The Effect of Teacher Qualifications on Student Achievement Gains: Evidence from the FUNDESCOLA Schools in Brazil, 1999-2003

Raquel Guimaraes[*]

Abstract

It is recognized that teachers have a crucial role in improving learning outcomes. Because quality differs by teacher, their potential impact on student outcomes may also differ. However, measuring teacher quality and its influence on learning has been challenging in educational research. In this paper I explore the effect of teacher qualifications on the rate of learning in Mathematics and Portuguese in Brazil. Drawing on longitudinal data from the *Fundo de Fortalecimento da Escola* (FUNDESCOLA) program from 1999 to 2003 and using a dose-response analysis and a latent trait model, I conclude that teacher qualifications positively influence the pace of learning in Mathematics. However, results for Portuguese were not significant, which may indicate that subjects require different models of teacher effectiveness. Also, I find that students' socioeconomic status does not enhance the effect of teacher qualifications on the pace of learning in Mathematics and Portuguese.

1. Introduction

The increasing quest for development, both economic and social, has intensified concern around the world about the importance of human capital accumulation. This issue is especially appropriate to developing countries, which lag behind in terms of the level of schooling of their populations (Benhabib and Spiegel 1994). Even though there has been significant improvement in terms of expanding access to education in many countries (Barro and Lee 1993; Behrman 1987), quality enhancements are not prevalent across developing societies (Hanushek and Woessmann 2008). In other words, more widespread effort is necessary to improve the quality of education in order to promote development.

The need to enhance the quality of education has intensified in Brazil since the end of the 1990s (Rios-Neto and Guimaraes 2010). Despite being a major economic power and having universalized access to primary education, Brazil still stagnates in cross-country comparisons of poor students' achievement, as shown by PISA results (OECD 2011). Consequently, there is an emerging government agenda aimed at improving the quality of education, notably in public schools (Soares, 2007).

For these policymakers, quality means that students are succeeding on standardized tests. In this context, teachers are assigned a central role in delivering a high-quality education by policymakers and international organizations (OECD 2005; UNESCO 2006). Therefore,

[*] Ph.D. candidate in Demography at Federal University of Minas Gerais, Brazil. raquel@cedeplar.ufmg.br

increasing the quality of education requires understanding the means to increase teacher quality and how teachers can be effective in promoting learning.

Amongst education researchers, teacher quality is widely considered an important school factor – and maybe even the most important factor – of student learning (Ladd 2008; Rivkin, Hanushek, and Kain 2005). In the large body of research that has addressed the importance of teacher quality and effectiveness in the learning process, the United States educational system appears most prominently (Boyd et al. 2008; Aaronson, Barrow, and Sander 2007; Desimone and Long 2010; Ladd 2008), and to a lesser extent studies include other countries, such as South Africa and Botswana (Carnoy et al. 2011). In Brazil, research has mainly explored the influence of school factors on student achievement (Soares 2006; Soares 2007), but there are only a few studies that have conducted a systematic analysis of teacher effects (Fonseca 2011; Menezes-Filho and Pazello 2007).

This study explores the role of teacher qualifications in enhancing the pace of learning in Brazil from the schools participating in the *Fundo de Fortalecimento da Escola* (FUNDESCOLA), or School Improvement Plan[1] from 1999 to 2003. This program, funded by the World Bank, started in 1997 and aimed to improve schools in six states in the country with some the lowest educational indicators: Rondônia, Pará, Pernambuco, Sergipe, Mato Grosso do Sul, and Goiás. In order to evaluate the effectiveness of the FUNDESCOLA program, a longitudinal survey was then carried out from 1999 to 2003.[2]

Scholars have benefited from the FUNDESCOLA data to analyze the impact of the program (Carnoy et. al., 2008) and the factors associated with school achievement (Macedo 2004; Pereira 2006) and school retention (Luz 2008). Yet, a more specific analysis of distinct aspects of teachers and their effects on the pace of student learning in the FUNDESCOLA schools was not yet been explored.

Taking advantage of the FUNDESCOLA survey and drawing on an Educational Production Function (EPF) framework, this study examines the relationship between teacher qualification and pace of learning in Portuguese and Mathematics for students in fourth to eight grade by testing two hypotheses. First, I test whether students who learned faster are the ones who, on average, were exposed to high-quality teachers, that is, those with higher qualifications. By assuming that teacher qualification is an exogenous variable, it is possible to infer whether a teacher with higher qualifications predicts larger student achievement gains. Next, I test whether the socioeconomic status (SES) of a student enhances the effect of teacher qualifications on his/her pace of learning, as those from lower SES groups may require greater teacher effectiveness in order to achieve the same achievement gains as their higher SES counterparts (Brophy 1986).

In order to test these hypotheses, I propose two methodological innovations. First, I estimate "teacher qualifications" and socioeconomic indexes using a latent trait model under the Item Response Theory approach. Next, I employ a dose-response analysis as an identification strategy to derive the causal effect of teacher qualifications on the pace of learning in Portuguese

and Mathematics. This study, therefore, provides an important contribution for the assessment of teacher effects in Brazil.

2. Overview of Brazil and FUNDESCOLA states' teaching force at the elementary level

Starting from the 1980s, teachers assumed a central role in Brazilian education policy as agents responsible for improving the quality of education in the country (Brasil 1988; Brasil 1996). Since then, the Brazilian government has strived to improve the quality of elementary school teachers and their working conditions. For instance, the Brazilian Law for Basic Education of 1996 required a college degree for teachers of basic education. In 2009, the government released the National Policy on Teacher Education, which aimed to organize the initial and continuing training of basic education teachers (Brasil 2009). Basic education teacher training is recognized as a public commitment of the State, seeking to ensure the rights of children, youths, and adults to a high-quality education (Brasil 2009, sec. 2).

Despite its importance and government intentions, the teaching profession in Brazil remains unattractive: poor working conditions and low levels of earnings and prestige prevent the best students from choosing teaching careers (Louzano et. al., 2008). This situation is even worse in the public sector, where teachers must also address demands that are beyond their training, due to the needs of the students (Oliveira 2004).

Given the centrality of teachers in the learning process in Brazil and because this paper focuses on specific states in the country (those evaluated by the FUNDESCOLA survey), it is important to investigate whether there are vast differences in teacher qualifications and in their distribution based on their quality. From this evidence, I can conclude to what extent the teachers assessed by the FUNDESCOLA survey are more or less qualified than the average Brazilian teacher.

Drawing on Brazil's National Household Survey (PNAD) from 1992 to 2009 and on the national assessment of public school students (Prova Brasil) in 2007, I describe the main characteristics of the elementary-level teachers in FUNDESCOLA and non-FUNDESCOLA states: their educational composition, wages and qualifications.

Figure 1 shows the average number of years of schooling for elementary-level teachers[3] from 1992 to 2009 in FUNDESCOLA and non-FUNDESCOLA states using PNAD data. It can be drawn from the figure that the proportion of teachers with a college degree has increased significantly. This has resulted from a change in the educational law (LDB) in 1996, which required a minimum tertiary education level for elementary school teachers. Although both regions had improved the average qualification of their teachers, FUNDESCOLA states consistently have, on average, teachers with lower levels of schooling than the rest of the states. This may indicate that FUNDESCOLA teachers have lower quality when compared to their other Brazilian counterparts. However, the gap in average schooling between the states has decreased since 2005.

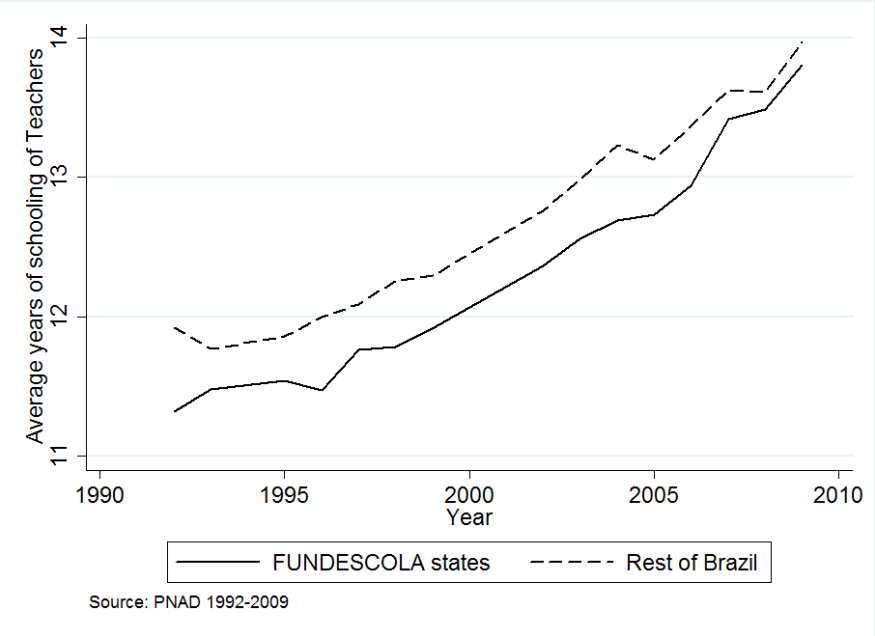

Figure 1. Average years of schooling of elementary-level teachers in FUNDESCOLA states and in the rest of Brazil, 1992 to 2009.

Wages reflect an important aspect of teacher quality. In general, high-quality teachers would be attracted to work in schools that pay the highest wages (Loeb and Page 2000). In Brazil, the wages of public school teachers at the elementary level are set by the local government (either at the state or municipal level) according to the teachers' qualifications. Therefore, it can be expected that the higher the wage of teachers in one region, the more it attracts high skilled teachers. Figure 2 shows the log of standardized (hourly) wages for elementary-level teachers from 1992 to 2009 in FUNDESCOLA states and in the rest of Brazil using PNAD data. It can be drawn from the figure that FUNDESCOLA states pay, on average, lower standardized wages for their elementary-level teachers and the differential between FUNDESCOLA and non-FUNDESCOLA states remains relatively stable over time. This evidence also indicates that FUNDESCOLA teachers may have lower quality when compared to their Brazilian counterparts.

