Over-estimating the effects of teacher attributes on school performance in the Chilean education system*

Sobrestimación del impacto de atributos docentes en el desempeño escolar en el sistema educativo chileno

Gabriela Toledo Román*
Juan Pablo Valenzuela**

Abstract

This article investigates the biases involved in estimating the effects of teacher attributes on school's performance. The study was performed for Chilean educational system, where student distribution is differentiated by schools and teachers are not randomly assigned to them. Findings showed that teacher attributes which favored learning appeared more frequently alongside higher socioeconomic status students. When correcting the bias, results showed that the effects of teacher attributes have been overestimated for the vast majority of characteristics. Nonetheless, attributes such as teaching experience, being a woman, having short-term specific professional training, and having a greater curriculum coverage continued to have positive impacts on the performance of 4th grade students.

Key words: *Teacher attributes, sorting, school segregation, school performance.*

JEL Classification: *A2, I2.*

Resumen

Este artículo investiga el sesgo presente en las estimaciones del efecto de los atributos docentes en el desempeño escolar en el sistema educativo chileno, donde estudiantes y profesores no son aleatoriamente asignados a las escuelas. Los resultados señalan que los estratos socioeconómicos altos presentan con mayor frecuencia atributos docentes que promueven el aprendizaje. Cuando se corrige por este sesgo, los resultados mostraron una sobrestimación del impacto de la

* Funding from PIA-CONICYT Basal Funds for Centers of Excellence Project BF0003 is gratefully acknowledged. The paper benefited from helpful comments of anonymous referees.

** Pontificia Universidad Católica de Chile. gabitoledo@gmail.com

** Centro Investigación Avanzada en Educación (CIAE) [Center for Advanced Research in Education (CARE)] and Economics Department, Universidad de Chile.

gran mayoría de atributos docentes. Sin embargo, atributos como la experiencia docente, ser mujer, con capacitación profesional específica y tener una mayor cobertura curricular, continúan siendo atributos que impactan positivamente en el desempeño escolar.

Palabras clave: *Atributos de profesores, segregación escolar, desempeño escolar.*

Clasificación JEL: *A2, I2.*

1. Introduction

When studying the impact that a teacher can have on a school's academic performance, one of the main difficulties is in knowing which teacher attributes significantly improve learning (Goldhaber and Brewer 1997, Rivkin, Hanushek and Kain 2005, Sanders and Rivers 1996). Currently, two principal methodological challenges exist for addressing this issue. The first is based on data of teacher attributes that effectively reflect quality (Goldhaber 2008, Goldhaber and Anthony 2007), but the available information is, in general, generic such as demographic variables, educational level, and experience, among others. This is not the exception, to include characteristics like motivation and abilities will be a great improvement in measuring teacher's impact on school performance. The second challenge considers the distribution of teacher attributes in relation to student characteristics, given that when the distribution of teacher and school attributes is not random, the estimations of teacher impact will be biased (Clotfelter, Ladd and Vidgor 2006).

The present study addressed the second methodological challenge by first describing and documenting the extent of segregation between teacher attributes and student and school characteristics in Chile, and secondly, by estimating the impact of teacher attributes on academic performance through a methodology that corrects for nonrandom assignment. This work presents novel evidence for sorting within schools by linking teacher and student attributes in the classroom and complements the existing literature on sorting and teacher effectiveness, which has mainly been investigated in developed countries.

Additionally, Chile, due to its particularity, is an interesting study case. The Chilean educational system allows the coexistence of public schools, subsidized and private schools, offering an almost universal coverage and a variety of educational project. Private subsidized schools in Chile, which account for 53% of school enrolment, are allowed to charge a supplemental fee of up to USD$160 per month in addition to the received public subsidy, but discounting an increasing percentage of co-payment from public subsidy[1]. Close to 50% of

1 Discounts from monthly public voucher for student are 0% of the public voucher for monthly charges lesser than USD$20, an additional 10% for charges between USD$20-USD$40, 20% for charges between US$40-USD$80, and 35% for the charge which is between USD$80-USD$160, charges up to this threshold imply do not receive the public subsidy.

private subsidized schools charge mandatory fees to parents which represent the 70% of enrollment of this type of schools. Thus, private subsidized schools with family co-payments can affect their social composition between schools because magnitude of charges operates as an entry barrier for families with lower financial capacity (McEwan *et al.*, 2008). On the other hand, public schools receive transferences of resources from local and regional governments which are not available for private subsidized schools.

The system evidences a strong correlation of school performance with socioeconomic status (SES) and there is comparatively high socioeconomic school segregation (Valenzuela *et al.*, 2009, OECD, 2013). Moreover, Hsieh & Urquiola, 2006 states that the introduction of voucher system in Chile produced a large-scale segmentation of the educational system.

Previous studies note that teachers with attributes associated to higher educational quality are concentrated in schools and classes with students of higher SES and ability (Lankford, Loeb and Wyckoff 2002, Clotfelter, Ladd and Vidgor 2006). Considering that the socioeconomic segregation of the Chilean school system is among the highest in the world[2], it is worth examining how this relates to the segregation of teacher attributes among students in this country (OECD 2009, Valenzuela, Bellei and De los Ríos 2009), as is in fact the case with infrastructure, teachers and teacher aid support (García-Huidobro 2007, González, Mizala and Romaguera 2002, Hsieh and Urquiola 2006, Mizala and Torche 2012).

Learning inequality can also occur within educational institutions. In the Chilean education system, parent's choice is the motor for competition, and schools have the motivation to pursue better results and to advertise them (Mizala and Urquiola 2007). This pressure for better outcomes can lead to internal sorting of students according to cognitive and non-cognitive capabilities. In the presence of this situation, it is likely that better teachers will be assigned to students who already have a better learning potential (Dupriez, Dumay and Vause 2008).[3]

To document the extent of nonrandom assignment of teacher and student attributes between and within schools, this study used data from the Education Quality Measurement System (EQMS)[4] on school learning together with complementary information on student and teacher characteristics. The aim of this research was to estimate the level of sorting among teachers and students and, for the first time, to conduct estimates for Chile with partially restored random conditions by comparing the effects of sorting and the effects of teacher

2 The relationship between schooling segregation and student achievement has been widely documented internationally (Hanushek and Woessmann, 2011)

3 For an international review of matching students between and within institutions, see Dupriez, Dumay and Vause (2008).

4 The Education Quality Measurement System (SIMCE in Spanish) is a standardized, annual test for all 4th grade students in language and mathematics.

attributes on school performance. This paper presents the theoretical framework, methodology, results and discussion related to this research referred.

2. Theoretical Framework

There is consensus about the impact that teachers and schools have on moderating the effects of SES on learning. Consistently, many authors have found that after student and peer attributes, the teachers' quality is the main variable that drives student academic achievement (Rivkin, Hanushek and Kain 2005, Goldhaber and Brewer 1997, Sanders and Rivers 1996, Rockoff 2004, Aaronson, Barrow and Sanders 2007). Also, on a larger scale, the effectiveness of teachers is considered to be one of the key factors explaining the success of educational systems around the world (Barber and Mourshed 2007, OECD 2010).

However, the leadership role played by teachers in the process of learning is not always highlighted when identifying observable attributes that foster student learning (Goldhaber 2008, Goldhaber and Anthony 2007, Goldhaber and Brewer 1997). In what follows we present the theoretical framework used in the current study, based on education production functions and, in particular, on teacher quality research. Subsequently we review the available literature regarding the assignment of teachers to students and provide evidence from Chilean research.

Learning and school achievement are typically measured by standardized tests or measures of internal efficiency, such as dropout rate, grade repetition, and opportune secondary school completion. Identifying a production function for educational outcomes is the prevailing methodology used to estimate the determinants of school performance and the magnitude of their influence, such as in the following:

$$(1) \qquad A_{ijt} = \alpha A_{ijt-1} + \beta X_{it} + \gamma S_{jt} + \omega T_{ijt} + \varepsilon_{ijt}.$$

The education production function shown in Equation (1) incorporates the following principal educational variables: A_{ijt} is the test result of the student *i* in the school *j* at the time *t*. As controls it includes the variable of the result lagged by one period (A_{ijt-1}) as a measure of student ability, the family level *X*, school level *S*, and teachers. This specification is usually applied to cross-sectional data without information on the prior or systematic abilities possessed by each individual student, because of the difficulties of having access to longitudinal data.

Within education production functions, two approaches are generally used with relation to teacher quality, both of which are related to student learning. The **direct approach** to teacher quality is defined as the value added by a teacher to a student. In this case, the effect of the teacher is generally estimated with models using records of student academic performance. This approach avoids the problem of relating observable variables to performance but reduces the possibility for conclusions and policy recommendations based on teacher attributes. The literature, which applies this approach, focuses on capturing the importance of the teacher in academic performance and not necessarily on what are the determinants of teacher quality. Several studies concerning this approach have obtained positive, but relatively modest, effects (Rockoff 2004, Rivkin, Hanushek and Kain 2005, Aaronson, Barrow and Sanders 2007). The **indirect approach** involves quality as measured by observable teacher attributes, such as having a professional degree, continuous training, professional experience, teaching methods, and standardized teacher evaluations. Results of this approach are mixed given that there is a weak link between the quality of teachers and measurable variables or attributes (Rivkin, Hanushek and Kain 2005). Consequently, the results of these studies are sensitive to the context of the study and the methodology used. For example, Jacob and Lefgren (2004) conclude that there is no evidence that the continuous training of teachers increases teacher productivity in terms of students' results. In contrast, Clotfelter, Ladd and Vidgor (2007) find that the certifications awarded in professional development programs do affect school performance in the United States, but that teachers possessing these qualifications are unequally distributed in schools, impacting negatively on vulnerable, low-income students enrolled in schools with a high percentage of non-white students.

