004; Rivkin, Hanushek, and Kain 2005; Aaronson, Barrow, and Sander 2007; Chetty, Friedman, and Rockoff 2014a, 2014b; Jackson 2018; Gilraine and Pope 2021; Petek and Pope 2023). Moreover, research consistently demonstrates that the learning environment and resources provided by schools profoundly impact student outcomes (Jackson, Johnson, and Persico 2016). However, understanding the complex dynamics between schools, teachers, and student achievement remains an essential question with implications for policy initiatives and educational reforms.

This paper examines the factors that determine student achievement, focusing on the impact of teachers and schools on students' outcomes. To address this question, we use a unique dataset from Chile, which gathers administrative information from multiple sources. The extensive dataset includes records of students, teachers, and schools. However, what makes this data unique is the availability of detailed variables describing teachers' performance in high-stakes college admission assessments at the age of 17-18, how high they ranked "education" as their career of choice when applying to college, their high school GPA, and detailed information on their professional degree in education. This granular level of information merged with student-level results on Chile's college admission exams allows us to go beyond what the literature has explored about the teacher-student dyad.[1]

We investigate whether there are differences among schools in their ability to enhance student academic performance and to what extent these differences can be attributed to teacher quality. To do this, we employ a multi-step approach. First, we estimate a production function for student achievement using a value-added specification. We then explore how each input contributes to reducing the performance gap between public and voucher school students. Subsequently, we examine whether teacher quality can account for the disparities in the test score distributions across different school types. To achieve this, we utilize both classical and RIF Oaxaca-Blinder decompositions. Finally, following the methodology outlined in Firpo, Fortin, and Lemieux (2018), we implement an empirical strategy that decomposes the achievement gap into a composition effect (due to differences in the distribution of observed characteristics) and a structure effect (due to differences in the productivity of observed characteristics). This approach enables us to analyze mean performance differences between students in each type of school and to explain the gap across the entire performance distribution.

Given the pronounced levels of segregation within Chile's educational system and the substantial disparities in student outcomes, this paper contributes to the literature on multiple fronts. While previous studies have highlighted the influence of socioeconomic factors in explaining the performance gap across school types (Mizala and Romaguera 2000; Contreras 2002; Bravo, Mukhopadhyay, and Todd 2010; Iturra and Gallardo 2022), we examine the role of teachers. We address this gap by delving into granular information within the performance production function, focusing on the college admission test performance as our variable of interest. Due to the complex, many-to-many nature of the student-teacher relationship and our high-stakes outcome, we concentrate on high school students. Our sample comprises over 400,000 test-takers between 2013 and 2021. A limitation of our study is that we only have access to information on college admission assessments from 2006 onward. Consequently, our analysis is confined to investigating the role of young teachers, as only for them can we observe their performance in the same test their students are taking.

---

1 For more structural analysis of the teacher-student relationship and the process of accessing higher education in Chile, see Montaño et al. (2023).

This research yields several findings. First, the value-added model demonstrates significant transmission of college admission test performance (PSU) from teachers to students. Additionally, other characteristics of teachers, such as being more experienced or having a higher proportion of them with a formal education degree, are associated with higher student performance. However, even after accounting for student background characteristics and a comprehensive set of teacher attributes, the type of school attended in high school continues to play an essential role in explaining PSU performance. This suggests that students' and teachers' characteristics alone cannot fully explain the performance gap observed between students attending public and private-subsidized schools in Chile.

Then, we delve into estimating the decomposition of the PSU-performance gap across school types into its contributing components, such as students' family characteristics and previous standardized test performance, teacher characteristics, and school-specific PSU take-up rate. The Oaxaca-Blinder analysis reveals that we can explain 15 out of the 23-point gap in math and 12 out of the 28-point gap in Spanish solely by accounting for the observed characteristics between the two groups. These differences emerge when comparing each group's mean college admissions test performance. When we apply the RIF Oaxaca-Blinder decomposition to the average performance, we find that teachers' influence prominently manifests in the form of a price effect, suggesting differences in the productivity of teachers by school type, particularly in math. This finding indicates that teachers with similar characteristics exhibit greater effectiveness in voucher schools, thereby contributing to performance disparities.

Exploiting the RIF Oaxaca-Blinder decomposition, we further examine the role of each contributing component throughout the entire distribution of test scores. For teacher characteristics, we identify the heightened significance of the structure effect in the high-end of the performance distribution. Specifically, we find that at the top 80% of the test score distribution, the teacher structure effect explains up to 30 points of the school-type performance gap in mathematics, indicating that teacher characteristics substantially influence student performance among high-achieving students. On the other hand, for Spanish, we document that the effect of teachers explaining the gap is more important on the lower part of the distribution but much smaller in magnitude. Finally, we find evidence suggesting complementarity between students' past performance and teachers' characteristics, suggesting that teachers' productivity effect is more prominent when students have better baseline performance.

To the best of our knowledge, this is the first paper to explore the transmission of teacher-student performance in the context of college admission tests in Chile. In a previous study, Contreras (2002) investigated the impact of school type on college admissions test scores; however, teacher characteristics were not included in the analysis. Other studies conducted in Chile have analyzed the effects of teacher characteristics on lower-stake exams (Toledo and Valenzuela 2015; Canales and Maldonado 2018; Barrios Fernández and Riudavets 2021). In addition to examining a higher-stake exam, our paper also considers teachers' performance on the college admissions test as a relevant factor in explaining the performance gap across students attending public and voucher schools.

The remainder of the paper is organized as follows. Section 2 describes the institutional background of the educational system in Chile. Section 3 summarizes the previous literature, and Section 4 describes the data. Section 5 presents the methodology and results of an exploratory analysis of the main factors that determine student performance on the college admission test. Section 6 presents the methodology and the decompositions of the achievement gap across school types for different moments of the distribution, and Section 7 concludes.

2. Institutional Background

The Chilean educational system consists of eight years of primary and four years of secondary education. There are three types of schools: public schools, funded and administered by the government; voucher (private-subsidized) schools, which receive partial funding from the government through a voucher system and are administered by the private sector; and private fee-paying schools, funded and administered by the private sector. Regarding the distribution of students, approximately 40% are in public schools, 50% are in voucher schools, and only 10% are in private fee-paying schools.

Throughout primary and secondary education, students undergo SIMCE examinations (Sistema de Medición de la Calidad de la Educación), standardized assessments conducted nationally to evaluate education quality and school performance. These assessments cover subjects relevant to each grade level, including mathematics and reading comprehension.

Successful completion of secondary education is a prerequisite for admission to higher education institutions in Chile. Most higher education institutions select their students using a centralized deferred acceptance admission system that only considers the performance of students in secondary education and a standardized national university entrance exam (PSU).[2,3] The PSU is

[2] There are some few exceptions that include Special Admissions, which are reserved slots for students who meet specific criteria, such as athletes, indigenous students, or students with disabilities, and Admission by Merit, reserved for students with exceptional academic achievements or talents in specific fields.

[3] In the recent years the PSU has been reformed, but the admissions system remained the same. For the years considered in this study, PSU is the relevant test.

usually taken during the last year of secondary education (12th grade) at the end of the academic year. It consists of two mandatory sections, Mathematics and Language and Communication (Spanish), and at least one of the other sections, Scientific Reasoning or History, Geography, and Social Sciences. Some private universities do not participate in the centralized system and have admission tests or criteria that may differ from the PSU.

It is crucial to note that many students attend "preuniversitarios", institutions preparing them for the PSU, offering content review, test-taking strategies, and simulations. While our study primarily focuses on teachers within traditional academic settings, we acknowledge the potential interplay with "preuniversitarios", despite a lack of available data to assess this issue.

Successful completion of secondary education is a prerequisite for admission to higher education institutions in Chile. Most higher education institutions are part of a centralized deferred acceptance admission system, in which students' performance in secondary education and the standardized national university entrance exam (PSU) are the main factors for acceptance. Student admission to each program depends on individual performance, the reported ranking of program-university bundle according to their preferences, and available slots.

Despite recent reforms addressing inequality, challenges persist in the Chilean education system, marked by disparities in access, educational quality, and funding among public, voucher, and private schools. While public schools often serve disadvantaged populations, voucher schools attract better teachers, enjoying more hiring autonomy and curriculum development flexibility (Elacqua 2012; Behrman et al. 2016). Efforts to mitigate disparities, including increased funding for disadvantaged students, are ongoing, but challenges remain, especially regarding school segregation despite the introduction of a centralized school admission system (Kutscher, Nath, and Urzúa 2023).

3. Literature Review

There is a long-standing literature documenting how the quality of teaching significantly impacts students' academic performance (Hanushek et al., 2007; Chetty et al., 2014). The evidence suggests that teachers are among the most influential factors in explaining student achievement (Hanushek 2011). In particular, several studies have shown that an improvement in teacher quality by one standard device leads to a roughly 0.1 standard deviation increase in student test scores (Rockoff 2004; Rivkin, Hanushek, and Kain 2005; Aaronson, Barrow, and Sander 2007).

The impact of teachers extends beyond academic performance. Research by Jackson (2018) and Petek and Pope (2023) reveals that teachers also influence nontest score behaviors, such as absences and suspensions. These dimensions of teacher quality have been found to have a lasting impact on students' long-term outcomes. In a different context, Chetty, Friedman, and Rockoff (2014b) show that students assigned to better teachers are more likely to go to college and earn higher salaries.

Although there is broad consensus on the importance of teacher quality, accounting for it remains challenging, and studies differ on the extent of specific teacher factors in enhancing students' outcomes. In recent years, the adoption of Value-Added Models (VAMs) has become prevalent in educational research. For example, for the United States, Chetty, Friedman, and Rockoff (2014a) estimate that a standard deviation improvement in teacher value-added increases normalized test scores by 0.14 and 0.1 standard deviations in math and English, respectively. However, a common criticism of VAMs is their limited focus, namely, identifying the general contributions of teachers to learning but providing little information on which teacher characteristics contribute more to improving student outcomes (Wei et al. 2012). We aim to contribute to this issue by analyzing the impact of different dimensions of teacher characteristics on high-stakes test score performance.

Most studies on the impact of teacher quality have focused on the US context. However, a handful of studies have focused on the case of Chile. For example, Canales and Maldonado (2018) finds that teacher quality significantly affects eighth-grade standardized test scores, especially in math. They found no significant effect of teacher credentials but showed that the impact of teachers increases with professional experience. Similarly, Toledo and Valenzuela (2015) show that attributes such as short-term specific professional training and better curriculum coverage positively impact the performance of fourth-grade students. Barrios and Riudavets (2021) conduct teachers' VAMs and find that higher-quality teachers positively affect student test scores, high school graduation, higher education attendance, and the type of higher education institutions attended.