Teacher qualifications, one important aspect of teacher quality, also differ among FUNDESCOLA and non-FUNDESCOLA states. The proportion of teachers who have a college degree, who have completed postgraduate studies, and who have participated in a professional development program are good predictors of the quality of the instruction provided to students in each region. Also, the type of higher education institution that the teacher attended and the level of experience are important correlates. Using data from Prova Brasil 2007, I compute these indicators for FUNDESCOLA and non-FUNDESCOLA states.

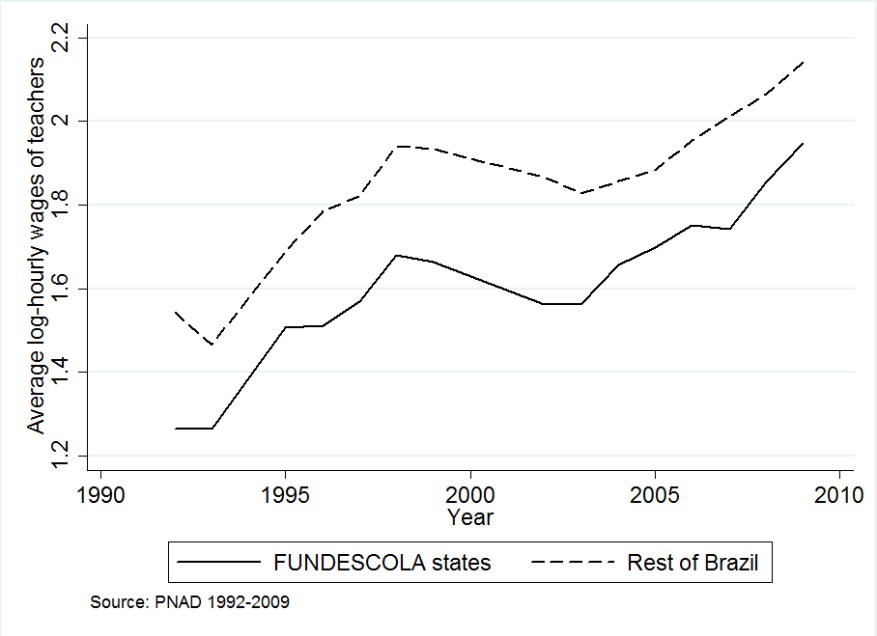

Figure 2. Average years of the log-hourly wages of elementary-level teachers in FUNDESCOLA states and in the rest of Brazil. 1992 to 2009

As shown in Table 1, the proportion of teachers with a college degree is higher for non-FUNDESCOLA states in fourth and eighth grade. Also, the proportion of teachers who attended public federal higher education institutions is higher in non-FUNDESCOLA states than in FUNDESCOLA states. In addition, the proportion of less experienced teachers is slightly higher for the FUNDESCOLA states. The proportion of teachers with some postgraduate study and professional development is similar for both regions.

In sum, Brazilian policymakers have recognized the role of teachers in the learning process and attempted to increase teacher quality. However, disparities and challenges still persist in the promotion of the teaching profession. In regards to regional disparities, FUNDESCOLA states have, in general, a less qualified teaching workforce than non-FUNDESCOLA states. This may influence my results: as teachers are less qualified, the effect of teacher qualifications on the pace of learning may be underestimated if contrasted with the average effect in Brazil. However, if the effects are significant for FUNDESCOLA states, then it can be inferred that they might be larger for Brazil as a whole, which raises the importance of investments in teacher qualifications.

Table 1: Teacher quality measures for 4th and 8th grade public school teachers in FUNDESCOLA and non-FUNDESCOLA states, 2007

		4th grade		8th grade	
		FUNDESCOLA states	Non-FUNDESCOLA states	FUNDESCOLA states	Non-FUNDESCOLA states
Highest level of schooling	Less than High School	0.2	0.1	0.1	0.1
	High School degree	68.2	61.5	8.0	7.9
	College degree	31.7	38.4	91.9	92.1
Type of HE institution	Public - Federal	14.4	9.5	28.2	13.3
	Public - State	28.3	23.3	25.7	21.3
	Public - Local	2.6	2.1	1.2	1.0
	Private	44.4	55.4	43.0	62.3
	Not completed college	10.3	9.7	2.0	2.1
Post-graduate studies	Yes	41.6	40.9	58.5	55.5
	No	58.4	59.1	41.5	44.5
Professional Development	Yes	85.4	85.0	76.4	76.1
	No	14.6	15.0	23.6	23.9
Experience	Less than 2 years	31.8	28.5	22.4	21.2
	3 to 4 years	20.3	20.2	17.6	17.8
	5 to 6 years	11.7	13.1	13.2	13.1
	7 to 8 years	7.3	8.1	9.9	9.7
	More than 8 years	28.9	30.1	37.0	38.2

Note: Author's own elaboration from Prova Brasil data (INEP/MEC), 2007

3. Teacher effectiveness and student learning

In this section, I present the current state of research in the areas of teacher effectiveness and teacher quality. First, I review studies that explore the relationships between teacher characteristics and student achievement. Then, I analyze research on teacher quality by describing the main determinants of teachers' effectiveness as well as measurement issues.

3.1. Challenges in the estimation of teacher effects

When considering school inputs to the learning process, it is widely accepted today that teachers are an important determinant of student achievement (Clotfelter, Ladd, and Vigdor 2007a; Boyd et al. 2008; Clotfelter, Ladd, and Vigdor 2007b; Ladd 2008; Nye, Konstantopoulos, and Hedges

2004). In the past, inadequate data and methodological limitations may have obscured the relevance of teacher effects.

The challenges that researchers find in the estimation of teacher effectiveness are not trivial. Ladd (2008), for instance, addressed the methodological limitations in assessing teacher effectiveness and argued that it is complex to isolate teacher effects from other potential confounders, such as family background, school resources, and unobserved traits of the student.

In the past decades, scholars have benefited from significant advances in methodology and the increased availability of longitudinal data for the estimation of teacher effects. By systematically controlling for student heterogeneity and for selection bias in the matching of students to teachers and schools, teacher effectiveness is now assessed in a more robust and consistent manner (Boyd et al., 2008; Clotfelter et al., 2007a; Ladd, 2008).

The literature on teacher effects is vast for the United States, perhaps as a result of a wide availability of longitudinal data. Taken together, the reviewed studies provide reasonable evidence that teachers' characteristics tend to be positively associated with student achievement (Clotfelter, Ladd, and Vigdor 2007a; Desimone and Long 2010; Rivkin, Hanushek, and Kain 2005).

However, this consensus ends when considering which are the most relevant predictors of teacher quality. Some authors found that teachers' experience, test scores, and regular licensure were positively associated with student achievement and that variations in teacher credentials had more explanatory power when considering students' performance (Clotfelter, Ladd, and Vigdor 2007b). However, some studies showed no predictive value of teacher credentials on variance of teacher effects (Rivkin, Hanushek, and Kain 2005; D. J. Boyd et al. 2009; Nye, Konstantopoulos, and Hedges 2004; Aaronson, Barrow, and Sander 2007). In the next section, I review the debate on the determinants of teacher quality.

3.2. How to define a high-quality teacher?

Given that teacher characteristics do influence student performance, it can be predicted that higher levels of teacher quality are positively associated with better student performance (Loeb 2001; Hanushek and Rivkin 2006). It is not surprising that this prediction is widely agreed upon. Yet, as teacher quality is a latent trait, some difficulties have arisen when researchers have attempted to estimate the real effect of teacher quality on students' proficiency. These difficulties in general are related to the choice and relevance of the variables used to approximate teacher quality (Eide, Goldhaber, and Brewer 2004; Brophy 1986; Hanushek et al. 2005; Hanushek and Rivkin 2006).

Overall, the studies reviewed differ on which proxy variables of teacher characteristics should be used to predict teacher quality. Some scholars focused on teacher credentials and experience as proxy variables for teacher quality (Boyd et al., 2009; Clotfelter et al., 2007b; Ladd, 2008). However, this approach is not without criticism because teacher effects, as

measured by teacher degree level and experience, have presented relatively small effect sizes (Ladd 2008).

On the other hand, some authors proposed the use of academic performance and teacher skills as proxies for teacher quality (Hill, Rowan, and Ball 2005; Eide, Goldhaber, and Brewer 2004). Debates persist about the appropriate balance between teachers' conceptual and procedural instruction when approximating teacher quality; it has not yet been determined which mix of content with which groups of students has what effect over what duration of time and under what circumstances (Nye, Konstantopoulos, and Hedges 2004). I concur with the studies cited above that pedagogical skills, that is, how teachers teach students, appear to be crucial for ensuring achievement gains. Nevertheless, it should be highlighted that teachers' pedagogical skills are not easily or often measured on surveys, which makes it difficult for researchers to evaluate and predict their real impact on student performance gains.

Altogether, these studies demonstrate that teacher credentials and academic knowledge are necessary but not sufficient conditions for student learning. In this paper I test whether teacher qualifications positively affect the pace of learning for students evaluated by the FUNDESCOLA data.

3.3. Teacher quality: inequality of access and effectiveness

As I have argued, teachers are considered to be an important (or the most important) determinant of student learning, even though there is no consensus on which characteristics define a high-quality teacher. However, how students are differently affected by a teacher quality is a relevant question for educational policy. For a given teacher quality, students from different social classes and ability levels may be not equally affected. For instance: are less advantaged students more sensitive to high-quality instruction than their high advantaged peers? Or are students from upper social strata more privileged when it comes to the effect of a high-quality teacher? Despite its relevance, there are few strong theoretical or empirical perspectives on how the effect of a high-quality teacher might differ according to students' background.