In a meta-analysis of studies conducted in the United States, Goldhaber (2008) analyzes the impact of five observable teacher aspects on the academic performance of their students.[5] He concluded that professional experience has a nonlinear positive effect during the first five years of a teacher's career, and that scores obtained in standardized teacher evaluations also have a systematic, positive effect on school learning. Other studies show that obtaining a postgraduate certificate has a small but significant negative effect on academic performance (Clotfelter, Ladd and Vidgor 2007), whereas in schools where the teacher specializes in one subject, there is a positive but modest effect (Monk, King and Ehrenberg 1994, Goldhaber and Brewer 1997).

Finally, recently available longitudinal data has allowed researchers to mix direct and indirect approaches. Longitudinal data allows for controlling of selection bias by including the fixed effects of the school, teachers, and students, thus leading to more reliable estimations (Harris and Sass, 2011). Findings support the positive and moderate effect of experience on achievement, although this effect does not last beyond the first few years of experience. This approach is much more desirable for estimating teacher effectiveness, but, regrettably, in Chile as well as in most developing countries, it is difficult to obtain longitudinal data.

Thus, although various studies recognize the role that teachers play in the resulting performance of the school in standardized tests, it is difficult to determine which attributes lead to increases in performance. What they do find is that teachers and their attributes are not randomly distributed with respect to the students they teach, thus possibly increasing inequality in children learning opportunities (Clotfelter, Ladd and Vidgor 2006).

---

[5] In his review, Goldhaber (2008) considers professional experience, academic degree, primary subject taught, teaching certifications, and demonstrated academic capacity (i.e., results in standardized tests, such as PSU scores (SAT equivalent) for acceptance to Chilean universities). Higher teacher quality is measured through improved student performance on standardized tests in relation to early teaching experience.

Non-random Assignment of Teachers to Students: Sorting Effect

An active process of nonrandom assignment or sorting is often observed among teachers, students, and schools in educational systems, in which two main sources of sorting have been identified. The first involves pairing between teachers and schools, that is, "sorting between schools," where the selection of teachers by schools is carried out according to the attributes offered by both parties. Schools offer a specific salary, certain labor conditions, and particular student qualities, and teachers offer their attributes and expectations. The second source is "sorting within schools," which occurs when teachers within a school are assigned to a specific class of a grade level, thus purposefully matching teacher and student attributes. In both cases, the effects of sorting may be positive or negative.

Types and Effects of Sorting

Sorting between/within schools is positive when the best teachers work at schools or in classes with students with the most favorable characteristics for learning. It is negative if these teachers are at schools or in classes with the least favorable conditions and attributes for learning.

It has been suggested that in the presence of sorting, the parameters of teacher impact on school achievements will be biased in both the direct and indirect approaches referred to above. Positive sorting typically occurs in socioeconomically segregated school systems, both between and within schools (Boyd *et al.*, 2008, Clotfelter, Ladd and Vidgor 2006). This is partly because schools with the best working conditions, in terms of salary, school environment, and prestige tend to hire teachers with better personal and academic attributes as well as better teaching effectiveness. This situation is further explained by school administrators or superintendents who believe that global performance is increased when assigning the best teachers to the best students. The Program for International Student Assessment (PISA) scores indicate that 12.9% of 15-year-old students in OECD countries are grouped in courses according to their academic capacity, and Chile ranks above average in this category with 30.3%, only below Luxembourg, the Netherlands, and Switzerland (OECD 2010).

Teacher characteristics and Sorting in Chile

The more heterogeneous the quality of teachers' training more relevant will be the sorting between teachers and students in the school system. In Chile, differences in quality of teacher education are remarkable, for example, of those teachers enrolled in education by 2012 a 29% did not take the university selection test (PSU) and only a quarter of them were among the 30% of students with the highest scores in the test (MINEDUC, 2013). In this context, there has recently been established a test, INICIA[6], to measure and control the quality of teaching new teachers. Data generated from this test indicates that new primary teachers reached on average only 50% of correct answers in the test of disciplinary

[6] Test INICIA is applied to recently graduated teachers to assess their performance.

knowledge and for those with a major in mathematics only 33% (MINEDUC, 2009), and that these results have remained over time.

Moreover, there is a systematic sorting process of new teachers among schools by their students´ socioeconomics and performance conditions. Those teachers coming from more selective colleges, with higher scores in test INICIA and others better observable attributes are much more concentrated among primary and secondary schools with better average performance or higher social status, and this sorting process is accentuated when teachers gain more experience in the educational system (Cabezas *et al.*, 2011; Meckes and Bascopé, 2010; Ortúzar *et al.*, 2012; Paredes *et al.*, 2014; Rivero, 2012; Ruffinelli and Guerrero, 2009).

Taking into account of this teacher-school sorting process, there are previous Chilean studies on the effects that teachers have on a student's school performance that handled biases produced by sorting between schools. For example, Bravo *et al.* (2008) estimated the effects of various teacher certification processes on student standardized test scores. To correct for the effect of sorting, they introduced a variable built on past evaluations that summarizes the student's school performance. In so doing, they found that teacher qualifications may identify the most efficient teachers in terms of standardized testing results. Also, Lara, Mizala and Repetto (2010) estimated the impact of teacher attributes on student academic performance. They corrected for the effect of nonrandom sorting between teachers and students through including the school's average performance, and they handled sorting within schools by sampling schools that have only one class per grade level. The researchers conclude that children who have female teachers perform better, that having a qualified teacher can increases student performance, and that teacher experience is nonlinear, that is, its increasing effect on student academic performance is only observed during the first years of teaching. Finally, Paredes (2014) states that female teacher have a positive impact on girls, and that these results is due to role model not to teacher bias effect.

To conclude, Chilean studies discussing the link between teacher characteristics and academic performance have found that sorting produces an important impact on the educational system. The present study carries previous studies further by correcting for both types of sorting, between and within schools, thus showing the impact that sorting has on the unequal distribution of educational results in Chile.

3. Data and Methodology

3.1. Data sources

The following two data sources were used for the present analyses: the EQMS scores of 4th grade students from 2005 and 2006 (*SIMCE*), and data from the National Teachers Registry for the same years. Using data from two years allowed for a greater number of analyzed observations and reduced the cohort effect, which could bias outcomes. The EQMS provided a rich database of standardized test results that were inter-temporally comparable. This study used results from the reading and mathematics tests, as well as background data from the teachers, students, and students' families. The sample consisted of only urban schools given that most rural schools are small, and, as such, EQMS results from these schools are not entirely reliable (Kane and Staiger, 2002).

To characterize sorting, nine teacher attributes were grouped in the following four main categories: demographic (age and gender); formal education (type of educational institution attended and level of first degree); professional development (years of experience, certificates or credentials that reflect recent improvement, and whether the teacher specialized in a single subject); and pedagogical management (amount of curriculum coverage achieved in language and math for the given year). All of these attributes were relevant to educational performance based on comparative evidence. However, it is necessary to declare the absence of quantitative information on motivation and ability of teachers, which are very important attributes when making career path decisions. The type of professional degree achieved by the teacher was obtained from the National Teacher Registry, and the amount of curriculum coverage achieved was constructed based on a series of questions in the EQMS questionnaire for teachers, which asked them how many of the various compulsory topics from a specific subject (e.g. language and mathematics) were covered in their class.

TABLE 1
ATTRIBUTES AND CATEGORIES ANALYZED OF TEACHERS

Variable type	Variable	Categories
i) Demographic	1) Gender	Male, female
	2) Age	Categorical
ii) Initial training	3) Type of teacher education institution attended	University, Normal School, Professional Institute, No professional studies.
	4) Level of teacher qualification	Primary education, Secondary education, Differential education, Preschool education, No qualification, Qualification in others areas.
iii) Teacher professional development	5) Teacher years of experience	0-2 years of experience, 2-5 years of experience, 6-12 years of experience, 13-24 years of experience and more than 25 years of experience;
	6) Professional development credentials	No diploma or certification, Short course, Diploma, Master or Doctorate.
	7) Subject specialization	Teacher teaches only language, teacher teaches only math.
iv) Curricular coverage in the year of assessment.	8) Language Curriculum Coverage	More than one Std. Dev. from average, within one Std. Dev. from average, less than one Std. Dev. from average.
	9) Math Curriculum Coverage	More than one Std. Dev. from average, within one Std. Dev. from average, less than one Std. Dev. from average.

3.2. Methodology

The methodology of this research was based on an adaptation of Clotfelter, Ladd and Vidgor (2006) and addressed the problem of sorting between and within schools. First, matching of teacher and student attributes both between and within schools was documented. Within schools sorting was documented for first time for the Chilean school system. Subsequently, estimates were made for the impact of teacher attributes on academic performance and included strategies designed to isolate the problems related to this type of assignment.

Matching Teacher and Student Attributes

Sorting between schools

To investigate whether there was sorting between schools, the distribution of student characteristics was analyzed among the categories of a specific teacher attribute. An F-test was used to establish the significance of these relations, the null hypothesis being that the distribution of student characteristics was similar along the different categories of a determined teacher attribute.