In a recent study, García-Echalar, Poblete, and Rau (2023) used VAMs to investigate the impact of teachers on gender gaps in standardized test scores. Their results reveal that, in general, teachers do not account for the existing math or Spanish score gaps between the genders. Interestingly, their research uncovers variations dependent on school type, with teacher value-added measures mitigating gender gaps in voucher schools but showing no such effect in public schools. This finding also motivates us to examine the impact of the school type in our context.

In this paper, we take one step further and analyze whether teacher quality can explain the performance gap observed by different types of schools in Chile. Previous studies have examined the test achievement gap across school types in primary and secondary education in Chile, using standardized test scores for students in the fourth, eighth, or tenth grades. For example, Bellei (2005) explored the relationship between school type and student performance in the fourth and tenth grades. Their findings indicate that, once accounting for sorting students due to selective admission processes and the exclusion of retained students, private schools are not more effective than public schools and may be less effective. Furthermore, Mizala and Romaguera (2000) analyzed the performance gap in the SIMCE test scores. Their research revealed that the test score gap between vouchers and public schools disappears when controlling for family socioeconomic characteristics.

Investigating college admission test results is pertinent, as they represent a high-stakes assessment in the educational context. Consistent with this, Contreras (2002) explores the influence of the type of school on college admission tests in conjunction with other SES variables. The findings reveal that the school's effect on student performance in college admission tests is notably substantial and statistically significant, even after controlling for parental education levels.

In this paper, we exploit a much richer dataset that allows us to control for a more comprehensive set of teacher variables, including teachers' performance in college admission test assessments. Recent evidence by Neilson et al. (2022) shows a positive and concave relationship between pre-college academic achievement and subsequent teacher productivity. Their evidence suggests that college entrance exams could be helpful to select or recruit students entering teacher colleges. This result underscores the potential role of including teachers' standardized college admission performance as a proxy for their productivity.

4. Data

We integrate data from multiple sources to investigate the factors influencing students' PSU performance. A time-invariant individual masked identifier allows us to establish connections between students, their teachers, and their historical performance and educational decisions, remaining consistent across various administrative datasets and over time. This section details the information we can extract from each dataset and the sample restrictions required to define our study sample.

We have access to DEMRE (Departamento de Evaluación, Medición y Registro Educacional) data on the national college admission test results for all students taking the PSU between 2006 and 2021. We use these data to identify teachers' performance on this test before entering higher education and assess students' performance in cohorts between 2013 and 2021. As some students retake the PSU, we only keep their first scores. It is important to note that not all graduating students take the PSU, as it is not mandatory.

We merge eighth-grade SIMCE records for each student in math and Spanish tests and information on their gender and their mothers' highest educational degree attained (high school, technical, professional, post-graduate), which we will use as a proxy for socioeconomic status (SES). Due to the SIMCE assessment design, only six out of nine cohorts of students with PSU score data underwent an eighth-grade SIMCE examination. Our methodology, relying on a value-added specification, considers the entire history of students' past input before high school. Consequently, we limit our analysis to students who completed the SIMCE test in eighth grade and subsequently took the PSU on time in their senior year, while also attending school each year of high school.

We retrieve each teacher's subject information for each classroom, grade, and year from administrative records, identifying whether a teacher is responsible for teaching Spanish or math to the students in our sample. These records include additional attributes such as gender, age, years of teaching experience, and whether they have a formal degree in Education. DEMRE datasets provide information related to teachers' PSU performance, how high they rank Education as their program of choice ranking in college applications, and the institution they select.

Given the multidimensional context, where each student can potentially have multiple teachers for various subjects, and each teacher instructs several students, we aggregate teacher characteristics throughout their secondary education. This involves sequentially averaging the characteristics of teachers of the corresponding subject at the classroom level for each grade. If no teacher information is available for a classroom, we attribute information based on the average characteristics of other classrooms in the same cohort, grade, and school, and the school average across grades and years if information is missing. This approach considers each student as the primary unit of observation.

It is essential to acknowledge some sample limitations. First, our analysis is restricted to exploring the impact of young teachers since we only have PSU data from 2006 onward. We use teachers' performance in this test as a critical determinant of students' PSU performance. Therefore, we can only use the subsample of teachers observed as students taking the test and, years later, as teachers in a secondary school classroom. Second, as not all students have a young teacher, we assign information on the average teacher characteristics at the cohort-school level to those for whom we do not observe the actual teacher's characteristics. Third, for comparability in the results, we focus on students in the regular education system, excluding those attending special education due to a disability or incarceration and those attending night school.

Consequently, our final sample comprises 428,973 observations for math and 415,315 for Spanish, with 307,169 students appearing in both subject samples.

4.1 Descriptive Statistics

Table 1 presents the summary statistics of all student characteristics (Panel A) and average teacher characteristics (Panel B) for the sample of students considered in the analysis. Columns (1) to (4) refer to characteristics of public school students in the sample, while columns (5) to (8) refer to attributes of voucher school students, in the sample, with columns (1) and (5) showing average values for the math sample in each type of school, and (3) and (7) for Spanish.

The first row in Panel A presents the statistics for the main dependent variable, the PSU score. We see that the difference in average math and Spanish scores between voucher and public schools is about 23 and 28 points, respectively, representing a difference of about 0.2-0.3 standard deviations. To further put into perspective how large this gap is, consider that the average difference in year-to-year changes in cutoffs for admission into undergraduate programs is about 15 points, and the median of this difference is only 10 points. Figure 1 presents the distribution of the PSU scores for the students in the sample in math and Spanish, showing that the gap between schools is present not only for the mean value but for most of the distribution.

The remaining rows in Panel A present additional characteristics. On average, public school students outperformed voucher school students in the eighth-grade SIMCE knowledge test by 0.22 standard deviations in math and 0.12 standard deviations in Spanish. These differences can be seen in Figure 2, where the difference in the distributions across groups for math is much more severe than that for Spanish. The lagged fraction taking each subject-specific PSU in each type of school also differs, with only 72% of students in public schools taking the tests, while the proportion in voucher schools reaches 78%. Other characteristics appear to be much more balanced across the school types, with around 55% of the test-taker students being female, and with 7-10% of mothers holding a technical degree, 28-30% holding a professional degree, and 3-5% of them having a post-graduate degree and the rest, 55-62% holding at most a high school degree.

From Table 1, Panel B, we also learn about the differences in the average characteristics of teachers. Voucher school teachers, especially Spanish teachers, score much higher than public school teachers in the PSU of the subject they teach. We also observe this pattern in Figure 3. At the same time, public school teachers had higher grade point averages when graduating from high school than voucher school teachers, although these differences are minor compared to the ones observed for the PSU scores. Additionally, public school teachers are two to three percentage points less likely to hold a degree from a highly selective institution but slightly more likely to have an education degree. Interestingly, it is more common for Spanish teachers to list education as a top 3 choice in their college application ranking than for math teachers, and teachers in public schools show a lower tendency to list education in their top application ranking than voucher school teachers. Finally, we observe that the teachers in the sample are about 31 years old and have only three to four years of teaching experience in both types of schools. This pattern is consistent with the fact that the teachers in our sample took the college admissions test after 2005, so they are relatively young. We should remember this fact when interpreting the results, as we cannot easily extrapolate the findings to all teachers in the system.

5. Predicting Academic Success: Exploratory Analysis

In this section, we explore the factors that influence student achievement in college admission tests, particularly emphasizing the impact of teachers and schools. It is well-established that socioeconomic characteristics, schools, and particularly teachers, strongly predict students' performance. When examining the characteristics of teachers that predict student performance, previous research has often concentrated on years of experience and academic credentials. In addition to these usual teachers' characteristics, we also study the potential role of teachers' performance in the college admissions test and whether education was among their preferred choices when applied to college.

We incorporate teachers' performance in the college admissions test and students' application preferences when they apply to college, which is a novel contribution to the existing literature. To the best of our knowledge, we are the first to attempt to study the relationship between student performance in college admission tests and teachers' performance in the very same test.

We estimate the following VAM separately for each subject, math and Spanish, to understand college admission performance:

$$(1)\quad Y_{i,s,t} = \beta_0 + \beta_1 Voucher_{i,s,t} + \beta_2 \mathbf{SES}_{i,s,t} + \beta_3 \mathbf{Teacher}_{i,s,t} + \beta_4 SIMCE^{8b}_{i,s,t} + \beta_5 PSU^{take-up}_{s,t-4} + \gamma_t + \varepsilon_{i,s,t}$$

where i denotes a student, s a school, and t a year where the outcome of interest $Y_{i,s,t}$ corresponds to the PSU score of the student i' in the year t graduating from high school s. $Voucher_{i,s,t}$ is an indicator variable taking value one if students attended a voucher school in high school, zero if it is public.[4] This variable captures any gap between comparable students in each type of school. $\mathbf{SES}_{i,s,t}$ is a categorical variable we construct from the student's mother's highest level of education (no higher education degree, technical tertiary degree, university degree, or graduate degree). We use this variable to proxy for the socioeconomic status of the student. $\mathbf{Teacher}_{i,s,t}$ represents a vector that encompasses the mean characteristics of the teacher observed throughout the high school years of a student. This vector incorporates several factors, including the average performance of teachers in the PSU subject in which they teach, the average high school GPA (measured on the PSU scale), the fraction of teachers who have an education degree, the fraction of teachers who ranked education among their top three choices in college applications, and the proportion of teachers who attended a selective university.[5] Additionally, this variable includes the usual characteristics used in the literature, such as average teacher experience, age, and the proportion of female teachers. The variable $SIMCE_{i,s,t}^{8b}$ captures the students' performance on the eighth-grade SIMCE test for the corresponding subject. Including a baseline performance measure allows us to interpret the results as a value-added specification, where SIMCE is a sufficient statistic for the educational input of students before high school (Todd and Wolpin 2003). Lastly, to control for possible selection on the PSU take-up across schools, we control for the proportion of students in the school who took the PSU test four years earlier, captured by $PSU_{s,t-4}^{take-up}$. We calculated it lagged to reduce concerns about potential endogeneity issues.[6] We also include in the estimations the gender of the student and year-fixed effects, γ_t, to capture aggregate shocks to PSU results at the national level.