Some scholars posit that students from higher social strata benefit more from high-quality teachers than students from lower social classes. These researchers argue that the students' social and cultural background, which define their pre-instructional understanding of the subject, shape their understanding of information presented in classrooms by teachers (McDiarmid, Ball, and Anderson 1989, 88:19). Therefore, social background is a proxy for differentiated demands of the students of teacher support. For instance, high-SES students are more likely to be participative and self-confident in class and often do not need encouragement. On the other hand, low-SES students are more likely to demand attention and support from their teachers in addition to good instruction (Brophy 1986, 1073).

On the other hand, some researchers contend that students from low socioeconomic strata may benefit more from high-quality teachers. If high-quality teachers are able to identify

children's needs and adapt their teaching skills, then disadvantaged students may benefit more from teacher quality (Darling-Hammond 2006).

In this paper, I test whether students with high SES, as a proxy for social background, benefit more when it comes to the effect of a high-quality teacher than their low-SES counterparts.

3.4. Teacher effectiveness in Brazil

In Brazil, only a few studies have offered systematic analyses of teacher effects. By exploring an exogenous change in the resource allocation for teacher wages in the Brazilian public sector, Menezes-Filho & Pazello (2007) found that an increase in teacher salaries had a positive impact on proficiency. Using a national assessment, Fonseca (2011) showed that public school students in Brazil had higher achievement when exposed to teachers who studied in public institutions of higher education, instead of private institutions.

The scarcity of studies on teacher effects in Brazil may be explained by the lack of longitudinal data in the country. As presented in the last section, an accurate assessment of teacher effects requires longitudinal data following students throughout the school year (Ladd 2008). In this paper, I contribute to the existing literature by assessing teacher effects in Brazil on students' pace of learning using FUNDESCOLA data and a measure of teacher quality based on teacher qualifications.

4. Methods and data

In this study, I seek to understand the effect of teacher qualifications on changes in student achievement gains. My first hypothesis is that *higher levels of teacher qualifications improve the average rate of learning*. During the school trajectory from fourth to eighth grade, students can be exposed to teachers of different quality. Therefore, one interesting question would be whether the exposure to a specific level of teacher qualifications are related to a change in achievement gains. The second hypothesis in my study is that *students' SES enhance the effect of teacher qualifications by improving the average rate of learning*. The outcome of interest in this study is the growth in students' scores in Portuguese and Mathematics.

4.1. Identification strategy

This study aims to assess whether having a teacher with high qualifications affects the pace of students' learning in Portuguese and Mathematics. In order to achieve this goal, I address two issues: measurement of teacher quality and causal inference.

The first methodological challenge in this paper is measurement. Teacher quality is a latent trait. In surveys, information was collected on a set of teacher characteristics, such as schooling, content knowledge, experience and certification. However, these variables try to capture what is underlying teacher qualifications. I solve this problem by building a teacher qualifications index based on relevant variables present in the questionnaire. If modeling

assumptions are satisfied, this latent trait variable has the advantage of summarizing relevant information on teacher qualifications in a continuous scale. I discuss this method in more detail below.

The second challenge of this paper is to assess the causal effect of teacher qualifications on the pace of learning. For this task, experiments are recommended because they avoid bias caused by omitted variables and model specification (Krueger 1999). However, FUNDESCOLA data collection was conducted using an observational design, as students were not randomly assigned to high- or low-quality teachers. Despite its observational nature, causal inference using FUNDESCOLA data is possible by employing quasi-experimental design (QED). QED has been widely used by researchers as an alternative to experiments. However, this design requires that certain assumptions are met as well as that appropriate models are used (Stuart and Rubin 2008).

In this study, selection bias plays a central role in limiting causal inference. This bias occurs because, in general, students are not randomly assigned to teachers. Even though Brazil does not have an institutional system of student tracking in public schools, it is probable that school managers and teachers do assign students to classrooms based on certain decision rules. Moreover, students may be sorted by principals into classrooms according to their achievement. Also, parents may contest teacher assignments if they believe that these are related to their children's academic performance (Rothstein 2009). If these conditions hold, it is extremely difficult to isolate the effect of a teacher from pre-existing differences among students.

Also, the choice of the specification of the Educational Production Function in which student achievement gains are used as dependent variable may induce a downward bias in the estimates. First, it is worth mentioning that this strategy solves an important limitation of VAM, as regressions on current achievement on baseline achievement usually lead to misleading estimates of the parameters of interest: in the presence of measurement error, the observed final and initial scores are a poor estimate of the correlation between the true final and initial scores (Rogosa 1988). However, by using gains as a dependent variable, the effect of teacher qualifications will be biased downward. Clotfelter, Ladd, and Vigdor (2007) demonstrated that, in the gains equation, the estimated coefficients will be biased downward if there is not perfect correlation between initial and final test scores. Therefore, the effect of teacher qualification on students' pace of learning in this study is likely underestimated.

4.2. Empirical specification

4.2.1. Latent trait model for teacher qualifications and SES

The use of Item Response Theory (IRT) to model latent traits has been increasingly popular among scholars in various fields, and there is substantial evidence that it can be appropriately applied in educational research (Stout 2007; De la Torre and Douglas 2004; Yen 2005). IRT stands for a collection of mathematics and statistical methods suited to analyze items and scales, to create ability measures, and to measure individuals on latent constructs (Reise, Ainsworth, and

Haviland 2005). In regards to construct building, measures obtained by Item Response Theory present at least two advantages: they provide an adjustment to the continuous latent trait distribution, and they reduce dimensionality by offering relevant information about the trait of interest (Kolen and Brennan 2004; Samejima 1969; Stout 2007; Yen 2005).

Given a set of relevant variables that explain teacher qualifications and SES, I used a continuous latent trait IRT approach to build continuous indexes. Item response theory (IRT) models are designed to relate a subject's response on each of the items of a questionnaire to a single underlying latent trait (Kolen and Brennan 2004; Stout 2007). In this study teachers' responses to each item were intended to provide a piece of information about their qualifications, and students' responses to each item provided information about their SES. If these items are capturing the same latent trait (unidimensionality assumption), then it is possible to aggregate item responses into a single continuous variable that measures the latent trait.

The first step in the latent trait modeling was to select relevant variables. Variables should capture one single latent trait. Having selected the variables, the next step was to analyze the correlations between them. Statistical procedures, such as factor analysis of the polychoric matrix, provided validation diagnosis of variables that satisfied the unidimensionality assumption.[4] I estimated the polychoric matrix using the *polycor* package available in R (Fox 2010). The usual procedure for the validation of the unidimensionality assumption for a set of chosen variables consists of verifying two conditions are met:

- All the variables should have the same direction of the biserial correlations[5] in the polychoric matrix. Discrepant variables should be dropped;
- Examination of the relative sizes of the eigenvalues associated with the principal components analysis of the polychoric matrix. A rule-of-thumb suggests that as many latent traits are present as there are eigenvalues greater than one (Loehlin 1998).

Having selected the variables according to the checks described previously, the next step was to estimate the IRT model. The Graded Response Model (GRM) proposed by Samejima (1969) is suitable for items with graded or ordinal responses, which is usually the case of teacher characteristic variables. In this paper I used the *Latent Trait Model* package (*ltm*) available in R to estimate the GRM (Rizopoulos 2006). Then, it was necessary to check the unidimensionality assumption of the IRT model. For this purpose, the Cronbach's alpha provided an adequate measure of internal consistency and unidimensionality (Sijtsma 2009). In the literature, a rule-of-thumb regarding whether the IRT latent trait model has a good fit is to verify if the Cronbach's alpha is greater than 0.80 (DeVellis 2012).

4.2.2. Individual dose-response analysis under the mixed-effects approach

Individual dose-response modeling is commonly used in epidemiology and clinical studies and provides an appropriate causal framework for understanding the relationship between dose and

response (Edler 2008; Holmgren and Koch 1997). The proof of the existence of a dose-response relationship is a basic element of the proof of causality in experimental sciences (Edler 2008). This method is widely used in toxicology and clinical studies. In this paper, this framework is applied to understand the how the pace of learning in Portuguese and Mathematics (the responses) are related to the amount of exposure to teacher quality (the dose provided).

This study addresses the problem of selection bias by estimating individual dose-response models. In this approach, the relationship between teacher quality and achievement gains is estimated for each student: during the longitudinal follow up, the student was exposed to different teachers, who may have had different qualifications. Assuming that there is no sorting mechanism based on student gains, each student was randomly exposed to teachers of different qualifications during his/her trajectory. Therefore, each student trajectory serves as his or her own natural experiment.

The basic requirements for dose-response modeling are the availability of dose-response data and an assumption in regards to the functional form of the relationship. For this study, I assume a linear function of the dose-response. This means that teacher quality does not have a threshold, that is a specific level of exposure below or above which no effect occurs (Cox 1987).

To test these hypotheses, I employ a linear mixed-effects model. Mixed-effects models are appropriate when a researcher wants to study the dependence of a response on covariates and the unexplained variation in the response (Bates, 2005). The fixed-effects coefficients represent the average rate of change in student achievement gains with respect to teacher qualifications. Also, it is possible to account for the effect of time-invariant covariates to induce changes in the response of the time-varying covariate by the inclusion of random-effects terms.