For example, to assess the educational level of the student's mother we used the question "Did the student's mother finish secondary education?" In turn this variable was examined along categories of teacher attributes, as for example, years of teacher experience, (<2 years; 2-5 years; 6-12 years; 13-24 years; and >25 years). The analyses were repeated with other student attributes, such as "Did the student's family receive social aid?", "Does the student belong to an ethnic group?", and "Does the student's family belong to the first socioeconomic quintile?".

Sorting within schools

To examine whether there were differences in the distribution between classes based on student attributes, a χ^2 test was conducted for student samples enrolled in schools with two or more sections per grade level. Differences were compared among students from the 4th grade.[7] If the randomness test was rejected, this indicated that there was sorting within the school, and so these institutions were subsequently eliminated from the sample. For different sections within the same grade, the test was conducted considering the following student attributes: mother's education, relatives receiving help from social services, number of books at home (used as a proxy of cultural capital, McEwan, 2001), and the family income quintile. As in sorting between schools, to assess the sorting within schools in its real dimension it would be necessary to have information on motivation and abilities also for the students.[8]

---

7 The null hypothesis of the test was that the characteristics of the students were randomly distributed between the samples.

8 A proxy for this variable is the history of grade repetition for the student. However, despite the relevance of showing prior student academic ability, this variable has not been systematically recorded in the EQMS tests. We performed the chi-square test and 15%

In addition, there was conducted an analysis of the distribution of teachers' attributes between sections within the same grade. For this purpose, a ranking of student attributes was developed among sections, with better attributes indicating a more privileged class. Subsequently, the distribution of some teacher attributes among these equivalent sections was compared.

Impact of teacher attributes on academic performance

The methodology designed by Clotfelter, Ladd and Vidgor (2006) was adapted to correct for positive matches resulting from the existing assignments of teachers to students in the Chilean school system. First, a baseline consisting of correlations between teacher attributes and schools' EQMS scores was estimated. Subsequently, the initial stage of estimation included detailed variables at the student level. In Clotfelter, Ladd and Vidgor (2006), the student's individual score from the previous testing period was included as a control parameter, but unfortunately this information is not available in Chile.[9] For this reason, for each year, only control variables related to students and their family environment were included. The individual variables included in the estimation were as follows: gender, preschool attendance, family income quintile, mother's educational level, and whether the family received help from the government. In addition, proxies for educational resources were included, such as internet availability and the number of books at home. Finally, peer effect variables were also incorporated and included the average years of maternal schooling for the entire class and classroom-level characteristics.

The second stage of analysis corrected for non-observable characteristics of schools, included a set of school fixed effect dummies that affect student performance. The sample was reduced to schools with more than two sections per grade level so as to allow for variability in teacher attributes within the school. In the third stage, corrections were made for the nonrandom assignment of teachers between different sections of the same grade level at the school. For this purpose, the sample was reduced according to the procedure described below so as only to analyze those schools where student attributes were not significantly different between equivalent sections.

As previously explained, the methodology implied working with different samples. The sample used for the first and second stages comprised 179,813 students. Surveys were conducted in urban schools with two or more sections for the 4th grade, and complete student information for all the variables considered was obtained. As shown in Table 2 these students represented 40.9% of the entire student population in the two years studied. In the third stage, however, only schools free of sorting were used.

of schools rejected randomness of distribution in this attribute, but it is only available for one year.

9 A proxy for this variable is the history of grade repetition for the student. However, despite the relevance of showing prior student academic ability, this variable has not been systematically recorded in the EQMS tests.

TABLE 2
SAMPLE DISTRIBUTION OF STUDENTS BY TYPE OF SCHOOL AND THE AVAILABILITY OF INFORMATION

Type of school	Information	2005	2006	Total	Percent of total students
Schools with 2 or more classrooms	Incomplete	87,547	76,049	163,596	37.2%
	Complete	86,512	93,301	179,813	40.9%
	Total	174,059	169,350	343,409	78.1%
Total of schools (one or more classrooms)	Total	221,575	217,970	439,545	100.0%

4. Results: School Sorting and Observable Teacher Attributes which Affect Student Performance

4.1. Assignment of Teachers to Students in Chilean Schools

Below, we provide a brief description of the characteristics of the teachers and students included in the study.

Fourth grade teacher attributes

Table 3 shows the characteristics of 4th grade teachers in urban Chilean schools. These attributes were ordered according to the different groups of schools. Analysis was first done with the whole sample (Column 1), followed by analysis of schools with only one section per grade (Column 2), with two or more sections per grade (Column 3), and without sorting within the school (Column 4).

It was possible to observe that only in the group of schools with one section per grade, teacher attributes were significantly different in comparison to the total number of schools. This finding may be because these schools tend to be smaller and of lower socioeconomic level, thus representing a particular and specific sub-sample of the teacher population.

In the general sample, over 65% of the teachers were between 40- and 60-years-old. In schools with only one section per grade most teachers had lower qualification levels and attributes. They were also less likely to have participated in additional professional development courses and covered less of the mandated curriculum in their classes.

Results from this sample indicated that 36% of the teachers had 25 or more years of professional experience, while just over half (57%) had not participated in professional development activities during the last two years. A large proportion of these teachers (84%) was university graduates, and certified for elementary teaching.

TABLE 3
DISTRIBUTION OF TEACHERS' CHARACTERISTICS BY TYPE OF SCHOOL FOR BOTH YEARS

	Number of classrooms in fourth degree	Total of schools	Only 1 classroom	2 or more classrooms	2 or more classrooms without sorting within
Number of teachers		8,828	2,025	6,802	2,911
Gender	Percent female	85.68%	87.68%	87.68%	85.92%*
Age	Less than 30 years	9.22%	12.05%*	8.38%	8.42%
	30-39 years	19.23%	22.12%*	18.38%	18.28%
	40-49 years	33.44%	32.59%	33.70%	33.56%
	50-59 years	31.15%	26.91%*	32.42%	32.53%
	More than 60 years	6.96%	6.32%	7.13%	7.21%
Teacher years of experience	0-1 years	4.96%	7.01%*	4.35%	4.77%
	2-5 years	13.08%	16.74%*	12.00%	12.30%
	6-12 years	17.50%	20.15%*	16.72%	16.04%
	13-24 years	28.09%	26.02%*	28.71%	29.75%
	More than 25 years	36.36%	30.07%*	38.22%	37.14%
Teacher Qualifications	None	57.03%	57.83%	56.79%	56.44%
	Degree or Diploma	34.88%	34.02%	35.14%	34.80%
	Certificates or Diplomas	5.27%	5.14%	5.31%	5.74%
	Master or Doctorate	2.82%	3.01%	2.76%	3.02%
Type of teacher education institution	University	83.88%	82.52%	84.28%	83.89%
	Normal School	10.60%	10.72%	10.57%	11.03%
	Professional Institute	4.77%	5.43%	4.57%	4.53%
	None	0.75%	1.33%*	0.57%	0.55%
Level of teacher education qualification	Primary education	94.26%	92.94%*	94.65%	94.6%
	Secondary education	0.37%	0.54%	0.32%	0.27%
	Special education	0.99%	1.19%	0.93%	0.89%
	Preschool education	2.6%	3.5%*	2.4%	2.34%
	No degree	0.7%	0.9%*	0.6%	0.76%
	Degree in others areas	1.09%	0.99%	1.12%	1.17%
Language Curriculum Coverage (from the average)	More than a Std. Dev.	18.71%	15.65%*	19.61%	19.51%
	Within a Std. Dev.	64.85%	65.28%	64.73%	64.89%
	Less than a Std. Dev.	16.44%	19.06%*	15.66%	15.60%
Math Curriculum Coverage	More than a Std. Dev.	19.11%	15.26%*	20.24%	21.26%
	Within a Std. Dev.	63.19%	64.10%*	62.92%	62.21%
	Less than a Std. Dev.	17.71%	20.64%*	16.83%	16.52%
Subject specialization	Teacher teaches only language	1.28%	2.32%*	0.97%	1.17%
	Teacher teaches only math	0.67%	1.04%*	0.56%	0.45%*

* Difference respect to total establishments is significant at the 5 percent significance level.

Description of students enrolled in 4th grade

The student characteristics variables used included family SES (mother's education and income quintile), cultural capital (number of books at home), internet availability, and student demographics (age, gender, and preschool education). However, student level variables on ethnicity and grade repetition were available for only one of the years analyzed and could not be used in

TABLE 4
CHARACTERISTICS OF STUDENTS OF URBAN SCHOOLS SAMPLE WITH 2 OR MORE CLASSROOMS PER GRADE FOR YEARS 2005 AND 2006

		Total	Free of sorting
Observations		179,813	76,754
Gender	Percent of female	50.00%	49.5%*
Social Program	Are you beneficiary of social programs?	17.80%	17.70%
Education of father	No secondary education	35.10%	33.4%*
	Secondary education	36.90%	36.80%
	Technical education	13.20%	14.0%*
	University education	11.20%	12.2%*
Education of mother	No secondary education	36.00%	34.3%*
	Secondary education	39.90%	40.0%*
	Technical education	15.60%	16.5%*
	University education	8.50%	9.2%*
Internet	Do you have internet?	21.10%	22.5%*
Books at home	Average books	30.83	31.98*
Income quintile	1st quintile	16.50%	15.5%*
	2nd quintile	22.60%	21.9%*
	3rd quintile	20.40%	20.10%
	4th quintile	19.80%	20.00%
	5th quintile (richest)	20.70%	22.4%*
Peer effect	Classroom average of mother education	11.5	11.7*
School status	Public education	46.70%	42.0%*
	Subsidized education	46.40%	50.0%*
	Private education	6.90%	8.0%*
Class size	Number of students per classroom	34.7	34.5*
School size	Number of students per level	190.3	186.8*
Etnicity[a]	Do you belong to any indigenous people?	8.40%	8.3%*
Individual performance[b]	Has the student repeated a grade?	6.20%	5.9%*

a. Variable is available only for 2006 year.
b Variable is only available for 2005 year.
* Differences in means are statistically significant at < .01 across schools that making sorting and those who do not.

teacher effectiveness estimations. Nonetheless, as ethnicity is related to spatial and socioeconomic segregation it was used to show the sorting between schools. Comparison of sample groups showed that, on average, the students enrolled in schools free of within-school sorting had a slightly higher SES and cultural level than peers from schools with sorting. Nonetheless, nothing conclusive could be said about within sorting and school's SES because the lack of data disaggregation in EQMS may not capture sorting in higher SES schools (see sorting within section). A more detailed description of within school segregation in the Chilean educational system can be found in Treviño *et al.* (2014).