5.1 Results

Table 2 presents the results of estimating equation (1) using students' PSU performance as the dependent variable. We begin by examining the impact of the type of high school attended by students and gradually incorporate other

---

4 Only 6.19% of students in our sample switched from a voucher to a public school, or vice-versa. For these students, we consider the school they graduated from.

5 The select universities considered are Pontificia Universidad Católica de Chile (PUC) and Universidad de Chile (UCH), which are the most selective institutions in the country (Bordón, Canals, and Mizala 2020).

6 As Table 1 indicates, approximately 70% of high-school students in our sample take the PSU, which might raise concerns about the impact of self-selection on test taking in the college admission test performance. Since we are interested in examining the impact of school type, we cannot use school-fixed effects to account for this potential source of bias. However, we interpret our lagged measure of school-specific PSU take-up rate and students' pre-high school performance as sufficient statistics accounting for this self-selection into testing.

relevant variables into subsequent columns. In all specifications, we include year-fixed effects. Column (1) presents the results, including only the voucher school indicator variable. The voucher coefficient indicates an average math PSU score difference of 23.3 points between the two school types. In Column (2), we observe that the PSU gap decreases to 22 points when controlling for student background variables, such as gender and maternal education.

Column (3) controls for teacher information. The average teacher's math PSU result is a robust positive predictor of student math PSU performance. A one-standard-deviation increase in the teacher's score is associated with a rise of 9 points, approximately 0.1 standard deviations. The proportion of teachers with a formal education degree also exhibits a positive, statistically significant relationship with the PSU score, indicating that a 10% increase in the fraction of teachers with an education degree corresponds to a 3.17-point increase in the PSU score. The interplay between average years of experience and average teacher age almost cancels each other out, possibly due to the teacher sample's youth and some collinearity. Notably, other teacher characteristics, such as having a degree from a selective institution, selecting education in their top 3 college application choices, and the proportion of female teachers, are not significant predictors after controlling for the aforementioned variables.

Even after accounting for student SES characteristics and a comprehensive set of teacher features, the type of school attended continues to play a crucial role in explaining PSU performance. In Column (4), with the inclusion of eighth-grade math SIMCE scores, the school gap reduces to 9.4 points. The coefficient for SIMCE in eighth grade suggests that a one-standard-deviation higher math SIMCE score is associated with a 62-point increase in the predicted math PSU score. This result indicates the significant role played by the educational inputs students received in primary education, reducing the relevance of the high school attended. Additionally, the inclusion of past test scores transforms the estimate into a Value-Added Model (VAM), indicating a substantial increase in the model's goodness-of-fit. It is essential to highlight that the eighth-grade test score not only captures pre-high school academic preparation but also incorporates other educational investments, such as parental involvement and innate talent, influencing better performance in standardized tests.

Finally, in Column (5), the inclusion of a proxy for the school's propensity to have students taking the PSU further reduces the gap to 7.5 points. The estimate for this lagged PSU take-up variable is statistically significant and positive. Consequently, the reduction in the gap aligns with the fact that voucher schools, as indicated in Table 1, are more likely to have their students take the PSU. By including this by-school measure of the tendency to take the test, we control for school-specific heterogeneity that affects students' likelihood of self-selecting into taking the PSU. The coefficients for teacher characteristics decrease in size after including SIMCE scores and lagged PSU take-up rates, but they remain statistically significant, albeit at around a third of their size in Column (3).

We find similar results when we compare the Spanish PSU performance in Table 3. In Column (1), there is an expected conditional voucher gap of 28.3 points, which reduces to 25.6 points in Column (3) after controlling for students' gender, socioeconomic characteristics, and teacher characteristics. The results closely mirror those of math PSU. Average Spanish PSU performance of teachers and the fraction of teachers with an education degree strongly correlate with the Spanish PSU performance of students. Column (4) reveals a considerable improvement in the model's goodness of fit, indicating a decrease of 7 points in the relevance of school type when including students' eighth-grade Spanish SIMCE scores. Finally, Column (5) shows that the school gap decreases an additional 2.5 points with the inclusion of the school PSU take-up proxy. The estimates for teacher characteristics decrease to about half the size observed in Column (3) but remain statistically significant.

The above specifications assume that the coefficients for each explanatory variable must be the same across both voucher and public schools. However, this might not be the case if there are productivity differences in using any of those variables. Guided by the results in Tables 2 and 3, we apply the methodology developed by Firpo, Fortin, and Lemieux (2018) and Rios-Avila (2019) to analyze whether teacher quality can explain the performance gap observed by different types of schools. In the next section, we estimate the decomposition of the school type gap in the average PSU performance by the contribution of socioeconomic characteristics, teachers, own past performance, and propensity to take the PSU test, allowing more flexibility to explore the isolated effects coming from the different elements of the model.

6. Bridging The School-type Gap: O-B Decompositions

The regression analysis in the previous section identifies many influential factors that explain the gap in PSU performance, measured as the difference in the means of student scores in the two types of schools. We implement a classical Oaxaca-Blinder decomposition to separately estimate how much of the difference comes from the composition effect, i.e., the differences in covariates between the two groups, and how much comes from the structure effect, i.e., the estimated coefficients, which in this educational setting is akin to the productivity of the covariates. Then, we take the analysis one step further by implementing a Recentered Influence Functions (RIF) Oaxaca-Blinder Decomposition, which allows us to go beyond simple mean comparisons to consider gaps in other statistics independent of the decomposition's sequential order. We follow Firpo, Fortin, and Lemieux (2018) to perform the RIF Oaxaca-Blinder decomposition analysis to explain the differences in the PSU performance of students between schools for both mean and quantiles, separating the differences in the distributions into composition and structure effects, decomposing each effect by the contribution of each covariate, combining the RIF Oaxaca-Blinder analysis and the reweight strategy proposed by DiNardo, Fortin, and Lemieux (1996).

Classical Oaxaca-Blinder Decomposition

We first implement a conventional Oaxaca-Blinder decomposition of the form:

$$(2)\ E\left[PSU|X,\text{Voucher}\right]-E\left[PSU|X,\text{Public}\right]=\underbrace{\bar{X}^{V}\left(\hat{\beta}_V-\hat{\beta}_P\right)}_{\text{Unexplained}}+\underbrace{\left(\bar{X}^{V}-\bar{X}^{P}\right)\hat{\beta}_P}_{\text{Explained}},$$

where $\bar{X}^k$ denotes a vector containing the averages of the independent variables for students enrolled in type k's schools, and $\hat{\beta}_k$ is the associated vector of point estimates obtained from a linear regression model (k $\in$ {Public,Voucher}). As is standard in the literature, the school gap attributed to the differences in means is denoted by the "explained" gap, while the part attributed to the coefficient discrepancies is the "unexplained" gap. Taking into account the evidence presented in Tables 2 and 3, we exclude from the characteristics of the teachers the nonsignificant variables (selective education dummy, education as top choice dummy, and teacher's gender) to minimize the noise in the estimations.

RIF Oaxaca-Blinder Decomposition

We then implement the RIF Oaxaca-Blinder decomposition in the following manner, following the terminology laid out in Rios-Avila (2019). Assume that there is a joint distribution that describes all relationships between PSU scores, Y, exogenous characteristics, X, and the categorical variable indicating the types of schools to be compared, T. Then, we can rewrite the PSU distribution conditional on school type as:

$$(3)\qquad f^k_{Y|X}\left(y,x\right)=f^k_{Y|X}\left(Y|X\right)f^k_X\left(X\right),$$

$$(4)\qquad F^k_Y\left(y\right)=\int F^k_{Y|X}\left(Y|X\right)dF^k_X\left(X\right),$$

where k indicates whether the density is conditional on the type of school, T = k with k $\in$ {0,1}.

Then, differences in any distributional statistic v can be calculated as:

$$\begin{aligned} \Delta v &= v_1 - v_0 \\ &= v\left(F_Y^1\right) - v\left(F_Y^0\right) \\ &= v\left(F_{Y|X}^1\left(Y|X\right)dF_X^1\left(X\right)\right) - v\left(F_{Y|X}^0\left(Y|X\right)dF_X^0\left(X\right)\right). \end{aligned} \tag{5}$$

From equation (5) it follows that the differences in statistics v can arise from differences in average characteristics $\left(dF_X^1\left(X\right) \neq dF_X^0\left(X\right)\right)$, or differences in coefficients $\left(F_{Y|X}^1\left(Y|X\right) \neq F_{Y|X}^0\left(Y|X\right)\right)$. To separately estimate how relevant the composition and structure effects are in separately explaining the school-type gap, it is needed a third statistic, a counterfactual one, that permits the consideration of step-wise variations:

$$v_c = v\left(F_Y^c\right) = v\left(F_{Y|X}^0\left(Y|X\right)dF_X^1\left(X\right)\right)$$

With this counterfactual statistic, we can decompose Δv in equation (5) as:

$$\Delta v = \underbrace{v_1 - v_c}_{\Delta v_s} + \underbrace{v_c - v_0}_{\Delta v_x},$$

where Δv_s denotes the structure effect and Δv_x represents the composition effect. However, v_c is, by definition, a counterfactual statistic and, therefore, not observable in the data. This unobservability represents an empirical challenge that the methodology sorts out by approximating the relevant distribution as follows.

$$F_Y^c\left(y\right) = F_{Y|X}^0\left(Y|X\right)dF_X^1\left(X\right) \approx F_{Y|X}^0\left(Y|X\right)dF_X^0\left(X\right)\omega\left(X\right),$$

where the weights, $\omega\left(X\right)$, can be obtained applying Bayes rule:

$$\omega\left(X\right) = \left[\frac{1-P}{P}\right] \times \left[\frac{P(T=1 \mid X)}{1-P(T=1 \mid X)}\right],$$

where P is the proportion of students in school type T = 1, and P(T = 1|X) is the conditional probability that someone with characteristics X belongs to a school type T = 1. Thus, by reweighting $dF_X^0\left(X\right)$, we can proxy for v_c.