Under the mixed-effects approach, I estimate two specifications:

Unconditional model: $g_{it} = \gamma_{00} + \gamma_{10} TQ_{it} + \epsilon_{it} + u_{0t} + u_{1t}$

Conditional model: $g_{it} = \gamma_{00} + \gamma_{10} TQ_{it} + \gamma_{11} TQ_{it} \times SES_{it} + (\epsilon_{it} + u_{0t} + u_{1t} \times SES_{it})$

Where g_{it} is the gain in equalized test scores for student *i* between time *t* and time *t-1*; TQ_{it} refers to the teacher qualifications index of student *i* at time *t*; SES_{it} refers to the SES index of student *i* at time *t*; ϵ_{it} gives the residual variance; u_{0t} gives the variance of the intercept γ_{00}. In the unconditional model, and u_{1t} gives the variance of the slope γ_{10}. In the conditional model, this variance is given by $u_{1t} \times SES_{it}$.

The models can also be written in a multilevel form:

Unconditional Model:

Level 1: $g_{it} = \alpha_{0t} + \alpha_{1t} TQ_{it} + \epsilon_{it}$

Level 2: $\alpha_{0t} = \gamma_{00} + u_{0t}$

$$\alpha_{1t} = \gamma_{10} + u_{1t}$$

Conditional Model:

$$\text{Level 1:} \quad g_{it} = \alpha_{0t} + \alpha_{1t} TQ_{it} + \epsilon_{it}$$

$$\text{Level 2:} \quad \alpha_{0t} = \gamma_{00} + u_{0t}$$

$$\alpha_{1t} = \gamma_{10} + \gamma_{11} SES_i + u_{1t}$$

The first equation in each model in the multilevel form refers to the individual dose-response regression of gains on the teacher qualifications index. The second equation postulates that the intercept of the individual dose-response curve, α_{0t} -- the mean student improvement when teacher quality is equal zero, may vary across students. Finally, the third equation posits that the gradient of teacher qualifications on student gains, the slope α_{1t}, can vary across students. The unconditional model is the baseline specification, in which the increment in gains given a level of teacher quality α_{1t} and the mean student improvement when teacher quality equals zero α_{0t} are randomly distributed. The gradient of teacher qualifications for student gains, α_{1t}, is the key parameter of my interest, and γ_{11} represents the average gradient in teacher qualifications for all the students.

In the conditional model, I used SES as a predictor of change in the variability in the rate of improvement in learning given a certain level of teacher qualifications. Therefore, the conditional model allowed me to test whether context-specific effects enhance or minimize the effect of teacher qualifications on student achievement gains.

To interpret γ_{11} as the average causal effect of teacher qualifications on student gains in each grade, I assumed that students are assigned to teachers on the basis of their time-invariant characteristics (for example, race, parents' education) rather than on qualities that vary with time (such as their performance). Some authors criticize this assumption by arguing that school managers may assign students to classrooms on the basis of dynamic evaluation of student performance (Rothstein 2007; Rothstein 2009; Rothstein 2010). Because it is not possible to predict the criteria that school managers employ when assigning students, I assumed that assignments were made on the basis of students' time-invariant characteristics. However, I relaxed this hypothesis by allowing the rate of improvement to differ according to student's SES as expressed in the conditional model.

For the estimation of the mixed-effects models, I used the *lme4* package in R (Bates, Maechler, and Bolker 2011).

4.2.3. Dataset

The FUNDESCOLA longitudinal study included six rounds of data collection, following students by grade and year, starting in fourth grade in 1999 and continuing until eighth grade in

2003. In 1999, students were evaluated at the beginning of the school year (April) and at the end of the school year (November). For the next rounds, students were evaluated in November of each year (2000, 2001, 2002 and 2003). The sample included schools from six states, namely Rondônia, Pará, Pernambuco, Sergipe, Mato Grosso do Sul and Goiás, which were in the regions served by the FUNDESCOLA program. To be selected in the sample, schools needed to meet the following requirements: they were public schools; all their elementary classes were held during the day; they had at least 200 students; and they were located in the regions of the capital with administrative facilities belonging to each state.

The FUNDESCOLA survey included information on test scores in Mathematics and Portuguese and on socioeconomic characteristics of students (Student Questionnaire), teachers (Teacher Questionnaire), and schools (School Questionnaire). In regards to students' testing, items were based on those from the National Assessment of Basic Education (SAEB). During the longitudinal study, different populations were examined with different forms, but with some common items.[6] In order to ensure comparability, scores in each round were equalized using statistical equating methods.[7]

During the longitudinal follow-up, information was collected on all students present in the classrooms at the time of the application of the instruments, regardless of their participation in the previous rounds. Thus, in addition to students observed across all included grades (fourth to eighth grade), it is possible to identify students who were present in one, two, or three of the rounds.

4.2.3.1. Sample and statistical filters

A potential problem in longitudinal studies is attrition, and the FUNDESCOLA data is not an exception. The survey was carried out in six rounds, and many students dropped out of the dataset because of migration (schools or towns), grade repetition, and death. Also, students may have been missing information on test scores because they did not attend class the day of the assessment. In addition, there was a considerable non-response rate among teachers, and therefore, information gaps may be even more pronounced. In this section I present details on the sample and the filters needed to carry out my analysis.

In this study I analyzed two different datasets according to the subject evaluated: Portuguese and Mathematics. For these subjects, the number of students in the longitudinal panel was 29,778. This number reflects students who were evaluated by the FUNDESCOLA survey at least in one of the six rounds. Next, I applied filters to carry out the statistical analysis. Those are analyzed hereafter:

- The first filter applied was to select individuals who have non-missing information on test scores. Of this number, 23,288 students had information on Mathematics test scores and 24,633 on Portuguese.

- The second filter consisted in the selection of students who were evaluated in at least two adjacent rounds, as a requirement for the computation of the gains variable. 9,733 students satisfied this requirement in Mathematics and 11,476 in Portuguese.
- The third filter was applied to satisfy the dose-response model: for a within-student regression model, at least two observations on the dependent variable (gains) with no missing observation for the independent variable (gains). Hence, 2,178 students in Mathematics and 2,563 students in Portuguese were selected.
- The fourth filter was intended to keep in the sample only students who had some variability in teacher quality. If students did not have variability in the treatment condition, exposure to teacher quality, then it would not be possible to estimate the dose-response relationship. Hence, I computed a within-student indicator of the teacher qualifications variability during the longitudinal follow-up -- the range.[8] From the students who remained in the sample after the third filter, 927 students had no variability in teacher quality (42.56%) for Mathematics and 1,133 students in Portuguese (44.21%). Therefore, my final sample is 1,251 students in Mathematics and 1,430 students in Portuguese.

From the description above, it is clear that the final sample of this study has important characteristics that make it unique in relation to the students assessed by the FUNDESCOLA survey. First, they are students who survived many risks: the risks of grade retention, death, and migration. Therefore, it is necessary to investigate whether my final sample is a non-random sample from the initial one.

Students who were selected in the final sample (fourth filter) did not differ from students who had participated in at least two adjacent rounds (second filter) in terms of their gains in Mathematics and Portuguese, as shown in the kernel density curve for gains by group in Figure 4 for Mathematics and Figure 5, for Portuguese. Therefore, I conclude that students in my final sample did not have significant differences in the outcome that may bias my results.

However, students who were selected in the final sample did have in general lower scores in the SES index obtained by Item Response Theory[9] than students who had participated in at least two adjacent rounds (2nd filter). Results show that this condition is true for Mathematics (Figure 6) and Portuguese (Figure 7), since the kernel density curve for the final sample is shifted to the right compared to students selected in the second filter. This may indicate that, despite lower SES, students in my final sample had a similar pace of learning to the students who had participated in at least two adjacent rounds.

In sum, sample restrictions were necessary to carry out the dose-response model. Those restrictions seem to affect the representativeness of students for those with similar gains and lower SES conditions. This may indicate, therefore, that students in my final sample have higher

ability than their counterparts because they had similar gains distributions in a more adverse SES condition.

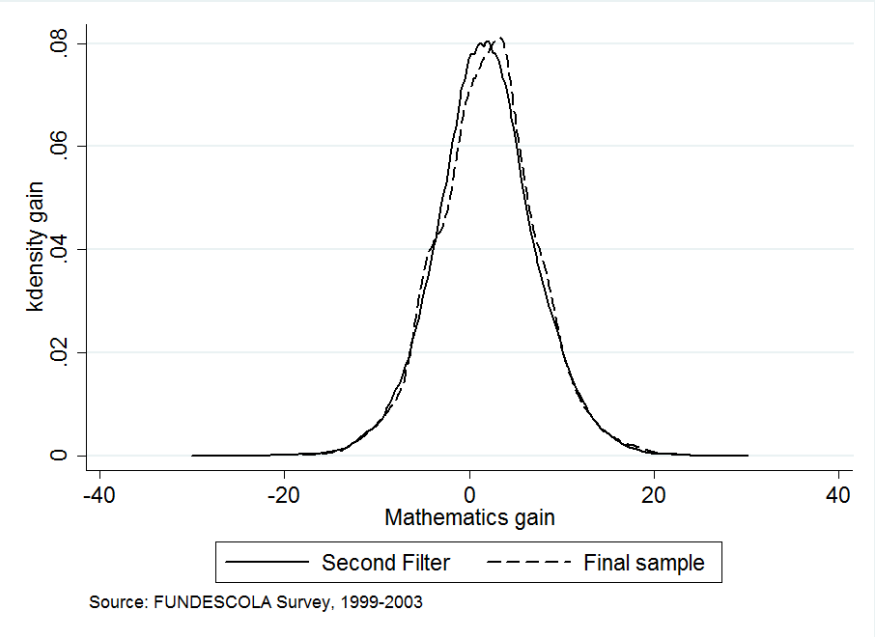

Figure 3. Kernel density estimator of the gains in Mathematics for students present in at least two adjacent tests (second filter) and students selected for the final sample (fourth filter).