Evidence of Sorting in the Chilean School System

As indicated in note 1, in Chile there are three types of schools: public-municipal, private subsidized and non-subsidized private schools. These schools are very different in their financing and flexibility in teacher hiring decisions, and they usually show a high correlation between SES and average student performance. In the light of these distinctions it was assumed that if the market was capable of identifying the characteristics of high-quality teachers, such teachers would then be selected more frequently by schools that offered better monetary and non-monetary labor conditions, which in turn would be schools with higher student SES.

Sorting between schools

To describe sorting between schools, the distribution of student characteristics among categories of teachers' attributes was analyzed using an F-test.

Table 5 quantifies the magnitude of sorting in urban schools with two or more sections per grade level and at least 15 students in each course.[10] The rows represent teacher characteristics and the columns represent the proportion of student characteristics related to such teacher attributes. The values shown in the table reflect the weighted averages of student characteristics as related to the proportion of teachers at each school with the specific attribute. This weighted average better represented the effect of sorting and showed that, in a school with a greater proportion of teachers with a given attribute, a greater proportion of students had favorable characteristics. For example, the cell of the first row and column indicates that female teachers are less likely to teach students whose mothers did not finish their secondary education (38.5% v. 41.2% for male teachers). Despite the difference being relatively small, the F-test indicated a statistical significance, suggesting that male teachers tend to teach in more vulnerable school environments than female teachers.

Teaching experience was the attribute which most affected teacher sorting across student socioeconomic characteristics. In general, students taught by teachers with over 25 years of experience were considerably more disadvantaged than those with less experienced teachers. Thus teachers with the longest experi-

10 This selection was performed in order to avoid an additional bias in the estimates, given the measurement error of school performance due to transitory shocks (Kane & Staiger 2002).

TABLE 5
EVIDENCE OF SORTING BETWEEN SCHOOLS: CHARACTERISTICS OF STUDENTS AND THEIR FAMILIES BY ATTRIBUTE OF PROFESSOR IN THE URBAN SAMPLE OF 4TH GRADE SIMCE 2005

		Mother without secondary education	Beneficiary of social programs	belongs to an ethnic group	belongs first socio-economic quintile
Gender	Female	38.5%	18.7%	8.8%	18.5%
	Male	41.2%	19.0%	9.2%	19.6%
	Reject F-test	yes***	yes***	yes***	yes***
Teachers' years of experience	0-1 years	33.97%	19.69%	8.71%	17.96%
	2-5 years	35.29%	18.87%	8.17%	18.12%
	6-12 years	35.72%	18.97%	8.18%	18.52%
	13-24 years	37.32%	19.24%	9.24%	20.75%
	More than 25 years	48.45%	21.70%	9.84%	29.02%
	Reject F-test	yes***	yes***	yes***	yes***
Type of Teacher Education	University	40.47%	20.17%	9.15%	22.97%
	Professional Institute	41.29%	18.28%	8.83%	22.11%
	Normal School	43.64%	20.36%	9.66%	24.11%
	None	44.51%	18.44%	10.06%	24.12%
	Reject F-test	yes***	yes***	yes***	yes***
Level of teacher qualifications	Primary education	41.18%	20.08%	9.21%	23.32%
	Secondary education	20.67%	21.33%	5.80%	11.79%
	Special education	34.98%	17.12%	6.03%	16.42%
	Preschool education	37.76%	20.36%	9.45%	21.18%
	None	41.69%	20.55%	10.87%	20.01%
	Degree in others areas	41.89%	16.42%	10.08%	19.79%
	Reject F-test	yes***	not	yes***	yes***
Professional Development certificates or credentials	None	40.97%	20.13%	9.02%	23.24%
	Courses	41.95%	20.09%	9.61%	23.67%
	1 year diploma	36.24%	20.80%	9.02%	19.24%
	Master or Doctorate	33.40%	18.16%	7.48%	18.87%
	Reject F-test	yes***	yes**	yes***	yes***
Language Curriculum Coverage (Standardized)	More than a Std. Dev	40.44%	21.03%	9.47%	20.32%
	Within a Std. Dev	41.40%	21.37%	8.98%	20.45%
	Less than a Std. Dev	44.42%	22.29%	9.42%	22.45%
	Reject F-test	yes***	yes***	yes***	yes***
Math Curriculum Coverage (Standardized)	More than a Std. Dev	37.66%	19.03%	9.20%	21.45%
	Within a Std. Dev	40.59%	19.99%	9.10%	22.77%
	Less than a Std. Dev	45.64%	21.76%	9.58%	26.05%
	Reject F-test	yes***	yes***	yes***	yes***

* F-test of null hypothesis that student's characteristics are equals across teacher attributes.
** F-Test P-Value< 0.05.
*** F-Test P-value < 0.01.

ence were teaching classes with 29% of the students coming from the poorest quintile, compared to the teachers with five years or less experience that were teaching classes with only 18% of their students coming from the poorest quintile. This may be caused by the school system itself, as only public schools pay based on seniority and thus retain the more experienced teachers, and because of the legal constraints involved in firing a teacher. It is also possible that the systematic migration of students from municipal to private-subsidized schools, a more flexible sector as far as hiring is concerned, could be leading to the hiring of less experienced teachers that in turn earn lower salaries.

There were also substantial differences in reading and mathematics curriculum coverage, reflecting positive sorting across schools. However, these variables also depend on student attributes, as smarter classes generally learn faster and vice versa. Thus, the distribution of this variable may be endogenous, that is, the less favorable the students' attributes are the harder it is to advance in curriculum coverage even in the absence of teacher sorting. The relevance of this variable is that it reflects qualitative characteristics more than any other variable in the analyzed databases. In any case, the eventual bias of an endogenous relationship should substantially decrease when correcting for socioeconomic characteristics and fixed effects. However, this variable could still be endogenous with students' characteristics like abilities, or behavior, so results must be analyzed considering these data restrictions.

The type of teacher education institution from which teachers graduated also showed significant differences. Teachers with a university degree were far less likely to teach the most vulnerable students. However this was not the case with teachers who, graduated from a non-university tertiary institution (known as Professional Institutes). This last finding suggests that differences in the type of teacher education institution are not reflected in the educational labor market. As for the type of teacher specialization, most teachers were qualified for elementary education. Teachers qualified as secondary, pre-school or special education constituted a minority in the sample of teachers, but they were more likely to teach less vulnerable students. The same holds true for teachers with postgraduate studies (diplomas, master's and Ph.D. degrees).

In conclusion, the positive sorting of teachers in relation to students was analyzed, and findings showed that male, older, and/or less trained teachers were most likely to work in more vulnerable schools. Teachers in vulnerable schools also covered a smaller share of established curriculums as compared to other teachers.

Sorting within schools

To determine the existence of sorting within urban schools with two or more 4th grade sections, the characteristics of students were examined to confirm random distribution between the sections. For this purpose, a χ^2 test was performed using the null hypothesis that this distribution was similar across sections of the same school. In schools with cases of rejection, there was a possible nonrandom sorting within the school for the analyzed student characteristics, and therefore, these schools were removed from the sample so as to eliminate a potential bias in the parameters of teacher attributes.

TABLE 6
SUMMARY OF CHI-SQUARED OF STUDENT'S RANDOM ASSIGNMENT ACROSS SCHOOLS IN SCHOOLS WITH TWO CLASSES IN FOURTH GRADE

Number of tests failed	2005	2006	Total	Percent of schools (%)
0 de 4	1,102	1,112	2,214	65.70
1 de 4	492	456	948	28.10
2 de 4	90	90	180	5.30
3 de 4	14	15	29	0.90
4 de 4	0	1	1	0.00
Total of schools	1,698	1,674	3,372	100
Total of students	85,384	95,044	180,428	100

Four socioeconomic student characteristics were considered when identifying sorting within schools.[11] Three of these characteristics coincided with those described in the analysis of sorting between schools. However, in order to show sorting within the school, the variable representing student's ethnic origin was substituted for the number of books at home as being more related with student's cultural capital, and with more variability within the school. It is worthy to note that even if this variable shows more variability within school than ethnicity is far from being the best proxy of cultural capital heterogeneity and could lead to some under-estimation of sorting in higher SES schools.[12] Table 6 shows the distribution of schools according to the number of rejected tests.

The schools selected for the estimation were those without student sorting within the school for both years analyzed. The results indicated that the Chilean school system has a high level of segregation within schools given that 34.3% of applicable schools rejected at least one test.[13] This result moreover indicates that segregation within schools exists even in the early years of schooling. For over one third of these schools, it may be possible that students were distributed in a way directly or indirectly linked to educational performance.