We estimate P(T = 1|X) using a logit model, includings as the main explanatory variables the percentage of voucher schools in the municipality of residence of students, the total number of voucher schools in the municipality of residence of students, mother's education (less than high school, technical, professional or graduate degree), eighth-grade student-specific SIMCE scores on both subjects, gender, the average experience of teachers in the municipality of residence, and average PSU scores of teachers in the municipality of residence. We also include year fixed effects and interactions between the proportion of voucher schools, the total number of voucher schools, SIMCE scores, and the average characteristics of teachers with the SIMCE scores of students and the level of mother's education.[7] Thus, we have:

$$v_k = E\left(RIF\left(y_k; v\left(F_Y^k\right)\right)\right) = \bar{X}^{k\prime}\beta_k \qquad \text{for } k \in \{0,1,c\}.$$

We can then decompose the gaps in the $\hat{\text{P}}$SU scores between the two types of schools, public and voucher, as follows:

$$(6) \qquad \Delta v = \underbrace{\bar{X}^{1\prime}\left(\hat{\beta}_1 - \hat{\beta}_c\right)}_{\Delta v_s^p} + \underbrace{\left(\bar{X}^1 - \bar{X}^C\right)\hat{\beta}_c}_{\Delta v_s^e} + \underbrace{\left(\bar{X}^c - \bar{X}^0\right)\hat{\beta}_0}_{\Delta v_x^p} + \underbrace{\bar{X}^{c\prime}\left(\hat{\beta}_c - \hat{\beta}_0\right)}_{\Delta v_x^e},$$

where the structure effect is further divided in pure structure $\left(\Delta v_s^p\right)$ and a reweighting error $\left(\Delta v_s^e\right)$. Likewise, the composition effect is separated into the pure composition effect $\left(\Delta v_x^p\right)$ and a specification error $\left(\Delta v_x^e\right)$.

The idea behind the pure composition effects, Δv_x^p, is to capture differences in PSU performance between groups that can only be explained by the fact that the two groups are different. For example, voucher school students score higher in 8th-grade knowledge exams than public school students. Therefore, we expect them also to have an advantage on subsequent college admission test results over public school students. This kind of difference between groups is isolated in the composition effect. On the other hand, the pure structure effects, Δv_s^p, indicate differences in PSU scores due to factors that are more productive for one type of school than the other, leading to better PSU results under the same levels of factors. If, for example, keeping all student characteristics constant, having a more experienced teacher is more advantageous (productive, as measured in PSU score points) for voucher school students, this would be captured in the structure effect.

The two additional estimates from the RIF OB decomposition are the reweighting and the specification error, Δv_s^e and Δv_x^e. The reweighting error comes from the selection of the variables and interaction terms included to compute the counterfactual statistic by estimating P(T = 1|X). It should go to zero in large samples. Of course, a tension exists between a higher Pseudo–R2, a common support, and a perfect prediction, which is undesirable (Firpo, Fortin, and Lemieux 2018). Lastly, the specification error comes from deviations from linearity in the conditional expectation and the fact that $F_Y^c\left(y\right)$ is an approximation, so we should expect this error to be different from zero.[8]

7 This richer specification with many interaction terms is needed to improve the fit in the reweighting process. See also Lemieux (2002).

8 As Firpo, Fortin, and Lemieux (2018) point out, how large the error should be remains an open empirical question.

In the next section, we estimate the classical and RIF Oaxaca-Blinder decompositions to explain the mean gap in the math and Spanish results. Following Firpo, Fortin, and Lemieux (2018). Additionally, by implementing RIF Oaxaca-Blinder decompositions, we break down the observed gap by quantiles of PSU performance. As before, we exclude non-significant components (selective education dummy, education as a top-choice dummy, and teacher's gender) from teacher characteristics to minimize noise in the estimations.

6.1 Mean Differences

Table 4 presents our results for the conventional Oaxaca-Blinder decomposition. Panel A displays the overall gaps. Consistent with the summary statistics, voucher schools have unconditional average advantages in math (column 1) and Spanish (column 2) PSU scores. However, 69% and 45% of these gaps, respectively, are explained by the average differences in the observed characteristics. Panel B shows that eighth-grade test scores and school-level historic PSU take-up are the most important contributors for both subjects. High school teachers' characteristics, a compound of the different variables in this category, play a relatively minor role.

The analysis of the contributors to the unexplained gaps delivers a different story. Panel C of Table 4 suggests that the coefficients associated with the characteristics of the teacher contribute more than 10.9 points to the math PSU score gap (favoring voucher schools), being the largest contributing factor. This result suggests that public and voucher schools produce different outcomes equipped with the same inputs, indicating differences in math PSU's productivity levels across school types. In the case of Spanish, we observe that the coefficient associated with teachers' characteristics is negative. However, its size is much smaller than that of the intercept, which is by far the largest contributor to the unexplained component. This result suggests that the covariates included in the analysis are not compelling enough to comprehend the performance gap in the Spanish PSU. Therefore, the estimates presented here should be interpreted with caution.

Table 5 presents the results of the RIF Oaxaca-Blinder decomposition for the mean distribution, which includes the reweighting scheme (expression (6)). Columns 1 and 2 display the results for math and Spanish, respectively. Panel A shows that the observed characteristics (composition) explain more of the total gaps than the parameters (structure), similar to the results in Panel A of Table 4. The analysis of the Composition effects (Panel B) indicates that 8th-grade test score is the most important contributor, followed by the School's PSU take-up (lagged). Teachers' characteristics contribute with less than one PSU point to closing the gap, which is only statistically significant for Spanish. Panel B also provides insight into the model's goodness-of-fit, which is captured by the specification error. We observe a close-to-zero coefficient that is nonsignificant at conventional levels for math and a large, highly significant positive coefficient for Spanish. This result indicates caution in interpreting results for Spanish, given that, unlike the case for math, the model cannot completely capture all the nuance in the factors that might explain the performance gap across students in both types of schools. Consistent with Panel C in Table 4, the coefficients associated with teachers' characteristics play an essential role in closing the average gaps for math, contributing almost 10 points. The school PSU take-up rate significantly contributes to widening the gap, confirming the role of selection in the PSU discussed above. SES characteristics and eighth-grade test scores contribute only marginally to this component.

These results confirm that pre-high school test scores and the school's (predetermined) college admission test take-up emerge as the most critical differences in characteristics explaining the average PSU gap between public and voucher schools. This result underlines the limits of how much schools can modify and adjust input (e.g., teacher characteristics) to reduce the gaps in a specific cohort. Now, since the differences in coefficients in Equation (6) can be interpreted as proxies for the differential productivity levels of schools, our findings also suggest that, with equal input, voucher schools are better at producing higher PSU scores. This result represents a central challenge for public policies and is consistent with the long-standing evidence documenting the advantages of voucher schools' unconditional test scores.

6.2 Beyond The Mean: Quantile Differences

Decomposing the mean differences in PSU scores between public and voucher schools is informative of the factors driving these gaps and the effectiveness of public initiatives to close them. However, this approach does not reveal the factors that affect students at different levels of the academic performance distribution. For example, for low-performance students, the drivers of gaps between public and voucher schools could differ from those affecting students in the middle or at the top of the distributions. To examine this, we implement the RIF Oaxaca-Blinder decomposition introduced in section 6, which expresses the differences in any distributional statistic as the sum of the structure and composition effects.

Figures 4 and 5 represent the results for math and Spanish, respectively. Given the similarities in their messages and the better goodness-of-fit of the model, we focus mainly on math and discuss any disparities between the two subjects.

Panel A of Figure 4 presents the overall difference in math PSU between public and voucher schools, decomposing it into composition and structure effects in each quantile using the reweighting procedure described in equation (5). The estimated overall difference (red line) is more or less stable across the distributions of the PSU scores, and it slightly decreases as we move up across the quantiles. The range lies in the 20 to 35 interval, with an average of 30 points. The stability suggests that the distribution of student-level scores of voucher schools is mainly shifted to the right relative to public schools. This pattern is consistent with the evidence in Figure 1, which shows the distributions. The positive and increasing composition effects (blue dotted line) indicate that this component increasingly explains the gaps, with the observed characteristics increasingly favoring voucher schools as one moves up in the distributions. We come back to this point below. Finally, the structure effects, depicted by the dashed green line in Panel A, partially compensate for the composition effects, showing a declining slope toward the highest quantiles.

Panel B of the same figure presents the contribution of the different sets of factors to the overall gaps between school types by quantiles. Although socioeconomic characteristics have almost no role in explaining the gaps, teachers are the main drivers in expanding them, particularly in the upper half of the distribution. This pattern is not the case for Spanish, which we discuss later in this section. Additionally, heterogeneity from the pre-high-school test scores (SIMCE) explains between one-third and two-thirds of the gaps across the whole distribution. Finally, consistent with the findings of Tables 2 and 3, accounting for PSU enrollment reduces the advantage of voucher schools, further increasing the gap as we move up in the distribution.

Panels C to E complement the previous results and provide further insight into the relative magnitude of the different effects. Specifically, Panel C shows that most of the composition effects come from the pure explained component; meanwhile, Panel E shows that this is not the case for structure effects, where pure explained and residual effects mostly net out each other. Panels D and F highlight the importance of pre-high-school test scores comes from the composition instead of structure effects while confirming that the school-type-specific estimated coefficients (structure effect) of teachers' characteristics and predetermined PSU take-up rates are essential drivers of disparities. It is interesting to observe that the contribution of teachers to the structure effect is somewhat different between math and Spanish (see Panel F of Figure and 5). For math, we see that this component is crucial in explaining the gap in the upper half of the distribution, with voucher schools being relatively more productive. For Spanish, teachers are more important in explaining the gap in the lower half of the distribution, with voucher schools being relatively more productive. The pattern is much noisier in the upper part of the distribution.

6.3 Complementarities

Figure 2 shows that students from voucher schools outperform public school students on the SIMCE test. This fact holds for both subjects, although the gap is more prominent for math. In this subsection, we analyze how teachers' contribution to explaining the PSU gap changes for students at different baseline performance levels.

We start by analyzing the point estimates of Equation 1 when introducing two interaction terms into the model. The first is the interaction between students' performance on the eighth-grade SIMCE test and their average teachers' performance on the college admissions test. We also control for the interaction between students' performance on the eighth-grade SIMCE test, their teachers' performance on the college admissions test, and the voucher school indicator variable. Table 6 reports these results.

Columns (1) and (4) present the original regression results, including all explanatory variables, in the estimations explaining the math and Spanish PSU scores, equivalent to the results in column (5) in Tables 2 and 3, respectively. Columns (2) and (5) present results when including the interaction term between the average teacher's PSU and the student's past SIMCE math and Spanish scores, respectively. We only observe positive and significant coefficients associated with the interaction term for math. This result suggests that the positive effect of having a teacher with a higher PSU score is amplified when the students have higher pre-high school test scores in the case of math PSU scores. Lastly, columns (3) and (6) include an additional interaction term, multiplying the interaction term by the voucher indicator variable. We observe that the coefficient associated with the interaction between student and teacher performance and the voucher dummy is negative for both tests, although only statistically significant for math. Thus, the amplification of the teachers' PSU effect is smaller for students attending voucher schools and nearly nonexistent for Spanish.