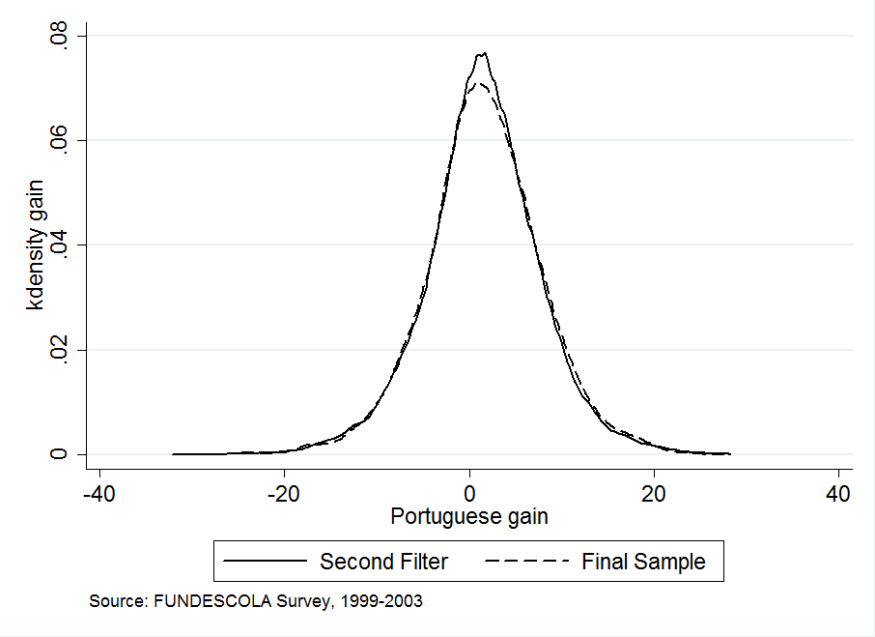

Figure 4. Kernel density estimator of the gains in Portuguese for students present in at least two adjacent tests (second filter) and students selected for the final sample (fourth filter).

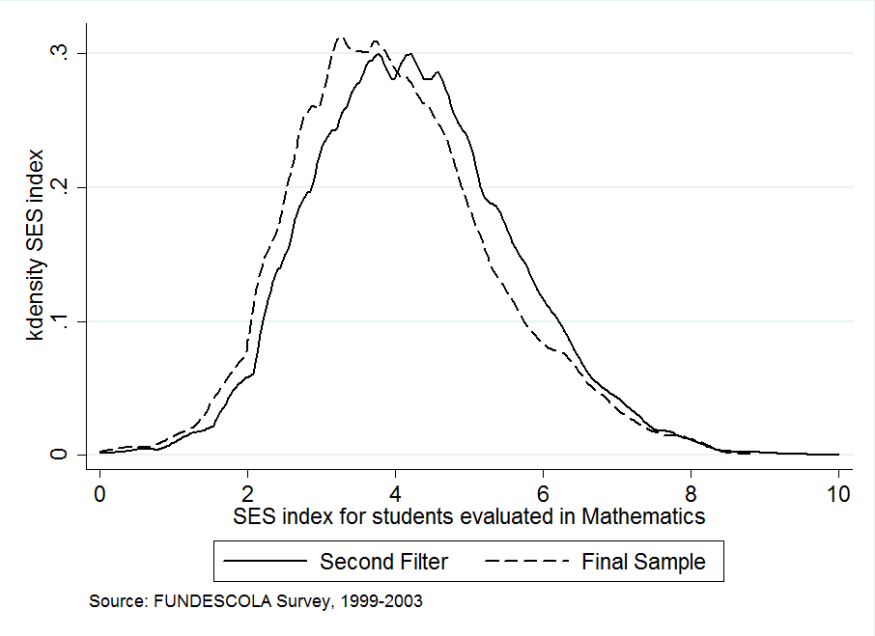

Figure 5. Kernel density estimator of the SES index for students present in at least two adjacent tests (second filter) and students selected for the final sample (fourth filter) in Mathematics

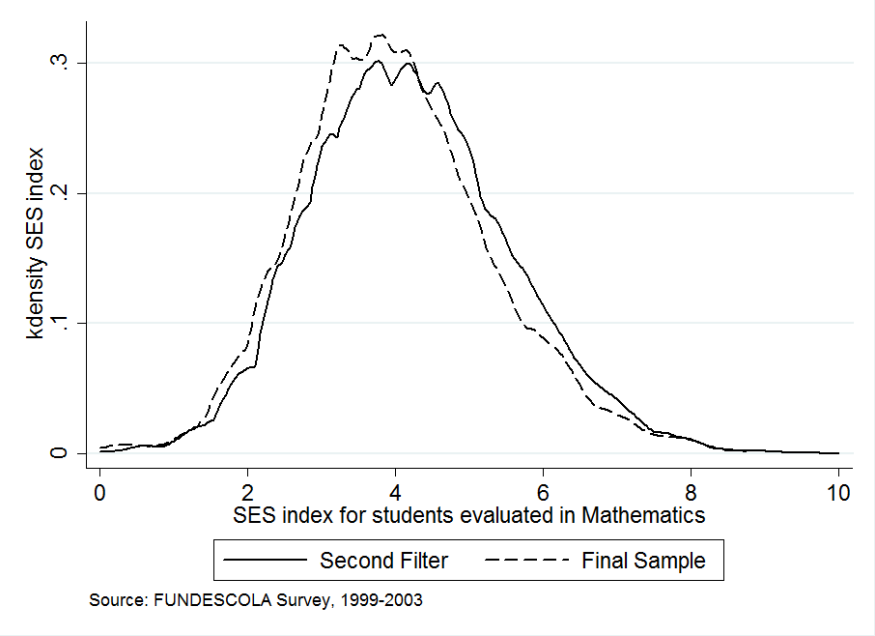

Figure 6. Kernel density estimator of the SES index for students present in at least two adjacent tests (second filter) and students selected for the final sample (fourth filter) in Portuguese

5. Findings

In this section I present the key findings of this paper. I first show results of the estimation of the latent variables -- teacher qualifications and SES index -- using Graded Response Models under the Item Response Theory approach. Then, I present the results of the dose-response model for the effect of teacher qualifications on the rate of improvement in Portuguese and Mathematics.

5.1. Latent trait models using Item Response Theory

I first discuss the results of the stepwise procedure for selecting the variables to be included in the teacher qualifications and socioeconomic indexes. In order to be selected for the preliminary model, variables needed to satisfy the single latent trait assumptions. First, all the variables needed to have the same direction as the biserial correlations.[10] The majority of the variables were positive, so variables with negative correlation coefficients were dropped. Next, the indicator needed to be indicative of a single latent trait, and the decomposition of the polychoric matrix of teacher qualifications proximates had to indicate that the first eigenvalue was greater than one and the second eigenvalue was less than zero.

For the teacher qualifications index, the first step included six variables present in the questionnaire that indicated teacher qualifications: highest level of schooling, type of higher education institution attended, postgraduate studies, professional development, experience and certification. Of these variables, type of higher education institution attended, postgraduate studies and professional development were dropped because of negative biserial correlations. Thus, the final IRT baseline model included three variables: highest level of schooling, experience and certification, which also had the first eigenvalue of the polychoric matrix greater than one. Table 3 presents the categories of each variable included in the IRT baseline model for the teacher qualifications trait.

Table 2: Variables and categories selected for the preliminary Item Response Theory Model: Teacher qualifications trait

Variable	Label	Category	Description
Highest level of schooling	hlschooling	1	Not completed elementary/Completed elementary
		2	Completed high school
		3	Completed college
Experience	experience	1	Less than 2 years
		2	2-5 years
		3	6-10 years
		4	11-15 years
		5	16-20 years
		6	More than 20 years
Certification	certification	1	No
		2	Yes

Notes: Source is FUNDESCOLA Survey, 1999-2003.

For the SES index, the first step included twelve variables that indicated the ownership of assets in the household. All the variables indicated positive biserial correlations, and the first eigenvector of the polychoric matrix was greater than one. Thus, all the variables were included in the IRT baseline model for SES. Table 4 presents the categories of each variable included in the IRT baseline model for the SES trait.

Having selected relevant variables for the baseline models, I estimated IRT latent trait models for teacher qualifications and SES. The model for SES with twelve variables presented Cronbach's alpha of 0.90. Therefore, this was considered the final model for the estimation of the SES index.

As described above, the baseline IRT model for the teacher qualifications variable was composed of three variables: highest level of schooling, experience and certification. In the baseline, the Cronbach's alpha for the Graded Response Model was 0.41. Stepwise procedures indicated that experience should be removed from the index. However, as the literature predicts that experience is an important indicator for teacher qualifications, I created a dichotomous variable for experience, which equals one if the teacher had six or more years of experience and zero otherwise. This new graded response model, which included a dummy variable for experience, certification and highest level of schooling, had a Cronbach's alpha of 0.67, still not adequate according to the literature. Then, I excluded experience, leaving the final graded response model for teacher qualifications to include certification and highest level of schooling, whose Cronbach's alpha was 0.80.