11 For an analysis of within school segregation in the Chilean educational system including ability variables is found in Treviño *et al*. (2014)

12 In higher SES schools every child may have more than 200 hundred books –the highest category– so there is no possible within sorting.

13 However, the nonrandom distribution of students within institutions could be greater than the estimated value. This conclusion is based on the lack of data for grade repetition for both years of the sample, which is a critical factor in explaining school performance and is strongly linked to matching within schools. In fact, a partial estimation, considering only the data from 2005, indicated that 15.2% of institutions showed a significant difference in the composition of children repeating grades between its classes. This situation is quite critical because grade repetition is associated with a standard deviation (SD) up to 55% (27.4 points) lower in the EQMS results after controlling for other individual and aggregate socioeconomic characteristics, as well as for teacher and institution attributes.

Accounting for teacher-student sorting within schools: a ranking approach

To illustrate the existence of teacher sorting within schools, we explore within the sample in which at least one of the randomness tests was rejected. The class with the most favorable learning conditions was identified and then plotted in relation to teacher attributes. To identify the class with the best characteristics within each school, a ranking of the classes was conducted in which all of the characteristics from the randomness test were considered. After creating the ranking, a probability distribution was performed for the teacher's attributes for both groups, where the first group corresponded to sections ranked in first place and the second group represented sections ranked in second place.[14]

Teacher experience was analyzed as it proved to be the most prevalent matching attribute between schools (see Table 5). As shown in Figure 1, it is possible to observe that classes ranked first were more likely to have teachers with 20 or more years of experience, while those ranked second were more likely to have teachers with fewer than 10 years of experience. This result shows that teacher experience is an attribute negatively sorted between schools –as can be observed in table 6– and positively sorted within schools. The main hypothesis for this apparent contradiction is that public system pays higher wages to more experienced teachers –so the negative sorting between teacher experience and students characteristics is policy driven, instead of a consequence of sorting driven by the market– which pays higher for quality attributes. On the other hand, in within sorting, better classes have a more experienced teacher, so a positive sorting can be found. This fact emphasizes the importance of controlling for both effects in estimating the impact of teachers' experience on academic performance. Also, the class ranked first had, on average, better curriculum coverage in mathematics than those ranked second, but the distribution is not superior in the entire range, as it can be seen in Figure 1. When analyzing the rest of the attributes of the teachers, it is possible to conclude that for schools that rejected randomness, teacher characteristics show no significant differences between the classes.

Finally, we analyzed school performance between classes in schools that did "sorting within". To do this, we used only those schools in which the randomness tests for students' attributes distribution were rejected, and also only the schools that have only two classes by level. Figure 2 shows how students' performance in classes ranked first had a higher density for better results in math and language standardized tests.

This finding characterizes that schools' sorting had consequences on results. This difference is relevant because it is occurring in equal conditions in non-observable schools' characteristics, the same school principal and policies. If there is an impact of sorting within schools on school performance probably it worked through student's characteristics, peer effect and through teacher-class sorting.

---

14 The analysis also included institutions with three or more classes, but final analysis included the first two classes of the ranking.

FIGURE 1
TEACHER-CLASSES SORTING

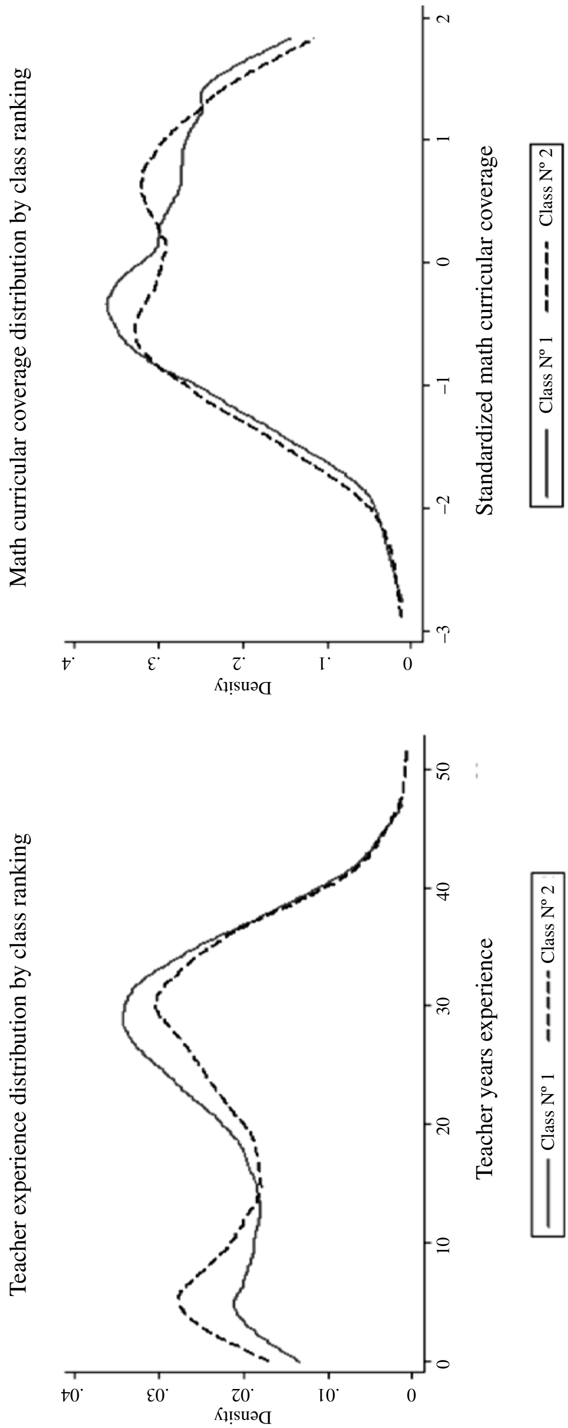
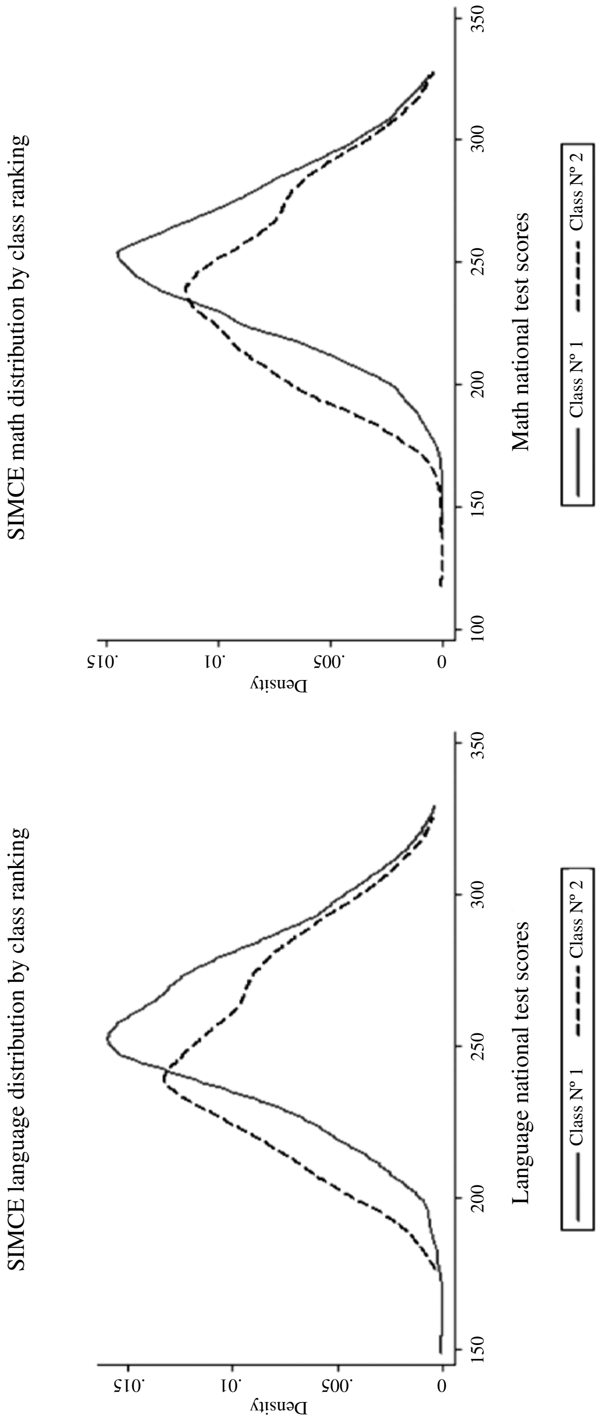


FIGURE 2
STANDARDIZED PERFORMANCE BY CLASS RANKING IN PRESENCE OF WITHIN SCHOOL SORTING



4.2. How do teacher's attributes really impact academic performance?

The previous section highlighted the positive matching of student and teacher characteristics between schools and showed that matching within a school is observable through teacher experience and mathematics curriculum coverage. Positive sorting generated a potential overestimation of the effects associated with teacher attributes on student learning as measured by EQMS scores. Below, results of the method used for correcting this bias are described.

Table 7 shows the combined results of the four models. In the first case, the results of the coefficients of the teacher attributes between Model 1 and Model 2 were compared, and neither model corrected for any type of sorting. Models 3 and 4 included corrections for sorting both between and within schools, and as the corrections were performed, changes in parameters of teacher attributes were identified.