Lastly, we return to quantile gap analysis to explore what would happen with the contribution of teachers if we abstracted the students' baseline performance. We re-estimate a RIF Oaxaca-Blinder decomposition of a new measure of student PSU performance orthogonal to SIMCE scores. We construct this measure by residualizing the PSU scores by their own SIMCE test scores and using this residualized measure as the new dependent variable. The results are graphically presented for quantiles of the distributions in Figure 6, with Panel A showing the results for math and Panel B for Spanish for both the original PSU score (red line, also shown in Panel F of Figures 4 and 5) and the new residualized version (navy line). In both panels of Figure 6, we re-scale the Y-axis to reflect the fraction of the total effect that is explained by teacher structure effects under each of the performance measures since the gap (in level) between vouchers and public schools might change once we decompose it using the residualized PSU measure.

The idea behind Figure 6 is the following. A positive teacher structure effect implies that teachers with the same characteristics have students with a better PSU performance in voucher schools than in public schools, i.e., they are more productive. Suppose there is a positive complementarity, such as teachers being more productive with more prepared students (i.e., higher SIMCE score). In that case, we should expect that the teacher's structure effect decreases once we take out the impact of the SIMCE test score on the PSU performance. The rationale is that students from voucher schools have higher SIMCE scores than students from public schools, especially in math. Then, once we isolate the fact that teachers in voucher schools work with students who are better equipped in terms of performance, the differences in productivity should be smaller. This pattern is what we observe in Panel A of Figure 6 for both the lowest and highest quantiles of math PSU performance. In Panel B, this pattern holds only for the lowest quantiles of Spanish PSU performance. These findings are consistent with Table 6, in which we show a positive complementarity between teachers' PSU performance and students' SIMCE score, but only statistically significant for math.

Finally, it is important to mention one caveat for the evidence shown in Figure 6. The analysis assumes that for each student, their corresponding quantile using PSU performance remains unchanged when using the residualized measure of performance and that for each of those quantiles, voucher students outperform public school students in terms of their SIMCE score. The correlation between students' ranking using the PSU and their residualized measures is about 0.7. Additionally, in Figure 2, we observe that for math and Spanish, the SIMCE test score distribution is shifted to the right for voucher school students, compared to public school students. This pattern suggests that the assumption above holds on average but is imperfect. Therefore, we should be cautious when interpreting the figure.

7. Conclusion

Our comprehensive analysis of the factors influencing student achievement in Chile's college admission test (PSU) provides insights into the intricate relationships between schools, teachers, and student outcomes. Our unique dataset includes matched teacher-student data, incorporating detailed information about teachers' performance in high-stakes college admission assessments.[9]

The main results, presented in Tables 2 and 3, demonstrate the persistent impact of the type of high school on PSU performance, even after controlling for student socioeconomic characteristics, teacher attributes, and eighth-grade test scores. Although factors such as teacher's subject-specific PSU performance, experience, and holding an education degree are highly relevant to predicting student success, the school-type gap remains substantial, underscoring the complexity of factors contributing to educational disparities. This fact indicates that addressing inequalities in teacher quality alone may not be sufficient to bridge the gap in student achievement across different school types.

The Oaxaca-Blinder decomposition, as outlined in Table 4, provides a nuanced understanding of the components that directly contribute to the performance gap between public and voucher schools. Although observed characteristics and eighth-grade test scores explain most of the average gaps, teachers' characteristics contribute substantially to the unexplained portion, especially in math. This evidence suggests inherent productivity differences between school types.

The reweighted Oaxaca-Blinder decomposition results (Table 5) emphasize the importance of pre-high school test scores and historic PSU take-up rates in understanding the average gaps. Thus, given equal input, voucher schools might exhibit higher productivity levels, representing a challenge for policy interventions seeking to alleviate educational inequalities. The RIF Oaxaca-Blinder approach enables us to look beyond the mean differences (Figures 4 and 5), adding another layer of complexity. While socioeconomic characteristics have minimal impact on the gaps across quantiles, teachers' characteristics become more pronounced in expanding the gaps, particularly at the high-end distribution of scores, emphasizing the influence of teacher characteristics on the performance of high-achieving students. These results highlight the need for targeted interventions that address the diverse needs of students at different achievement levels.

Finally, we explore complementarities between teacher and student baseline performance in predicting PSU scores and explaining the voucher-public

[9] Due to insufficient data, we cannot determine whether there is a difference in attendance to "preuniversitarios" for students attending both types of schools. Additionally, we cannot isolate the benefits of having access to "preuniversitarios" on PSU performance from our results.

school gap. We find that students with high SIMCE scores perceive a boost in their performance when paired with a high PSU-performing teacher and that the increase is twice as large among public school students. We examine complementarities in teacher-student interactions using the RIF Oaxaca-Blinder quantile methodology, considering all teacher characteristics but using a PSU measure orthogonal to the baseline performance level as measured by the SIMCE. We find evidence that teacher contributions to the gap disappear once baseline test scores are accounted for, underscoring the importance of understanding the interplay between prior achievement and teacher effectiveness in formulating effective educational policies.

Our study advances the understanding of educational disparities in Chile by revealing the persistent gap between school types, the influential role of teachers, and the presence of complementarities. The results emphasize the imperative for comprehensive interventions that combine targeted teacher training, early educational investments, and efforts to address existing productivity differences between school types. Policymakers must strategically combine these factors to foster a more equitable and effective educational system. Future research should deepen the exploration of these dimensions, proposing innovative strategies to improve educational outcomes and inform evidence-based policies.

8. TABLE

TABLE 1

PUBLIC VS. VOUCHER SCHOOLS: SUMMARY STATISTICS

	Public				Voucher			
	Math		Spanish		Math		Spanish	
	Mean (1)	**Std. Dev. (2)**	**Mean (3)**	**Std. Dev. (4)**	**Mean (5)**	**Std. Dev. (6)**	**Mean (7)**	**Std. Dev. (8)**
Panel A. Student-level information								
PSU results	484.67	102.84	477.95	103.95	508.03	99.00	506.10	98.76
Simce 8th grade	-0.13	1.03	-0.08	1.01	0.09	0.97	0.04	0.98
Fraction taking PSU (lagged)	0.72	0.20	0.72	0.20	0.78	0.20	0.78	0.21
Female	0.55	0.50	0.57	0.49	0.55	0.50	0.55	0.50
Mothers with a technical degree	0.07	0.26	0.07	0.26	0.10	0.30	0.10	0.31
Mothers with a professional degree	0.30	0.46	0.30	0.46	0.28	0.45	0.28	0.45
Mothers with a postgraduate degree	0.04	0.18	0.03	0.18	0.05	0.22	0.05	0.22
Year Cohort	2,018.75	2.31	2,018.76	2.33	2,018.68	2.41	2,018.68	2.40
Panel B. Teacher's information								
Teacher PSU	-0.03	1.02	-0.10	1.02	0.06	0.98	0.11	0.97
Teacher GPA in PSU scale (NEM)	0.07	0.99	0.06	1.04	0.01	0.98	0.03	0.97
Graduated from PUC or UCH	0.08	0.25	0.06	0.23	0.10	0.28	0.09	0.28
Teacher has education degree	0.87	0.32	0.97	0.15	0.86	0.32	0.95	0.20
Education as top 3 choice in ranking	0.72	0.41	0.82	0.35	0.75	0.39	0.81	0.35
Years of teaching experience	3.55	3.39	3.80	3.35	3.29	2.95	3.28	2.71
Age	31.45	4.00	31.41	4.08	31.22	3.59	30.92	3.38
Female	0.53	0.46	0.75	0.40	0.52	0.45	0.74	0.40
Number of Observations	155,615		140,947		273,358		274,368	

Note: Panel A displays summary statistics of students in the sample, divided by the type of school they attend, public or voucher schools, and the subject of interest, math, and Spanish. Panel B displays similar summary statistics by type of school and subject but for the average characteristics of teachers teaching math or Spanish to students during high school in Panel A. The total number of students considered is 537,119, while for 307,169 of them, we have information on their performance and teachers' characteristics in both subjects.

TABLE 2

PUBLIC VS. VOUCHER SCHOOLS: AVERAGE MATH PSU GAP REGRESSION ANALYSIS

	Math PSU				
	(1)	(2)	(3)	(4)	(5)
Voucher	23.261***	21.951***	22.005***	9.442***	7.469***
	(4.773)	(4.709)	(4.326)	(2.244)	(2.165)
Teacher's Math PSU			9.058***	3.632***	3.390***
			(1.537)	(0.831)	(0.802)
Teacher's GPA			0.465	0.381	0.419
			(1.985)	(1.083)	(1.029)
Teacher has Selective Education			4.899	0.924	-0.098
			(5.317)	(2.980)	(2.889)
Teacher holds Education Degree			31.772***	13.351***	12.524***
			(3.313)	(1.652)	(1.553)
Education as Top 3 choice			-2.773	-1.610	-1.631
			(3.152)	(1.585)	(1.521)
Years of Teaching Experience			2.647***	1.141***	1.079***
			(0.550)	(0.285)	(0.271)
Teacher's Age			-1.726***	-0.826***	-0.810***
			(0.422)	(0.221)	(0.213)
Fraction Female Teachers			-0.662	-0.453	-0.549
			(3.474)	(1.833)	(1.735)
8th grade Simce (Math)				62.190***	61.145***
				(0.816)	(0.737)
School's PSU Takeup (lagged)					41.695***
					(3.039)
Number of Observations	428,973	428,973	428,973	428,973	428,973
R-squared	0.025	0.032	0.054	0.413	0.418
Year FE	✓	✓	✓	✓	✓
SES Controls	✓	✓	✓	✓	✓

Note: The table presents the point estimates obtained from different versions of equation 1 using Spanish PSU as the dependent variable. The sample includes PSU takers covering the period 2013-2021. Year FE includes year-specific fixed effects for test-taking years. SES Controls include indicator variables for student gender and three indicator variables for maternal education categories: technical, undergraduate, and post-graduate degrees, with high school or less being the omitted category. Standard errors in parentheses clustered at the school level. * p < 0.1, ** p < 0.05,*** p < 0.01