Table 3: Variables and categories selected for the preliminary Item Response Theory Model: SES trait

Variable	Label	Category	Description
How many bathrooms do you have in your home?	nban	0	Do not have
		1	Have one
		2	Have two or more
How many radios do you have in your home?	nrad	0	Do not have
		1	Have one
		2	Have two or more
How many TVs do you have in your home?	ntv	0	Do not have
		1	Have one
		2	Have two or more
How many VCRs do you have in your home?	nvcr	0	Do not have
		1	Have one
		2	Have two or more
How many refrigerators do you have in your home?	ngela	0	Do not have
		1	Have one
		2	Have two or more
How many washers do you have in your home?	nmqlv	0	Do not have
		1	Have one
		2	Have two or more

How many vacuum cleaners do you have in your home?	naspo	0	Do not have
		1	Have one
		2	Have two or more
How many phones do you have in your home?	ntel	0	Do not have
		1	Have one
		2	Have two or more
How many mobiles do you have in your home?	ncel	0	Do not have
		1	Have one
		2	Have two or more
How many PCs do you have in your home?	npc	0	Do not have
		1	Have one
		2	Have two or more
How many cars do you have in your home?	ncar	0	Do not have
		1	Have one
		2	Have two or more
How many housemaids do you have in your home?	nedom	0	Do not have
		1	Have one
		2	Have two or more

Notes: Source is FUNDESCOLA Survey, 1999-2003.

After generating the final models for teacher qualifications and SES, I computed the scores for each variable, which provided continuous variables for each trait. As the scores ranged from negative to positive values, I rescaled each variable so their values were in the 0-10 range, where zero indicates the lowest score (lowest SES, lowest teacher qualifications) and 10 the highest (highest SES, highest teacher qualifications).

In order to provide an understanding of the distribution of the IRT scores, I first present the interpretation of the teacher qualification index in Table 5 and its absolute and relative frequency in my final sample for Portuguese and Mathematics. It shows that teachers with low scores in the teacher qualification index tended, for instance, not to have completed elementary level or are elementary graduates with no teacher certification. On the other hand, teachers with the highest scores had college degrees and certification. Also, the relative frequency of teachers in regards to teacher qualification was similar for both subjects: the category most prevalent in the final sample was the one with the highest score category (college degree and certified).

Table 4: IRT Score for teacher qualification index and descriptive statistics

Qualification	IRT Score	Portuguese		Mathematics	
		Absolute Frequency	Relative Frequency (%)	Absolute Frequency	Relative Frequency (%)
Elementary graduate/Not completed elementary level and not certified	0.00	105	2.38	138	3.64

High School graduate level and not certified	0.04	523	11.86	430	11.35
Missing information on schooling and not certified	0.17	26	0.59	34	0.9
College graduate and not certified	0.23	1,221	27.69	997	26.33
High School graduate and missing information for certification	5.19	36	0.82	45	1.19
Missing information on schooling and certification	5.74	42	0.95	28	0.74
College graduate and missing information for certification	5.94	36	0.82	30	0.79
Elementary graduate/Not completed elementary level and certified	9.73	87	1.97	108	2.85
High School graduate level and certified	9.79	811	18.39	670	17.69
Missing information on schooling and certified	9.93	112	2.54	76	2.01
College graduate and certified	10.00	1,411	32.00	1,231	32.51

Notes: Source is FUNDESCOLA Survey, 1999-2003.

As shown in Figure 5 presented previously, the SES index distribution was approximately normal [0,10].

5.2. Individual dose-response analysis

In this section I first present descriptive statistics for my final sample. As described in the section 5.2.3.1, the final sample size was 1,251 students in Mathematics and 1,430 students in Portuguese. The total number of observations was 3,787 in Mathematics and 4,410 in Portuguese. The sample presents information on 431 Mathemattics and 427 Portuguese teachers.

Table 6 presents information on student achievement gains from students in each school transition from fourth through eighth grade. Gains are based on equalized scores, so the between-grade growth has a meaningful interpretation. As shown in the table, FUNDESCOLA students in my final sample gained an average of 2 points per school transition in Mathematics and 1.6 in Portuguese. On average, students had the highest gains in their scores in sixth grade for Mathematics (3.68) and fifth grade for Portuguese (3.37), whereas they made their lowest gains in eighth grade for Mathematics (1.07) and fourth grade for Portuguese (-0.33).

Table 5: Descriptive statistics for the student achievement gains in Portuguese and Mathematics by school transition

School transition	Reference date	Mathematics			Portuguese		
		Mean	Std. Dev.	Abs. Freq.	Mean	Std. Dev.	Abs. Freq.
4th grade	April/1999-Nov/1999	1.51	4.84	79	-0.33	4.85	387
5th grade	Nov/1999-Nov/2000	2.48	5.20	183	3.37	6.03	232
6th grade	Nov/2000-Nov/2001	3.68	5.31	536	3.15	6.07	446
7th grade	Nov/2001-Nov/2002	1.29	4.74	708	0.55	6.72	813
8th grade	Nov/2002-Nov/2003	1.07	5.81	481	2.50	6.17	563
	Total	*2.00*	*5.32*	*1987*	*1.60*	*6.29*	*2441*

Note: Source is FUNDESCOLA Survey, 1999 to 2003

I present in Table 7 the average gain and standard deviation in Portuguese and Mathematics according to the level of the teacher qualification index. The analysis of this table allows us to speculate about sorting mechanisms based on gains, that is, whether students who had high or low gains were assigned to teachers with a specific qualification index. As is revealed by the table, there is not a clear pattern of increasing or decreasing gains as the teacher qualifications index increases or decreases. Therefore, I have no evidence that there is a non-random sorting mechanism of students to teachers based on their achievement gains.

Table 6: Descriptive statistics for the student achievement gains in Portuguese and Mathematics according to the level of the teacher qualification index

Teacher qualification index score	Mathematics		Portuguese	
	Average gain	Std. Deviation	Average gain	Std. Deviation
0.00	3.03	4.52	2.49	5.71
0.04	1.49	4.88	0.86	4.91
0.17	2.91	4.22	-0.06	8.39
0.23	1.38	5.36	1.48	6.23
5.19	-0.73	5.73	1.12	5.55
5.74	1.57	6.81	1.25	4.32
5.94	1.77	3.52	0.39	6.86
9.73	4.66	5.22	4.99	6.51
9.79	2.22	5.49	-0.14	5.72
9.93	2.21	3.60	-0.02	7.05
10.00	2.16	5.35	2.24	6.52
Total	*2.00*	*5.32*	*1.60*	*6.29*

Note: Source is FUNDESCOLA Survey, 1999 to 2003

Despite some evidence of random assignment of students to teachers in my sample, the individual dose-response model is still preferred because it has a clear and rigorous identification strategy for the estimation of causal effects. By assuming random variation in teacher assignment to learning gains by student, this method simulates a natural experiment and provides strong evidence for the effect of teacher qualifications on the pace of learning.[11] I expect that the higher the teacher quality, the higher the pace of learning for the student.

The results of the individual dose-response model with mixed-effects for the rate of improvement in Portuguese and Mathematics in the FUNDESCOLA sample are presented next. As described in section 5.2.2, I estimated two specifications. The unconditional model is the baseline specification, in which the gradient of teacher qualifications for student gains (α_{1t}) and the mean student improvement when teacher qualifications are equal to zero (α_{0t}) were randomly distributed. In the conditional model, students' SES is a predictor of changes in the gradient of teacher qualifications in the pace of learning.

Results of the conditional and unconditional model are presented in Table 8. For Mathematics, the effect of teacher qualifications on the average improvement in learning $\hat{\gamma}_{10}$ was found to be positive and statistically significant in the unconditional specification: a one-unit increase in the teacher qualifications index increased student gains by 0.07 points. This is a large effect, remembering that the average student gain per school transition was 2. The average gain in Mathematics for students exposed to teachers with qualification index equal zero[12] was 1.6. In the conditional model, where I controlled for the influence of SES in enhancing the effect of teacher qualifications, $\hat{\gamma}_{10}$ was positive but not statistically significant. Also, there is no evidence that SES enhanced the effect of teacher qualifications on student achievement gains, as the coefficient for the interaction between SES and teacher qualifications $\hat{\gamma}_{11}$was found to be not significant.

For Portuguese, the effect of teacher qualifications on the average improvement in learning $\hat{\gamma}_{10}$was found to be positive but not statistically significant in the unconditional specification: a one-unit increase in the teacher qualifications index increased student gains by 0.04 points. The average gain in Portuguese for a student exposed to a teacher with qualification index equal zero[13] was 1.4. In the conditional model, $\hat{\gamma}_{10}$ was positive (0.04) but again not statistically significant. Also, there is no evidence that SES enhanced the effect of teacher qualifications on student achievement gains in Portuguese, as $\hat{\gamma}_{11}$was found to be not significant.

Table 7: Results of the linear mixed-effects model. Dependent variable: Student Achievement Gains.

	Unconditional Model		Conditional Model	
	Mathematics	Portuguese	Mathematics	Portuguese
Teacher Qualification Index	0.071** (0.026)	0.038 (0.027)	0.090 (0.061)	0.044 (0.053)
Teacher Qualification Index * SES	-	-	-0.013 (0.014)	0.007 (0.012)

Intercept	1.593*** (0.188)	1.394*** (0.198)	1.394*** (0.198)	1.594*** (0.188)
Variance of the intercept	2.42	2.36	2.42	2.39
Variance of the slope	0.09	0.10	0.09	0.10
Residual variance	26.16	37.18	26.17	37.14
Deviance	12267.79	15897.74	13657	16104
AIC	12287.7	15917.39	13680	16128
BIC	12321.27	15952.19	13718	16167
Number of observations	1,987	2,441	1,696	2,009
Number of students	1,182	1,361	1,182	1,361

Note: Source is FUNDESCOLA Survey, 1999 to 2003. *** p<0.01 ** p<0.05 *p<0.10

In sum, this study shows that high levels of teacher qualification increased the pace of learning in Mathematics. For Portuguese, results were not statistically significant. However, it should be emphasized that, by using gains as dependent variable, my estimates are biased downward (Clotfelter, Ladd, and Vigdor 2007a). Therefore, the effects might be higher than those that I found, strengthening the relevance of teacher qualifications to the pace of learning.