In Model 1, which can be interpreted as a baseline reference, we observed a positive relationship between female teachers and test scores in both reading and mathematics. Likewise, having a secondary teaching degree was associated with better results on the EQMS tests when compared to having an elementary teaching degree. In contrast, not having a teaching degree had a negative effect, where a greater influence was observed for mathematics.

The positive effect associated with having a degree in secondary education for teachers of 4th grade students completely disappeared when correcting for the effect of student and family characteristics in Model 2. This finding suggests that this type of teacher most likely already teaches high SES students, and thus, these teachers do not present any relative advantages compared to those with a degree in elementary education. Moreover, the results of Model 2 indicated that teachers with a degree in preschool education or with no teaching degree had a negative correlation with student performance. However, it is important to note that less than 1% of Chilean teachers lack a teaching degrees, and it is therefore not a substantial problem for the school system.

When analyzing the correlation of a teacher's professional experience with EQMS results, the high concentration of the most experienced teachers (25 or more years) in schools catering to the most vulnerable students explained a high negative coefficient of this variable in Model 1. However, when including student and family characteristics (Model 2), the sign of the attribute changed and more experience was positively associated with academic performance in both reading and mathematics. A similar phenomenon was observed for teachers who only taught reading, where in Model 1, the effect of this attribute was 21.2 additional points on the EQMS test, whereas in Model 2, this effect was reduced to a negative, though non-significant, coefficient of 1.5 points. Both sign changes reflect the high rate of teacher and student attribute matching between schools.

Finally, achieving greater curriculum coverage had an important positive effect on the students' academic performance. Notably, this was considerably greater in mathematics than in reading, indicating a differentiated effect of the institution on academic performance. Estimations with Model 2 again showed that the impact of this variable was mitigated when correcting for student and family attributes, although it maintained its statistical significance. For students whose teachers achieved curriculum coverage 1 SD greater than the national average, the increase in mathematic performance was 10.4 points, an equivalent of 0.2 SD. In the case of reading, the impact was 4.8 points, an equivalent of 0.1 SD.

In Model 3, columns 5 and 6, the fixed effects of schools were included to correct for sorting between schools and teachers. Through this, the estimated teacher coefficients in a model with fixed effects are a product of only the variability within the schools with teacher attributes in different courses of the same grade (*within effects*). There was need to consider the high probability that within schools teachers would have comparable quality, given the lesser variability within schools in comparison with the overall sample. The consequence of this lower variability is that attributes within the school could lead to an underestimation of parameter values. This fact explains why Clotfelter, Ladd and Vidgor (2006) consider the coefficients resulting from this type of methodology to be the lower bound of the teacher effect and, likewise, the results without the fixed effects as the upper bound.

When correcting for fixed effects in Model 3, smaller coefficients were observed than those in the estimates of Model 2, but these still remained significant. The most relevant change corresponded to homogeneity in the effectiveness of teachers with more professional experience. After completing the first five years of teaching, the effect of experience does not increase. This situation was present both in mathematics and in reading and is similar to results found in the United States.[15] This result explains why improvements in teacher performance do not significantly increase after the first years of teaching, and why the greater heterogeneity observed in Model 2 occurred only among schools capable of strengthening and improving the capabilities of their teachers.

A second, relevant change observed in the fixed effect model corresponded to a reduction on the effects of greater curriculum coverage. In the case of mathematics, a curriculum coverage of over 1 SD, in comparison with a coverage level 1 SD below the average, had a reduced effect on EQMS scores, from 10.4 points to 4.8 points (Columns 4 and 6, respectively), whereas in reading, the effect went from 4.5 to 1.7 points.

When considering only the sample of students who attended schools free of sorting within the school (Model 4), unbiased estimators of the effects of teacher characteristics on student performance were used. The results indicated that female teachers had a positive effect on student learning in reading and mathematics (see Columns 7 and 8 of Table 5), but the effects were within a range of 7% to 8% of one standard deviation, which is in congruence with international and national literature for the effect of role model for female teachers (Paredes 2014).

In turn, the type of teacher educational institution attended was only relevant for mathematics, although the effect was very small and close to 3% of a standard deviation. No differences were found for reading. Likewise, not having professional qualifications had a negative but non-significant effect on mathematics.

A variable that increased both mathematics and reading test scores was initial teacher experience. This, however, was only relevant for those who had between 2-5 years of experience, as compared to those who had just started teaching. Again, the magnitude of the impact was slight and within 7% of a standard deviation for reading and 3% for mathematics. Estimates indicated that this result was maintained only for teachers with greater experience in reading,

---

15 In a recent literature review on the effects of professional experience on school performance, Goldhaber (2008) confirms the idea that this is nonlinear, and its positive effect is concentrated within the first five years of experience.

but this initial relative advantage disappeared in mathematics. These results are aligned with the opposite direction of sorting between and within schools found for experience by the previous analysis. At the beginning professional career teachers' experience is negatively sorted between schools (column 1 and 2), but after correcting by between sorting (fixed effects) is an attribute with positive impact. Then, the positive effects of more teacher experience disappear controlling by within sorting (using the sample without sorting).

A third attribute that improved test scores was teacher training. The results indicated that postgraduate qualifications (e.g., diploma, Master's degrees or PhDs) did not affect test scores, whereas short-term courses intended to develop skills directly related to classroom work) had a positive and significant effect within 5% of one standard deviation in both reading and mathematics. This result should be further analyzed controlling by motivation or teacher abilities, the main unobserved variables.

Furthermore, greater curriculum coverage only had a positive effect in the case of mathematics, achieving up to 8% of a standard deviation for those teachers that achieved coverage over the lower third of the national distribution. However, there was no difference in reading. These results reflect that up to 4th grade, the curriculum shortcomings of mathematics cannot be adequately resolved by family factors or non-observable institutional factors. Finally, the effect found for math should motivate future research to differentiate the effect of curriculum coverage from students' non observable abilities because in its current specification it is highly dependent on students' characteristics.

The obtained results were consistent with previous cases in the literature regarding teacher attributes and capacities of teachers that affect student performance, such as gender, initial professional experience, professional qualifications, and specialized professional development activities, also the magnitude of estimated effects on Chilean students achievement are similar with those calculated for the United States (Goldhaber 2008). Also, there was some of these attributes that have a small positive effect only in mathematics as in Goldhaber (2008). Comparing with other policies, the magnitude of estimated impact of teacher's attributes on students' performance is relevant in the Chilean context. For example, Bellei (2009) estimates that full school day reform in Chile raises 5%-7% of a standard deviation for reading test and 0%-12% for math test at student level, which are similar values to our estimates of teacher gender, experience or curriculum coverage.

Finally, the result linked to higher mathematics curriculum coverage may reflect both contextual teaching difficulties and the importance of strengthening the pedagogical capabilities and content knowledge in initial elementary teacher education. This is consistent with the generally lower results observed in both national and international tests of mathematical knowledge for new Chilean teachers.[16]

[16] Recent IEA TEDS-M study on future teachers mathematics content and pedagogical content knowledge in 34 Chilean universities (Ávalos and Matus, 2010), placed Chile as 15th out of 16 participating countries in primary level mathematics knowledge for teaching and last in lower secondary mathematics knowledge for teaching. Additionally, the results of a an exit test for future teachers, which also measures content and pedagogical knowledge, are unsatisfactory especially in mathematics (cited in Bellei and Valenzuela 2010).