TABLE 3

PUBLIC VS. VOUCHER SCHOOLS: AVERAGE SPANISH PSU GAP REGRESSION ANALYSIS

			Spanish PSU		
	(1)	(2)	(3)	(4)	(5)
Voucher	28.305***	26.615***	25.666***	18.753***	16.064***
	(4.544)	(4.462)	(4.385)	(2.613)	(2.495)
Teacher's Spanish PSU			8.062***	4.781***	4.282***
			(1.532)	(0.930)	(0.878)
Teacher's GPA			-0.406	-0.341	-0.548
			(1.751)	(1.092)	(1.023)
Teacher has Selective Education			14.569	10.837**	9.124*
			(9.178)	(5.489)	(5.048)
Teacher holds Education Degree			24.545***	13.936***	12.958***
			(4.354)	(2.613)	(2.366)
Education as Top 3 choice			1.957	0.678	0.416
			(3.892)	(2.363)	(2.203)
Years of Teaching Experience			2.645***	1.267***	1.129***
			(0.489)	(0.295)	(0.282)
Teacher's Age			-1.556***	-0.705***	-0.630***
			(0.459)	(0.252)	(0.232)
Fraction Female Teachers			-2.005	-1.823	-1.300
			(2.887)	(1.791)	(1.701)
8th grade Simce (Spanish)				64.664***	63.753***
				(0.525)	(0.461)
School's PSU Takeup (lagged)					50.849***
					(3.602)
Number of Observations	415,315	415,315	415,315	415,315	415,315
R-squared	0.019	0.029	0.043	0.433	0.441
Year FE	✓	✓	✓	✓	✓
SES Controls	✓	✓		✓	✓

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table presents the point estimates obtained from different versions of equation 1 using Spanish PSU as the dependent variable. The sample includes PSU takers covering the period 2013-2021. Year FE includes year-specific fixed effects for test-taking years. SES Controls include indicator variables for student gender and three indicator variables for maternal education categories: technical, undergraduate, and post-graduate degrees, with high school or less being the omitted category. Standard errors in parentheses clustered at the school level. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$

TABLE 4

MEAN DECOMPOSITION OAXACA-BLINDER WITHOUT REWEIGHTING

	(1)	(2)
	Math	Spanish
A. Overall Gap		
(a) Average PSU in Voucher Schools	508.028***	506.102***
	(0.189)	(0.189)
(b) Average PSU in Public Schools	484.667***	477.950***
	(0.261)	(0.277)
Gap in favor of Voucher Schools ((a)-(b))	23.361***	28.153***
	(0.322)	(0.335)
Total Explained	16.065***	12.686***
	(0.215)	(0.242)
Total Unexplained	7.296***	15.467***
	(0.258)	(0.266)
B. Contributions to the Explained Gap		
Student's SES	0.545***	1.549***
	(0.035)	(0.050)
Teacher	0.168***	0.617***
	(0.031)	(0.062)
Student's Simce	12.815***	7.863***
	(0.194)	(0.215)
Lagged School PSU Takeup	2.537***	2.657***
	(0.063)	(0.069)
C. Contributions to the Unexplained Gap		
Student's SES	-1.764***	-2.526***
	(0.353)	(0.360)
Teacher's characteristics	10.854***	-11.426***
	(2.951)	(3.076)
Student's Simce	0.278***	-0.124***
	(0.021)	(0.012)
Lagged School PSU Takeup	-5.224***	-13.093***
	(0.966)	(0.996)
Intercept	3.152	42.636***
	(3.186)	(3.298)
Number of Observations	428,973	415,315

Note: The table presents the Oaxaca-Blinder decomposition for PSU scores in math and Spanish by type of school (public or voucher). The sample includes PSU takers covering the period 2013-2021. Standard errors in parenthesis clustered at the school level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

TABLE 5

MEAN DECOMPOSITION OAXACA-BLINDER WITH REWEIGHTING

	Math	Spanish
A. Decomposition		
Total gap	23.361***	28.153***
	(0.31)	(0.317)
Composition	15.401***	12.237***
	(0.294)	(0.324)
Structure	8.282***	8.694***
	(0.3)	(0.308)
B. Contributions of Xs (composition)		
Student's SES	0.615***	1.639***
	(0.041)	(0.069)
Teacher's characteristics	-0.059	0.874***
	(0.036)	(0.057)
8th grade test score (Simce)	14.233***	9.101***
	(0.283)	(0.298)
School's PSU takeup (lagged)	0.611***	0.622***
	(0.039)	(0.042)
Specification error	-0.547*	7.476***
	(0.314)	(0.36)
C. Contributions of β s (structure)		
Student's SES	0.145	0.002
	(0.49)	(0.492)
Teacher's characteristics	9.937***	0.321
	(3.36)	(4.077)
8th grade test score (Simce)	0.051	-0.226***
	(0.032)	(0.029)
School's PSU takeup (lagged)	-10.17***	-19.241***
	(1.063)	(1.106)
Intercept	8.32**	27.838***
	(3.698)	(4.245)
Reweighting error	0.225	-0.254
	(0.314)	(0.248)
D. Total (composition + structure)		
Student's SES	0.759	1.64***
	(0.494)	(0.49)
Teacher's characteristics	9.878***	1.195
	(3.362)	(4.075)
8th grade test score (Simce)	14.284***	8.876***
	(0.285)	(0.29)
School's PSU takeup (lagged)	-9.559***	-18.618***
	(1.07)	(1.105)

Note: The table presents the RIF Oaxaca-Blinder decomposition for the mean PSU score (math and Spanish) by type of school (public or voucher). The sample includes PSU takers covering the period 2013-2021. Bootstrapped standard errors over the entire procedure (100 replications) were used to compute the p-values and are presented in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 6

PUBLIC VS. VOUCHER SCHOOLS: STUDENT-TEACHER INTERACTIONS AS A MECHANISMS REGRESSION ANALYSIS INTERACTION

		Math PSU			Spanish PSU	
	(1)	(2)	(3)	(4)	(5)	(6)
Voucher	7.469***	7.647***	7.873***	16.064***	16.092***	16.200***
	(2.165)	(2.150)	(2.153)	(2.495)	(2.501)	(2.519)
Teacher's PSU	3.390***	3.442***	3.551***	4.282***	4.266***	4.299***
	(0.802)	(0.800)	(0.809)	(0.878)	(0.871)	(0.883)
Teacher's GPA	0.419	0.456	0.465	-0.548	-0.547	-0.546
	(1.029)	(1.033)	(1.039)	(1.023)	(1.023)	(1.026)
Teacher has Selective Education	-0.098	-0.511	-0.637	9.124*	9.099*	9.051*
	(2.889)	(2.819)	(2.805)	(5.048)	(5.060)	(5.088)
Education Degree	12.524***	12.337***	12.360***	12.958***	12.945***	12.935***
	(1.553)	(1.539)	(1.539)	(2.366)	(2.362)	(2.364)
Education as Top 3 choice	-1.631	-1.624	-1.611	0.416	0.421	0.446
	(1.521)	(1.512)	(1.511)	(2.203)	(2.204)	(2.201)
Years of Teaching Experience	1.079***	1.072***	1.074***	1.129***	1.131***	1.133***
	(0.271)	(0.273)	(0.272)	(0.282)	(0.281)	(0.281)
Teacher's Age	-0.810***	-0.797***	-0.800***	-0.630***	-0.632***	-0.634***
	(0.213)	(0.212)	(0.212)	(0.232)	(0.231)	(0.230)

Fraction Female Teachers	-0.549	-0.620	-0.690	-1.300	-1.300	-1.308
	(1.735)	(1.755)	(1.758)	(1.701)	(1.701)	(1.702)
Student's Simce	61.145***	61.106***	61.127***	63.753***	63.753***	63.812***
	(0.737)	(0.725)	(0.722)	(0.461)	(0.461)	(0.455)
School's PSU Takeup (lagged)	41.695***	41.723***	41.574***	50.849***	50.845***	50.807***
	(3.039)	(3.026)	(3.020)	(3.602)	(3.603)	(3.607)
Teacher PSU × Student SIMCE		2.780*** (0.415)	4.070*** (0.719)		0.373 (0.365)	1.264 (0.795)
Teacher PSU × Student SIMCE × Voucher			-2.169** (0.860)			-1.454 (0.889)
Number of Observations	428973	428973	428973	415315	415315	415315
R-squared	0.418	0.419	0.419	0.441	0.441	0.441
Year FE	✓	✓	✓	✓	✓	✓
SES Controls	✓	✓	✓	✓	✓	✓

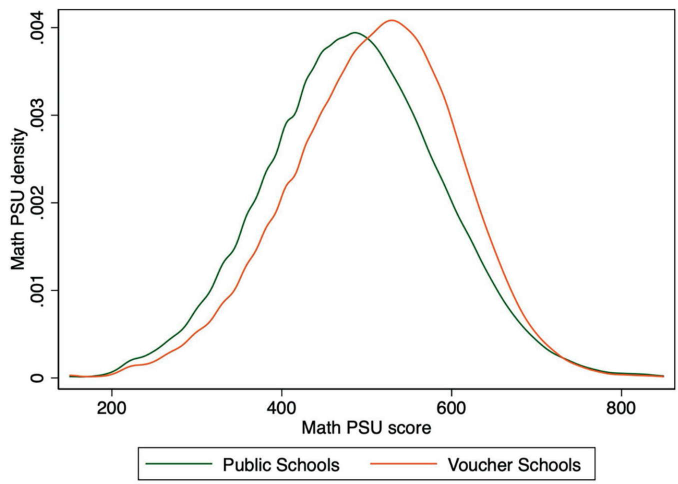
Note: The table presents the point estimates obtained from different versions of equation 1 using Spanish PSU as the dependent variable. The sample includes PSU takers covering the period 2013-2021. Year FE includes year-specific fixed effects for test-taking years. SES Controls include indicator variables for student gender and three indicator variables for maternal education categories: technical, undergraduate, and post-graduate degrees, with high school or less being the omitted category. Standard errors in parentheses clustered at the school level. * $p < 0.1$, ** $p < 0.05$,*** $p < 0.01$