This analysis extends research on the influence of teacher qualifications on student achievement in Brazil in three ways. First, it draws on a rich longitudinal survey, the FUNDESCOLA data, which allows for a precise estimation of teacher effectiveness. Second, it introduces a single latent trait measure to capture teacher qualifications, a measure of effectiveness, under the Item Response Theory approach. Teacher qualifications were defined by credentials indicators: highest level of schooling and certification. Finally, this paper employs an individual dose-response model with mixed-effects. This model worked in two steps: first, it estimated the relationship between teacher qualifications and achievement gains of an individual student from fourth to eighth grade; next, results for all the students were pooled to assess the average effect of teacher qualifications. In addition, I considered an alternative specification for the model which included a measure of the interaction between students' SES and teacher qualifications. This alternative specification tests whether SES would enhance teacher effectiveness in the learning process, serving as a context-specific catalyst for the interaction between teachers and students.

6. Discussion

During their school careers, students may be exposed to teachers of different quality. In this study I explore how students' learning gains respond to different levels of teacher qualifications. In this section I discuss the results of this paper. First, I explore the difference found in this study in the significance of teacher qualifications on student gains according to the subject evaluated: the statistically insignificant influence of teacher qualifications on Portuguese gains, and the

significant and positive effect of teacher qualifications on Mathematics gains. Second, I discuss the lack of significance of SES in enhancing teacher effectiveness in the learning process. Finally, I conclude by explaining that this study's findings in some instances confirm and in others contradict evidence from the pre-existing literature on teacher qualifications and by arguing that considering the relationship between the current study and earlier findings can provide valuable contributions to public policy.

This study shows that teacher qualifications matter for enhancing learning gains in Mathematics for students evaluated in FUNDESCOLA data. This finding is statistically significant but not robust to specifications that include and exclude the interaction between SES and teacher qualifications. This result confirms the literature on the relevance of teacher credentials for student learning (Clotfelter, Ladd, and Vigdor 2007a; Clotfelter, Ladd, and Vigdor 2007c; Ladd 2008) and supports the Brazilian literature on the important impact of teacher characteristics on student achievement (Fonseca 2011; Menezes-Filho and Pazello 2007). On the other hand, the results show that the parameters are positive but not statistically significant for Portuguese, suggesting that teacher qualifications do not influence the rate of learning gains for this subject.

Before speculating about the non-significance of the teacher qualifications index in regards to the pace of learning, it should be stressed that in this paper my estimates are biased downward, as there is no perfect correlation between initial and final test scores. Hence, true effects may be larger than those I was able to capture.

However, this differential pattern of teacher effects may support the literature that argues that teacher effectiveness may differ according to the subject taught (Campbell et al. 2003; Muijs et al. 2005). These differences may be related to factors such as the hierarchical nature of the subject as well as to different forms and kinds of representation of ideas (McDiarmid, Ball, and Anderson 1989, 88:10). Therefore, Mathematics and Portuguese might require different models of teacher effectiveness. It could be the case that Portuguese credentials are not enough to ensure that a teacher will be effective at enhancing student gains, and therefore this may explain the insignificance of the parameters for this subject. Other unobservable variables, such as the instruction procedure, content knowledge and attitude may play a crucial role in determining teacher effectiveness for Portuguese (Hill, Rowan, and Ball 2005). Unfortunately, such data were not available in the FUNDESCOLA dataset, so I was not able to test this hypothesis.

This paper shows that there is no evidence that students' SES enhances the effect of teacher qualifications on the rate of improvement in learning in Portuguese and Mathematics. This finding contradicts the argument proposed by Brophy (1986) that heterogeneity in students' learning demands according to their SES would require a different response by teachers, resulting in a differentiated effect of a given teacher according to the students' socioeconomic background. In other words, teacher qualifications are particularly important for students less likely to overcome bad instruction. This unclear result may be explained, first, because FUNDESCOLA students may have poor socioeconomic background in general and because

teachers do not need to cater to their individual needs. Therefore, differences observed in SES may not induce differences in demands and expectations by these students and, therefore, do not require different levels of teacher qualifications. Second, it is possible that the teacher traits captured by my teacher qualifications index do not enable teachers to identify and meet students' differentiated expectations according to SES groups.

In the interpretation of the results of this paper, it is worthwhile to mention that this study faces some limitations. First, my final sample is relatively small compared to the initial population of study, and this limits the generalizability of my results. Second, the estimate for the effect of teacher qualifications on student achievement gains may be biased if students are not assigned at random to teachers of different quality. For instance, if students are assigned to a certain level of teacher qualifications based on their gains, then the effect estimated in this paper is not causal. However, it is unlikely that principals assign teachers based on individual students' performance. It is more likely that principals do assign teachers to classrooms according to some criteria, but it would not be feasible to assign based on individual performance. Therefore, I believe that this limitation is unlikely to strongly affect my results.

Despite having found mixed results for the effect of teacher qualifications on student learning gains according to the subject taught, I argue that this paper suggests that highly-qualified teachers are an important school factor in determining student learning in Brazil. The use of individual dose-response analysis as an identification strategy for assessing the effect of teacher qualifications on student achievement gains is robust, although it provides some downward bias. Hence, even if my results show that teacher qualifications are not statistically associated with student learning gains in Portuguese, it does not mean that teachers do not matter for the pace of learning in this discipline.

Moreover, the positive and significant result of the effect of teacher qualifications on the rate of improvement in Mathematics learning represents strong evidence of teacher effectiveness for public school students at the elementary level. Also, because my estimates are biased downward, this effect may be larger. Therefore, this evidence suggests the important role of teachers in inducing learning gains in poor areas of Brazil, in this study represented by the states evaluated by the FUNDESCOLA survey.

7. Conclusion

The relationship between teacher quality and student achievement gains may be very obvious to a non-specialist. However, the task of assessing the causal effect of teacher quality on the rate of student learning is not easy. There is a boom of papers and research that try to infer the real impact of teacher quality on student learning using different research designs, statistical methods and datasets. Despite the mixed evidence for the influence of observable teacher characteristics on student learning gains, there is consensus that teachers matter for student achievement.

Drawing on the education production function framework, a rich longitudinal data, and a causal framework -- dose-response modeling -- I analyze the influence of teacher qualifications,

as a proxy for teacher quality, on the rate of improvement in Portuguese and Mathematics for students evaluated from fourth to eighth grade. I measure teacher qualifications as a latent trait composed by teacher credentials: highest level of schooling and certification. By employing an individual dose-response and assuming that the assignment of students to teachers did not occur in FUNDESCOLA schools based on students' time-varying characteristics, I could identify the causal effect of teacher qualifications on changes in students' scores.

My study strongly contributes to the literature on teacher effectiveness in Brazil. The causal inference about the influence of teacher qualifications on student learning gains would be impossible using cross-sectional data and standard regression techniques. This research advances the literature by employing a rich longitudinal dataset, the FUNDESCOLA survey, and a novel empirical strategy. Dose-response analysis is estimated under a mixed-effects approach and allows us to characterize the dependence of the pace of improvement in Portuguese and Mathematics on teacher qualifications. This estimation has only been possible after the development and availability of computational algorithms, for instance, the *lme4* package in R, which was used in this study (Bates, Maechler, and Bolker 2011).

The findings of this study confirm the Brazilian literature on teacher effectiveness (Fonseca 2011; Menezes-Filho and Pazello 2007) that teachers matter for student achievement in Brazil. My study reveals that teacher qualifications positively influenced the rate of learning in Mathematics.

The results of this paper also challenge existing literature by showing that the effect of teacher qualifications was not significant for Portuguese. Additionally, this paper challenges studies that explore the role of SES in shaping student demands for teacher qualifications (Brophy 1986; McDiarmid, Ball, and Anderson 1989): I found no evidence that students' SES enhances the effect of teacher qualifications on the rate of improvement in learning in Portuguese and Mathematics. This finding, however, may reflect the fact that socioeconomic conditions of students in schools evaluated by FUNDESCOLA survey are in general poor and thus are unable to enhance the learning process given teacher qualifications. Further research in these areas is greatly encouraged to provide insights about the role of social origins in shaping learning outcomes.

The results of this study can significantly influence educational policy in Brazil aimed at improving the educational achievement indicators of public school students at the elementary level. By showing that teacher qualifications positively influence students' pace of learning, I recommend that greater resources should be allocated to improving and enhancing teacher quality in Brazilian public schools, for instance, by attracting and retaining highly-qualified teachers through a restructuring policy of positions and salaries.

In sum, Brazil has universalized access to primary education, but its achievement indicators are extremely poor. In order to overcome this situation and ensure every child the right to learn, this study reveals that strong investments in teacher qualifications may represent the most effective resource.

Notes

1. Unless otherwise stated, all translations are my own. The program was designed in accordance with Loan Agreement Number 4311-BR.
2. The survey was first carried out by the Brazilian National Institute of Educational Statistics (INEP) and later by the Center for Regional and Development Planning (CEDEPLAR). In the earlier stages of the survey, schools were divided into treatment and control groups to provide impact evidence. These findings were presented by Carnoy et. al. (2008) and show that the program had a positive impact on school spending and student outcomes. However, during the longitudinal follow-up, all the schools ended up participating in the PDE program, which prevented further impact evaluations.
3. Elementary-level teachers were identified using compatible occupation codes in the PNAD series. For a description of the compatibility procedures, see Dias (2008).
4. The polychoric correlation matrix is an advancement of the Pearson correlation matrix and assumes that the response categories are proxies for unobserved and normally distributed variables.
5. Used to measure the association between an item and the test.
6. For the FUNDESCOLA data, ability differences for students taking different forms are confounded by differences in test difficulty. This is so called non-equivalent groups design (Kolen and Brennan 2004).
7. Statistical equating defines a functional relationship between multiple test score distributions. For the non-equivalent groups design, common items are also required. Statistical equating methods have some advantages over IRT equating methods because they require fewer assumptions (Kolen and Brennan 2004). I equalized the scores using the equipercentile equating under the chained method. Computations were conducted using the equate package in R (Albano 2011).
8. The range is defined by the difference between the maximum and the minimum.
9. More details about the construction of this index are provided in the Findings section.
10. Used to measure the association between an item and the test.
11. Note that this assumption may be violated as I discussed in the literature review and conceptual framework. However, the majority of empirical work also assumes that students are randomly sorted across teachers, and because I am not able to observe the mechanisms by which students are assigned to teachers, it is reasonable to assume they were assigned at random.
12. That is, Elementary graduate/Not completed elementary level and not certified.
13. Elementary graduate/Not completed elementary level and without teacher certification

References

Aaronson, D., L. Barrow, and W. Sander. 2007. Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics* 25,: 95–135. doi:10.1086/508733.