TABLE 7

SUMMARY OF ACHIEVEMENT ESTIMATIONS FOR 4TH GRADE URBAN SAMPLE, INCLUDING VARIABLES OF STUDENTS, TEACHERS AND SCHOOLS

		Teacher characteristics		Teacher, student and school characteristics		Teacher and student characteristics plus fixed effects		Teacher, student characteristics, fixed effects in free of sorting sample.	
		Model 1		Model 2		Model 3		Model 4	
		Language (1)	Math (2)	Language (3)	Math (4)	Language (5)	Math (6)	Language (7)	Math (8)
Teacher Gender	Percent female	7.32***	5.85***	4.35***	3.74***	3.20***	2.78***	4.45***	3.96***
		(–0.38)	(–0.38)	(–0.34)	(–0.35)	(–0.42)	(–0.42)	(–0.62)	(–0.62)
Type of degree obtained from professional studies (base is degree on Primary education)	Differential Education	0.49	4.57**	–2.01	–0.61	0.06	0.55	–1.06	–2.95
		(–2.23)	(–2.25)	(–2.01)	(–2.02)	(–2.39)	(–2.42)	(–4.00)	(–4.03)
	Secondary Education	12.96***	12.96***	–0.06	–1.04	0.98	–1.85	3.56	–0.1
		(–1.4)	(–1.44)	(–1.27)	(–1.30)	(–1.62)	(–1.64)	(–2.66)	(–2.65)
	Preschool Education	–4.64***	–4.11***	–3.35***	–2.74***	–1.07	–0.58	–2.47*	–3.04**
		(–0.84)	(–0.86)	(–0.76)	(–0.77)	(–0.94)	(–0.95)	(–1.50)	(–1.50)
	Degree in other area	1.33	–0.76	3.16*	3.06*	1.57	1.42	–2.28	–1.34
		(–2.00)	(–2.03)	(–1.80)	(–1.82)	(–2.21)	(–2.23)	(–2.80)	(–2.82)
	No professional degree	–5.23***	–11.49***	–0.78	–6.05***	0.57	–3.52***	–0.23	–1.68
		(1.16)	(1.17)	(1.05)	(1.05)	(1.24)	(1.23)	(1.92)	(1.86)
Type of institution of professional studies (base is university)	Normal School	–1.54***	–2.26***	–0.62*	–1.26***	0.38	–0.19	–0.7	–1.82***
		(–0.41)	(–0.41)	(–0.37)	(–0.37)	(–0.44)	(–0.44)	(–0.67)	(–0.68)
	Professional Institute	–6.12***	–6.62***	–2.37***	–2.38***	–1.66***	–1.65**	–1.15	–1.85*
		(–0.61)	(–0.62)	(–0.55)	(–0.55)	(–0.65)	(–0.65)	(–1.02)	(–1.03)
	No professional studies	–22.68***	–24.45***	–13.29***	–15.97***	–6.16***	–7.11***	–1.74	–4.33
		(–2.03)	(–2.02)	(–1.83)	(–1.82)	(–2.15)	(–2.13)	(–3.09)	(–3.00)
Teacher years of experience (base is one or less years of experience)	2- 5 years	0.13	0.62	2.25***	2.93***	3.03***	2.92***	3.83***	1.93*
		(–0.7)	(–0.71)	(–0.63)	(–0.64)	(–0.77)	(–0.77)	(–1.12)	(–1.12)
	6-12 years	0.23	0.73	3.26***	4.00***	3.28***	3.05***	2.83**	0.64
		(–0.69)	(–0.70)	(–0.62)	(–0.62)	(–0.76)	(–0.76)	(–1.11)	(–1.11)
	13-24 years	–0.3	0.53	4.01***	5.10***	2.77***	2.95***	2.80***	1.31
		(–0.66)	(–0.67)	(–0.60)	(–0.60)	(–0.74)	(–0.75)	(–1.08)	(–1.09)
	More than 25 years	–8.94***	–9.27***	4.19***	4.80***	2.74***	2.37***	2.67**	0.77
		(–0.65)	(–0.66)	(–0.60)	(–0.61)	(–0.76)	(–0.76)	(–1.12)	(–1.12)

Certificates or credentials (base is no certificate or credential)	Diploma	0.29	–0.12	0.76***	0.54**	0.08	–0.08	–0.51	–0.71
		(–0.27)	(–0.27)	(–0.24)	(–0.25)	(–0.3)	(–0.30)	(–0.47)	(–0.48)
	Post-title	4.47***	3.95***	0.92*	0.1	0.37	–0.42	2.49***	2.54***
		(–0.56)	(–0.57)	(–0.51)	(–0.52)	(–0.62)	(–0.63)	(–0.96)	(–0.97)
	Master or Doctorate	3.46***	2.52***	–1.12	–2.38***	–1.58*	–2.57***	0.93	–0.06
		(–0.76)	(–0.77)	(–0.68)	(–0.69)	(–0.82)	(–0.83)	(–1.28)	(–1.28)
A single subject teacher	Teacher teaches only Verbal	21.19***		-1.49		-4.82		-0.01	
		(1.78)						(-5.44)	
Language Curriculum Coverage (base is 1 desv. Est. less than average)	More than a Std. Dev.	7.79***		4.51***		1.70***		0.52	
		(-0.42)		(-0.38)		(-0.5)		(-0.78)	
	Within a 1 desv. Est	5.60***		2.74***		1.17***		1.02	
		(-0.35)		(-0.32)		(-0.41)		(-0.63)	
Math Curriculum Coverage (base is 1 desv. Est. less than average)	More than a Std. Dev.		19.05***		10.39***		4.80***		4.00***
			(-0.42)		(-0.38)		(-0.5)		(-0.77)
	Within a 1 desv. Est		11.88***		6.33***		3.67***		4.19***
			(-0.35)		(-0.32)		(-0.39)		(-0.619
Other Controls	2006 dummy	-4.07***	-0.51**	-3.63***	0.27	-4.57***	-0.84***	-3.95***	0.11
		(-0.25)	(-0.26)	(-0.27)	(-0.27)	(-0.28)	(-0.28)	(-0.44)	(-0.44)
	Students and family characteristics	No	No	Yes	Yes	Yes	Yes	Yes	Yes
	School characteristics	No	No	Yes	Yes	No	No	No	No
	School fixed effects	No	No	No	No	Yes	Yes	Yes	Yes
	Constant	252.21***	239.97***	157.35***	141.66***	145.65***	131.21***	205.61***	199.68***
		(-0.75)	(-0.76)	(-1.36)	(-1.38)	(-5.42)	(-5.47)	(-3.51)	(-4.22)
Number of observations		183792	184673	183792	184673	183792	184673	77767	78353
R^2		0.016	0.023	0.202	0.214	0.253	0.275	0.256	0.277

Standard errors in parenthesis *** p<0.01, ** p<0.05, * p<0.1

5. Discussion and Conclusions

This study showed for the first time the magnitude and characteristics of between and within school sorting using a large database for 4th grade Chilean students, and conducted estimates of teacher attributes on a student's academic performance with corrections made for sorting bias. Besides showing between and within school sorting in Chile, the novelty of this work is that elaborates a methodology that deepens the analysis on within school sorting, which has been hardly investigated.

Findings confirmed an important and positive sorting between schools based on teacher attributes and the socioeconomic and cultural conditions of students and their families (Meckes and Bascopé 2010, Ruffinelli and Guerrero 2009). On average, teachers with certain demographic and academic attributes, such as being young, having a college degree, achieving higher curriculum coverage, being female, and/or having conducted specialized studies, tended to teach students in schools with better socioeconomic conditions.

This type of positive sorting between schools was intensified by the frequent matching that also occurred within schools. The findings of this study revealed that 34% of urban schools with two or more 4th grade classes did not distribute their students homogeneously, but rather by student socioeconomic attributes. Evidence was also found for sorting within equivalent sections of a grade at a higher frequency in schools with less favorable learning conditions. These results suggest that the Chilean school system segregates students into classes early on, which is similar to other countries with high heterogeneity in teacher quality (Goldhaber 2008). Opportunities for improving educational quality are affected by the difficulty of getting better-quality teachers in schools with the most vulnerable students and also by the early segregation of students within these schools (Valenzuela, Bellei and De los Ríos 2009).

The partial identification of the effects associated with student and family attributes, as well as those of teachers and schools, leads to the conclusion that the nonrandom distribution of students and teachers entails not just an overestimation of the effects that some educational resources have on academic improvement, but also the possibility for mistakes in public policy design. This was the case for 4th grade teachers that taught just one subject or who had a degree qualifying them to teach middle school. The initial positive correlation of their attributes with standardized test results completely disappeared from the present analyses when correcting for the biases associated with student and school attributes.

Additionally, including fixed effects for schools in the estimations on academic performance led to the conclusion that schools add value to the process of student learning and that family and teacher attributes are not the only relevant factors for explaining student outcomes. The general estimates for 2005 and 2006 indicated that schools acted by increasing and balancing a teacher's potential, something observable in the reduction of the individual impact of experienced teachers with higher curriculum coverage, especially in mathematics, and in the reduction of the harmful effects of teachers with lower-quality academic training. This finding may be due to schools applying a number of diverse mechanisms for teacher, family, and student selection, and also because schools use different leadership, motivation, pedagogy, and administrative management strategies. The results obtained in this study show that these factors were generally important and could considerably benefit the academic performance of students, both directly and through higher teacher effectiveness.

In turn, this study has shown that teacher attributes needed to improve academic performance are scarce and that the overestimation of their impact is generalized due to the problem of nonrandom distribution of teachers between and within schools.

The results of this study allow concluding that appropriate school educational management can strengthen the performance of teachers and that student learning heterogeneity of academic results cannot be explained solely by strong school segregation. Among the other determinants of segregation was the positive sorting between schools and teachers who possess better professional and pedagogical attributes and families of a higher SES. Remarkably, the results observed in the various models indicate that the sorting of teachers to students according to socioeconomic characteristics, as it was observed in Model 2, has a larger effect than the biases associated with heterogeneity in the pedagogical management of Chilean primary schools, which was controlled by fixed effects in Model 3.

This study found four teacher characteristics with positive effects, although small, on the academic performance of 4^{th} grade students. Consistent with existing literature, female teachers outperformed their male peers by 8% of one standard deviation, and teacher experience had a positive impact on school learning, but only during the first years of experience. This finding suggests a serious need for greater teacher retention and building on professional teaching experience for improving educational effectiveness.

Moreover, the results indicated that professional development of teachers may improve school learning, although only specialized short-term courses proved to be the most effective. Finally, higher curriculum coverage in mathematics entailed an increase in the EQMS scores of up to 8% of one standard deviation. This finding did not occur in the case of reading, which again confirms the greater relevance of the role played by schools and teachers in mathematics education. This role is not easily compensated by the family, and shows a need for schools to widen the mathematics curriculum to foster improvement in student performance.

Although these results cannot be extrapolated to other levels of elementary or secondary education, it is possible to observe that an educational system with high levels of student segregation between and within schools, starting from the first years of schooling, reduces the possibility that the most vulnerable schools and students will have the best-quality teachers.

Consequently, there is a widening performance gap between students in different demographic conditions as well as an apparent reduction in the effectiveness of critical teacher attributes. In this context, to improve education for all children, the following strategies are recommended as essential: i) strengthening the mechanisms for attracting and retaining effective teachers in schools located in more vulnerable zones; ii) improving professional development opportunities for new teachers in ways that effectively benefit academic performance; and iii) improving pedagogical and directive management skills within schools, especially for those which teach to vulnerable students.