9. Figures

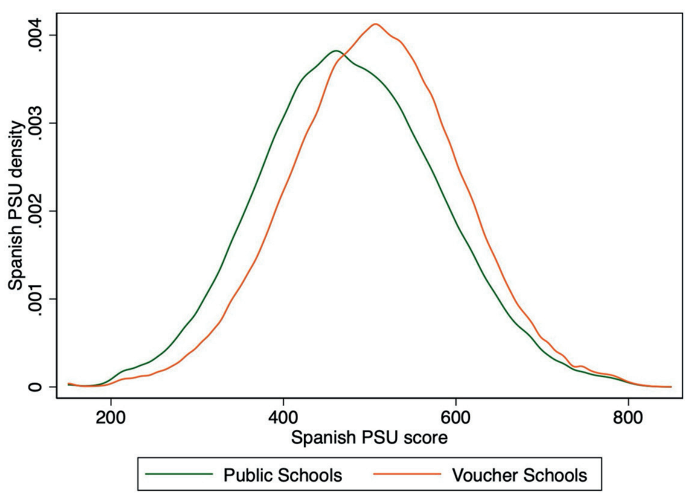
FIGURE 1

DISTRIBUTION OF STUDENT-LEVEL PSU SCORES BY SCHOOL TYPE

A. MATH PSU



B. SPANISH PSU

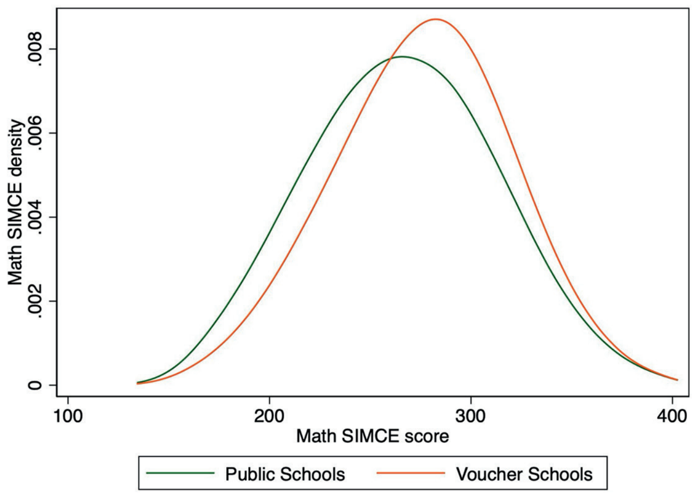


Note: Panel A (B) displays the distribution of math (Spanish) PSU scores computed from our sample of 428,973 (415,315) test takers between the years 2013 and 2021.

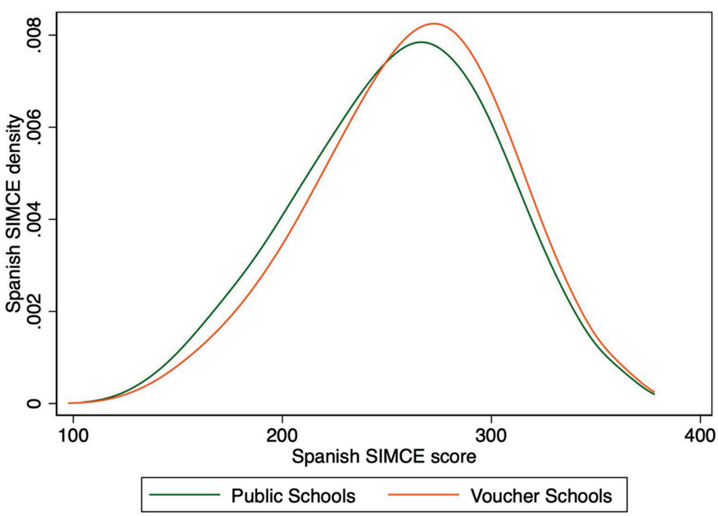
FIGURE 2

DISTRIBUTION OF STUDENT-LEVEL SIMCE SCORES BY SCHOOL TYPE

A. MATH SIMCE



B. SPANISH SIMCE

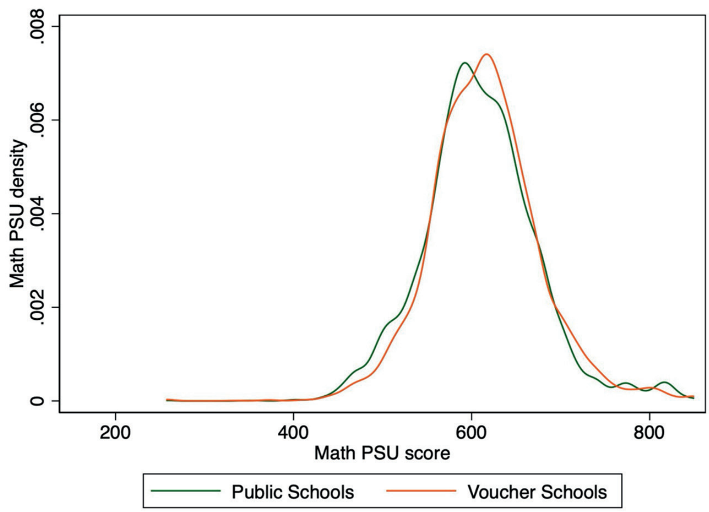


Note: Subfigure a (b) displays the distribution of PSU scores at the student level for math (Spanish) computed from our sample of 428,973 (415,315) test takers between the years 2013 and 2021.

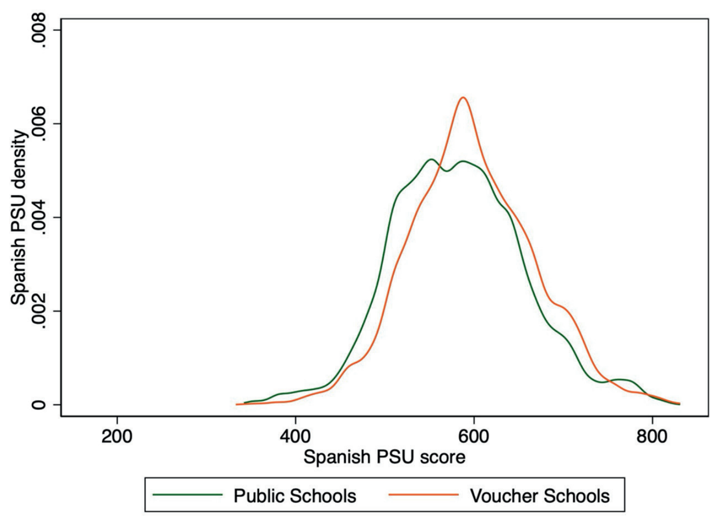
FIGURE 3

DISTRIBUTION OF AVERAGE TEACHERS' PSU SCORES BY SCHOOL TYPE

A. MATH PSU



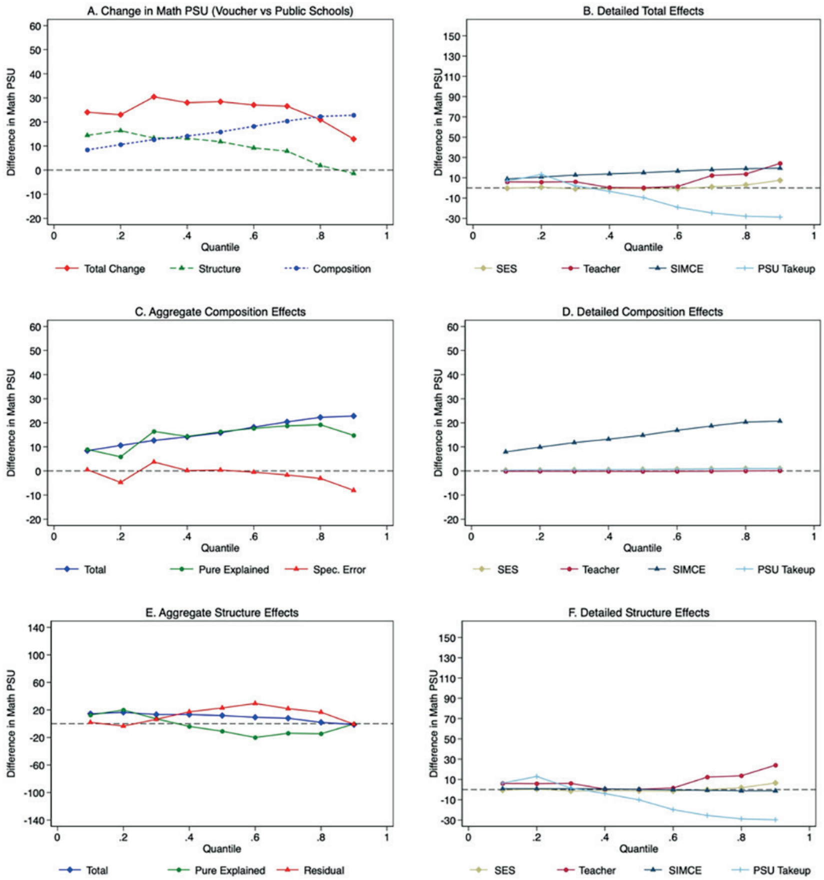
B. SPANISH PSU



Note: Subfigure a (b) displays the distribution of average teacher math (Spanish) PSU scores computed from our sample teaching 9,938 (9,574) school cohorts of student test takers between the years 2013 and 2021.

FIGURE 4

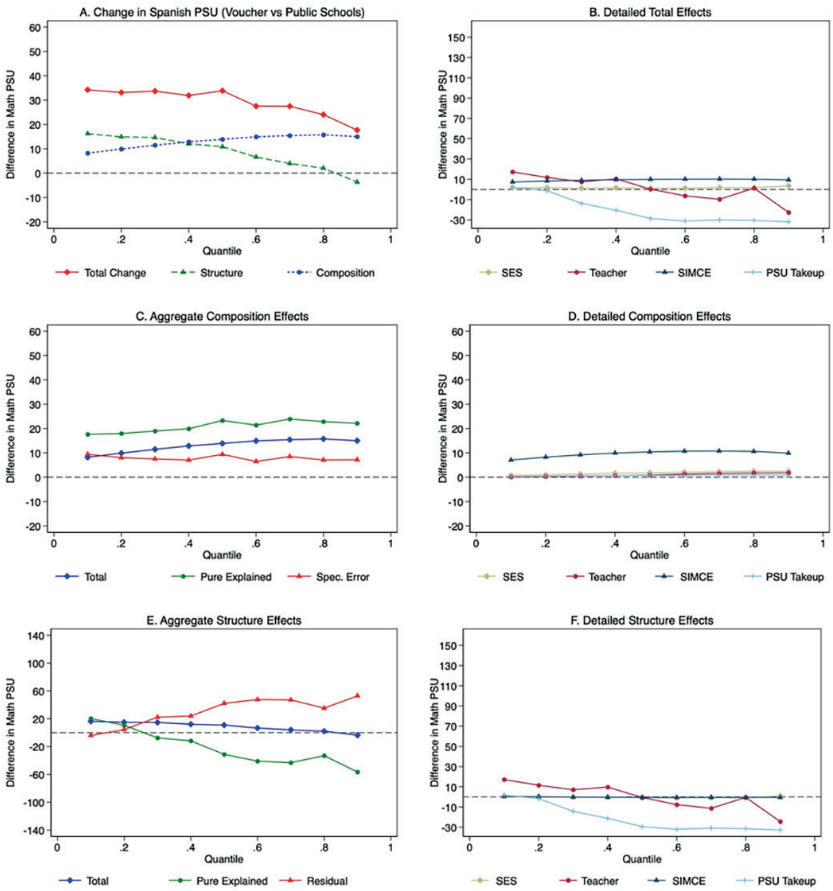
GAPS ACROSS THE MATH PSU DISTRIBUTION: TOTAL, COMPOSITION, AND STRUCTURE EFFECTS



Note: The figure presents the RIF Oaxaca-Blinder decomposition for quantiles of PSU score in math by type of school (public or voucher). Panel A displays the total gap and the portion of it that is explained by the composition and structure effect. Panel B decomposes the total gap by the contribution of each group of variables included in the analysis. Panel C displays the total composition effect, the portion of it that is purely explained, and the specification error. Panel D decomposes the purely explained composition effect by the contribution of each group of variables included in the analysis. Panel E displays the total structure effect and the portion of it that is purely explained and residual. Panel F displays decomposes the purely explained structure effect by the contribution of each group of variables included in the analysis. The sample includes 428,973 students covering the period between 2013 and 2021. See Sections 4 and 5 for a formal discussion.

FIGURE 5

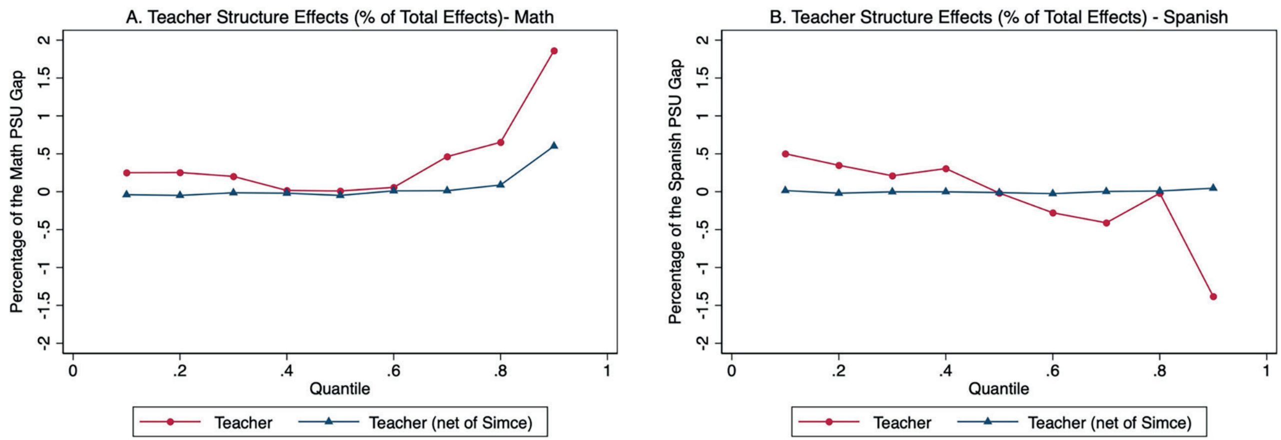
GAPS ACROSS THE SPANISH PSU DISTRIBUTION: TOTAL, COMPOSITION AND STRUCTURE EFFECTS



Note: The figure presents the RIF Oaxaca-Blinder decomposition for quantiles of PSU score in Spanish by type of school (public or voucher). Panel A displays the total gap and the portion of it that is explained by the composition and structure effect. Panel B decomposes the total gap by the contribution of each group of variables included in the analysis. Panel C displays the total composition effect, the purely explained portion, and the specification error. Panel D decomposes the purely explained composition effect by the contribution of each group of variables included in the analysis. Panel E displays the total structure effect and the portion of it that is purely explained and residual. Panel F displays the decomposition of the purely explained structure effect by the contribution of each group of variables included in the analysis. The sample includes 415,315 students covering the period between 2013 and 2021. See Sections 4 and 5 for a formal discussion.

FIGURE 6

TEACHERS CONTRIBUTION TO STRUCTURE EFFECT: TOTAL AND RESIDUALIZED OF PREVIOUS TEST SCORE



Note: The figure presents the contribution of teachers' characteristics to the purely explained structure effect on the PSU score gap and a residualized measure of the PSU score gap. This residualized measure corresponds to the PSU score minus the effect of the SIMCE test score in the eighth grade. Panel A displays the results for math. Panel B displays the results for Spanish. In both panels, the red line (Teacher) corresponds to the contribution of teachers' variables to the explained structure effect on the PSU gap by quantiles. In contrast, the blue line (Teacher res.) corresponds to the contribution of teachers' variables to the explained structure effect on the residualized PSU gap by quantiles. The sample includes 428,973 individuals for math and 415,315 for Spanish, covering 2013 to 2021. See Section 5.2 for a formal discussion.

References

Aaronson, Daniel, Lisa Barrow, and William Sander. (2007). "Teachers and student achievement in the Chicago public high schools." *Journal of labor Economics* 25 (1): 95–135.

Barrios, Andrés, and Marc Riudavets. (2021). "Teacher value-added and gender gaps in educational outcomes." Available at SSRN 3856935.

Behrman, Jere R, Michela M Tincani, Petra E Todd, and Kenneth I Wolpin. (2016). "Teacher quality in public and private schools under a voucher system: The case of Chile." *Journal of Labor Economics* 34 (2): 319–362.

Bellei, Cristian. (2005). "The private-public school controversy: The case of Chile." *In Conference on Mobilizing the Private Sector for Public Education*, October, vol. 5. 6. sn.

Bordón, Paola, Catalina Canals, and Alejandra Mizala. (2020). "The gender gap in college major choice in Chile." *Economics of Education Review* 77:102011.

Bravo, David, Sankar Mukhopadhyay, and Petra E Todd. (2010). "Effects of school reform on education and labor market performance: Evidence from Chile's universal voucher system." *Quantitative economics* 1 (1): 47–95.

Canales, Andrea, and Luis Maldonado. (2018). "Teacher quality and student achievement in Chile: Linking teachers' contribution and observable characteristics." *International Journal of Educational Development* 60:33–50.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014a. "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates." *American Economic Review* 104 (9): 2593–2632.

. 2014b. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American Economic Review* 104 (9): 2633–2679.

Contreras, Dante. (2002). "Vouchers, school choice and the access to higher education." School Choice and the Access to Higher Education (June 2002).

DiNardo, John, Nicole M Fortin, and Thomas Lemieux. (1996). "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64, no. 5 (September): 1001–1044.

Elacqua, Gregory. (2012). "The impact of school choice and public policy on segregation: Evidence from Chile." *International Journal of Educational Development* 32 (3): 444–453.

Firpo, Sergio P, Nicole M Fortin, and Thomas Lemieux. (2018). "Decomposing wage distributions using recentered influence function regressions." *Econometrics* 6 (2): 28.

García-Echalar, Andrés, Sebastián Poblete, and Tomás Rau. (2023). "Teacher Value-added and the Test Score Gender Gap." *Manuscript*.

Gilraine, Michael, and Nolan G Pope. (2021). Making teaching last: Long-run value-added. Technical report. *National Bureau of Economic Research*.

Hanushek, Eric A. (2011). “The economic value of higher teacher quality.” *Economics of Education review* 30 (3): 466–479.

Heckman, James J. (2000). “Policies to foster human capital.” *Research in economics* 54 (1): 3–56.

Iturra, Victor, and Mauricio Gallardo. (2022). “Schools, circumstances and inequality of opportunities in Chile.” *International Journal of Educational Development* 95:102668.

Jackson, C Kirabo. (2018). “What do test scores miss? The importance of teacher effects on non–test score outcomes.” *Journal of Political Economy* 126 (5): 2072–2107.

Jackson, C Kirabo, Rucker C Johnson, and Claudia Persico. (2016). “The effects of school spending on educational and economic outcomes: Evidence from school finance reforms.” *The Quarterly Journal of Economics* 131 (1): 157–218.

Kutscher, Macarena, Shanjukta Nath, and Sergio Urzúa. (2023). “Centralized admission systems and school segregation: Evidence from a national reform.” *Journal of Public Economics* 221:104863.

Lemieux, Thomas. (2002). “Decomposing changes in wage distributions: a unified approach.” *Canadian Journal of Economics/Revue canadienne d’économique* 35 (4): 646–688.

Mizala, Alejandra, and Pilar Romaguera. (2000). “School performance and choice: the Chilean experience.” *Journal of human Resources*, 392–417.

Montaño, Sebastián, Catalina Morales, Cristina Riquelme, and Sergio Urzúa. (2023). “Do high-performing teachers produce high-performing students?” *Manuscript*.

Neilson, Christopher, Sebastian Gallegos, Franco Calle, and Mohit Karnani. (2022). “Screening and recruiting talent at teacher colleges using pre-college academic achievement.”

Petek, Nathan, and Nolan G Pope. (2023). “The multidimensional impact of teachers on students.” *Journal of Political Economy* 131 (4): 1057–1107.

Rios-Avila, Fernando. (2019). “Recentered influence functions in Stata: Methods for analyzing the determinants of poverty and inequality.” *Levy Economics Institute*, Working Paper 927.

Rivkin, Steven G, Eric A Hanushek, and John F Kain. (2005). “Teachers, schools, and academic achievement.” *Econometrica* 73 (2): 417–458.

Rockoff, Jonah E. (2004). “The impact of individual teachers on student achievement: Evidence from panel data.” *American economic review* 94 (2): 247–252.

Todd, Petra E, and Kenneth I Wolpin. (2003). “On the specification and estimation of the production function for cognitive achievement.” *The Economic Journal* 113 (485): F3–F33.

Toledo, Gabriela, and Juan Pablo Valenzuela. (2015). “Over-estimating the effects of teacher attributes on school performance in the Chilean education system.” *Estudios de Economıa* 42(1).

Wei, Hua, Tracey Hembry, Daniel L Murphy, and Yuanyuan McBride. (2012). Value-added models in the evaluation of teacher effectiveness: A comparison of models and outcomes.