References

Aaronson, D.; Barrow, L. and Sander, W. (2007). "Teachers and Student Achievement in the Chicago Public High Schools", *Journal of Labor Economics*, University of Chicago Press, Vol. 25 Nº 1, pp. 95-135.

Ávalos, B. and Matus, C. (2010). *La Formación Docente en Chile desde un Óptica Internacional en Informe Nacional del Estudio Internacional IEA TEDS-M.*, Santiago.

Barber, M. and Mourshed, M. (2007). *How the world's best-performing school systems come out on top.*

Bellei, C. (2009). "Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile", *Economics of Education Review*, Vol. 28 Nº 5, pp. 629-640.

Bellei, C. and Valenzuela, P. (2010). "¿Están las condiciones para que la docencia sea una profesión de alto estatus en Chile?" (Martinic, S. and Elacqua, G., Eds.), *Fin de ciclo. Cambios en la gobernanza del sistema educativo*, UNESCO-PUC, pp. 1-26.

Boyd, D.; Lankford, H.; Loeb, S.; Rockoff, J. and Wyckoff, J. (2008). "The narrowing gap in New York City teacher qualifications and its implications for student achievement in high-poverty schools", *Journal of Policy Analysis and Management*, Forthcoming in the J. of Policy Analysis and Management, John Wiley & Sons Inc., Vol. 27 Nº 4, pp. 793-818.

Bravo, D.; Falck, D.; González, R.; Manzi, J. and Peirano, C. (2008). *La relación entre la evaluación docente y el rendimiento de los alumnos: Evidencia para el caso de Chile*, Santiago.

Cabezas, V.; Gallegos, F.; Santelices, V. and Zarhi, M. (2011). "Factores Correlacionados con las Trayectorias Laborales de Docentes en Chile , con Especial Énfasis en sus Atributos Académicos", *Fondo de Investigación y Desarrollo En Educación - FONIDE*, Santiago, Chile, Proyecto F.

Clotfelter, C.; Ladd, H. and Vidgor, J. (2006). "Teacher-Student Matching and the Assessment of Teacher Effectiveness", *Journal of Human Resources*, Vol. 41 Nº 4, pp. 778-820.

Clotfelter, C.; Ladd, H. and Vigdor, J. (2007). "How and why do teacher credentials matter for student achievement?", *National Bureau of Economic Research Working Paper Series*, NBER Working Paper, National Bureau of Economic Research Cambridge, Mass., USA, Vol. Nº 12828 Nº 2, available at: http://www.nber.org/papers/w12828.

Dupriez, V.; Dumay, X. and Vause, A. (2008). "How Do School Systems Manage Pupils ' Heterogeneity?", *Comparative Education Review*, Vol. 52 Nº 2, pp. 245-273.

García-Huidobro, J.E. (2007). "Desigualdad educativa y segmentación del sistema escolar: consideraciones a partir del caso chileno", *Revista Pensamiento Educativo*, Vol. 40, pp. 65-85.

Goldhaber, D. (2008). "Teachers matter, but effective teacher quality policies are elusive", in Ladd, H.F. and Fiske, E.B. (Eds.), *Handbook of Research in Education Finance and Policy*, Routledge, pp. 146-165.

Goldhaber, D. and Anthony, E. (2007). "Can Teacher Quality Be Effectively Assessed? National Board Certification as a Signal of Effective Teaching", *Review of Economics and Statistics*, Vol. 89 Nº 1, pp. 134-150.

Goldhaber, D. and Brewer, D. (1997). “Why Don’t Schools and Teachers Seem to Matter? Assessing the Impact of Unobservables on Educational Productivity”, *The Journal of Human Resources*, University of Wisconsin Press, Vol. 32 Nº 3, pp. 505-523.

González, P.; Mizala, A. and Romaguera, P. (2002). “Recursos diferenciados a la educación subvencionada en Chile”, *Serie Economía*, Vol. 150.

Hanushek, E.A. and Woessmann, L. (2011). “Chapter 2-The Economics of International Differences in Educational Achievement”, *Handbook of the Economics of Education*, Vol. 3, pp. 89-200.

Harris, D.N. and Sass, T.R. (2011). “Teacher training, teacher quality and student achievement”, *Journal of Public Economics*, Vol. 95 Nº 7-8, pp. 798-812.

Hsieh, C.-T. and Urquiola, M. (2006). “The effects of generalized school choice on achievement and stratification: Evidence from Chile’s voucher program”, *Journal of Public Economics*, Vol. 90 Nº 8-9, pp. 1477-1503.

Jacob, B.A. and Lefgren, L. (2004). “School Reform Efforts in Chicago The Impact of Teacher Training on Student Achievement Evidence from School Quasi-Experimental Reform Efforts in Chicago”, *The Journal of Human Resources*, Vol. 39 Nº 1, pp. 50-79.

Kane, T.J. and Staiger, D.O. (2002). “The Promise and Pitfalls of Using Imprecise School Accountability Measures”, *Journal of Economic Perspectives*, American Economic Association, Vol. 16 Nº 4, pp. 91-114.

Lankford, H.; Loeb, S. and Wyckoff, J. (2002). “Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis”, *Educational Evaluation & Policy Analysis*, Sage Publications, Vol. 24 Nº 1, pp. 37-62.

Lara, B.; Mizala, A. and Repetto, A. (2010). “Una Mirada a la Efectividad de los Profesores en Chile”, *Estudios Públicos*, Vol. 120 Nº Primavera.

McEwan, P.J.; Urquiola, M. and Vegas, E. (2008). “School Choice, Stratification, and information on School Performance: Lessons from Chile”, *Journal of LACEA Economia*, Vol. Spring.

Meckes, L. and Bascopé, M. (2010). “Distribución inequitativa de los nuevos profesores mejor preparados”, *Primer Congreso Interdisciplinario de Investigación en Educación*, p. 19.

MINEDUC. (2009). *Resultados Evaluación Inicia*, Santiago, Chile.

MINEDUC. (2013). *Presentación proyecto de ley sobre carrera profesional docente en el Congreso Nacional*.

Mizala, A. and Torche, F. (2012). “Bringing the schools back in: the stratification of educational achievement in the Chilean voucher system”, *International Journal of Educational Development*, Elsevier Ltd, Vol. 32 Nº 1, pp. 132-144.

Mizala, A. and Urquiola, M. (2007). *Parental choice and school markets: The impact of information approximating school effectiveness*, Documentos de Trabajo, Centro de Economía Aplicada, Universidad de Chile, available at: http://ideas.repec.org/p/edj/ceauch/239.html.

Monk, D.H.; King, J. and Ehrenberg, R.G. (1994). “Multi-level Teacher Resource Effects on Pupil Performance in Secondary Mathematics and Science: The Role of Teacher Subject Matter Preparation”, in Ehrenberg, R.G. (Ed.), *Contemporary Policy Issues Choices and Consequences in Education*, ILR Press.

OECD. (2009). *PISA 2009 Results: Overcoming Social Background*, *OECD*, OECD Publishing, Vol. II, pp. 1-224.

OECD. (2013). *PISA 2012 Results: Excellence through Equity*, Paris, France, p. Volume II.

Ortúzar, S.; Flores, C. and Milesi, C. (2012). "Distribución de profesores principiantes, jefes de UTP y directores en distintos tipos de establecimientos", *II Congreso de Educación CIAE-CEPPE, Santiago, Chile*, pp. 1-9.

Paredes, R.; Bogolasky, F.; Cabezas, V.; Rivero, R. and Zahri, M. (2014). "Los determinantes del primer trabajo para profesores de Educación Básica en la Región Metropolitana", *Evidencias para Políticas Públicas en Educación*, Sexto Concurso FONIDE, Ministerio de Educación, pp. 35-90.

Paredes, V. (2014). "A teacher like me or a student like me? Role model versus teacher bias effect", *Economics of Education Review*, Elsevier Ltd, Vol. 39, pp. 38-49.

Rivero, R. (2012). "The distribution of highly qualified teachers in Chile", *II Congreso de Educación CIAE-CEPPE,* Santiago, Chile, pp. 1-38.

Rivkin, S.G.; Hanushek, E.A. and Kain, J.F. (2005). "Teachers, Schools, and Academic Achievement", *Econometrica*, The Econometric Society, Vol. 73 Nº 2, pp. 417-458.

Rockoff, J.E. (2004). "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data", *American Economic Review*, American Economic Association, Vol. 94 Nº 2, pp. 247-252.

Ruffinelli, A. and Guerrero, A. (2009). "Círculo de segmentación del sistema educativo chileno: Destino laboral de los egresados de pedagogía en educación básica", *Calidad en la Educación*, Vol. 31 Nº 31, pp. 20-44.

Sanders, W.L. and Rivers, J.C. (1996). "Cumulate and Residual Effects of Teachers on Future Student Academic Achievement Summary of Findings", *University of Tennessee Value-Added Research and Assessment Center*.

Treviño, E.; Valenzuela, J.P. and Villalobos, C. (2014). "¿Se agrupa o segrega al interior de los establecimientos escolares chilenos?", Reporte VII Concurso FONIDE, Ministerio de Educación.

Valenzuela, J.P.; Bellei, C. and De Los Ríos, D. (2009), "Evolución de la segregación socioeconómica de los estudiantes chilenos y su relación con el financiamiento compartido", *Selección de Investigaciones Primer Concurso FONIDE: Evidencias para Políticas Públicas en Educación*, Santiago, Chile. Ministerio de Educación-FONIDE, pp. 229-284.