Albano, A. 2011. *Equate: statistical methods for test score equating*. http://CRAN.R-project.org/package=equate.

Barro, R., and J. Lee. 1993. International comparisons of educational attainment. *Journal of Monetary Economics* 32, no. 3: 363–394. doi:10.1016/0304-3932(93)90023-9.

Bates, D. 2005. Fitting linear mixed models in r. *R News* 5, no. 1: 27–31.

Bates, D., M. Maechler, and B. Bolker. 2011. *Lme4: linear mixed-effects models using S4 classes*. http://CRAN.R-project.org/package=lme4.

Behrman, J. 1987. Schooling in developing countries: which countries are the over- and underachievers and what is the schooling impact? *Economics of Education Review* 6, no. 2: 111–127. doi:10.1016/0272-7757(87)90045-8.

Benhabib, J., and M. Spiegel. 1994. The role of human capital in economic development evidence from aggregate cross-country data. *Journal of Monetary Economics* 34, no. 2: 143–173. doi:10.1016/0304-3932(94)90047-7.

Boyd, D. J., P. Grossman, H. Lankford, S. Loeb, and J. Wyckoff. 2009. Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis* 31,: 416–440. doi:10.3102/0162373709353129.

Boyd, D., H. Lankford, S. Loeb, J. Rockoff, and J. Wyckoff. 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* 27, no. 4: 793–818. doi:10.1002/pam.20377.

Brasil. 1988. Constitution of the federative republic of Brazil [In Portuguese.].

___. 1996. Law no. 9394: establishes the guidelines and bases for national education [In Portuguese.].

___. 2009. *Decree no. 6755* [In Portuguese.].

Brophy, J. 1986. Teacher influences on student achievement. *American Psychologist* 41, no. 10: 1069–1077. doi:10.1037/0003-066X.41.10.1069.

Campbell, R. J., L. Kyriakides, R. D. Muijs, and W. Robinson. 2003. Differential teacher effectiveness: towards a model for research and teacher appraisal. *Oxford Review of Education* 29, no. 3: 347–362. doi:10.1080/03054980307440.

Carnoy, M. 2011. The low achievement trap: comparing schooling in Botswana and South Africa. (mimeo). Stanford University.

Carnoy, M., A. Gove, S. Loeb, J. Marshall, and M. Socias. 2008. How schools and students respond to school improvement programs: the case of Brazil's PDE. *Economics of Education Review* 27, no. 1: 22–38.

Chetty, R., J. N. Friedman, and J. E. Rockoff. 2012. The long-term impacts of teachers: teacher value-added and student outcomes in adulthood. *National Bureau of Economic Research Working Paper Series* No. 17699. http://www.nber.org/papers/w17699.

Clotfelter, C., H. Ladd, and J. Vigdor. 2007a. *How and why do teacher credentials matter for student achievement? National Bureau of Economic Research Working Paper Series*. No. 12828

___. 2007b. *Teacher credentials and student achievement in high school: a cross-subject analysis with student fixed effects*. *National Bureau of Economic Research Working Paper Series*. No. 13617.

Clotfelter, C., L. Ladd, and J. Vigdor. 2007c. Teacher credentials and student achievement: longitudinal analysis with student fixed effects. *Economics of Education Review* 26, no. 6: 673–682. doi:10.1016/j.econedurev.2007.10.002.

Cox, C. 1987. Threshold dose-response models in toxicology. *Biometrics* 43, no. 3: 511–523. doi:10.2307/2531991.

Darling-Hammond, L. 2006. Securing the right to learn: policy and practice for powerful teaching and learning. *Educational Researcher* 35, no. 7: 13–24. doi:10.3102/0013189X035007013.

Desimone, L., and D. Long. 2010. Teacher effects and the achievement gap: do teacher and teaching quality influence the achievement gap between black and white and high- and low-SES students in the early grades? *Teachers College Record* 112, no. 12: 3024–3075.

DeVellis, R. 2012. *Scale development : theory and applications*. 3rd ed. Los Angeles :: SAGE,.

Dias, P. 2008. Mismatch between education and occupation: over-schooling or increasing eligibility requirements? [In Portuguese.]. Belo Horizonte: Universidade Federal de Minas Gerais.

Edler, L. 2008. Dose–response analysis. In *Encyclopedia of Quantitative Risk Analysis and Assessment*. John Wiley & Sons, Ltd. http://onlinelibrary.wiley.com/doi/10.1002/9780470061596.risk0264/abstract.

Eide, E., D. Goldhaber, and D. Brewer. 2004. The teacher labour market and teacher quality. *Oxford Review of Economic Policy* 20, no. 2: 230 –244. doi:10.1093/oxrep/grh013.

Fonseca, I. 2011. Teacher characteristics and differences in math and language achievement in Brazilian public schools. Masters monograph, Stanford: Stanford University.

Fox, J. 2010. *Polycor: polychoric and polyserial correlations*. http://CRAN.R-project.org/package=polycor.

Goldstein, H. 2001. Using pupil performance data for judging schools and teachers: scope and limitations. *British Educational Research Journal* 27, no. 4: 433–442. doi:10.1080/01411920120071443.

Hanushek, E. 1979. Conceptual and empirical issues in the estimation of educational production functions. *The Journal of Human Resources* 14, no. 3: 351–388. doi:10.2307/145575.

Hanushek, E., J. Kain, D. O'Brien, and S. Rivkin. 2005. The market for teacher quality. *National Bureau of Economic Research Working Paper Series* No. 11154,. http://www.nber.org/papers/w11154.

Hanushek, E., and S. Rivkin. 2006. Teacher quality. In *Handbook of the Economics of Education*, ed. E. Hanushek, S. Machin, and L. Woessmann, Volume 2:1051–1078. Elsevier. http://www.sciencedirect.com/science/article/pii/S1574069206020186.

Hanushek, E., and L. Woessmann. 2008. The role of cognitive skills in economic development. *Journal of Economic Literature* 46, no. 3: 607–668.

Hedges, L., R. Laine, and R. Greenwald. 1994. An exchange: part I: Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher* 23, no. 3: 5 –14. doi:10.3102/0013189X023003005.

Hill, H., B. Rowan, and D. Ball. 2005. Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal* 42, no. 2: 371 –406. doi:10.3102/00028312042002371.

Holmgren, E., and G. Koch. 1997. Statistical modeling of dose-response relationships in clinical trials— a case study. *Journal of Biopharmaceutical Statistics* 7, no. 2: 301–311. doi:10.1080/10543409708835188.

Kolen, M., and R. Brennan. 2004. *Test equating, scaling, and linking : methods and practices*. 2nd ed. New York: Springer.

Krueger, A. 1999. Experimental estimates of education production functions. *Quarterly Journal of Economics* 114, no. 2: 497–532. doi:10.1162/003355399556052.

Ladd, H. 2008. Teacher effects: what do we know? In *Teacher Quality: Broadening and Deepening the Debate*, ed. G. Duncan and J. Spillane, 3–26. Northwestern University.

Loeb, S. 2001. Teacher quality: its enhancement and potential for improving pupil achievement. In *Improving Educational Productivity*, ed. D. Monk, H. Walberg, and M. Wang, 254.

Loeb, S., and M. Page. 2000. Examining the link between teacher wages and student outcomes: the importance of alternative labor market opportunities and non-pecuniary variation. *Review of Economics and Statistics* 82, no. 3: 393–408. doi:10.1162/003465300558894.

Loehlin, J. 1998. *Latent variable models: an introduction to factor, path, and structural analysis*. 3rd ed. Mahwah, NJ: Lawrence Erlbaum.

Louzano, P., V. Rocha, G. Moriconi, and R. Oliveira. 2008. *Who wants to be a teacher in Brazil? are teachers being well prepared for the classroom? Attractiveness, selection and training*

of teachers in Brazil. Report elaborated for the Lemann foundation and Instituto Futuro Brasil.

Luz, Luciana. 2008. The impact of grade repetition in school proficiency: a longitudinal analysis of the performance of repeaters in 2002-2003 [In Portuguese.]. Belo Horizonte: Centro de Desenvolvimento e Planejamento Regional. http://www.cedeplar.ufmg.br/demografia/dissertacoes/2008/Luciana_Soares_Luz.pdf.

Macedo, G. 2004. Factors associated with academic performance of students in the fifth grade (2000): a value added and heterogeneity approach [In Portuguese.]. Masters thesis, Belo Horizonte: Universidade Federal de Minas Gerais.

McCaffrey, D. 2003. *Evaluating value-added models for teacher accountability*. Santa Monica, CA: RAND.

McCaffrey, D., J. Lockwood, D. Koretz, T. Louis, and L. Hamilton. 2004. Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics* 29, no. 1: 67 –101. doi:10.3102/10769986029001067.

McDiarmid, G., D. Ball, and C. Anderson. 1989. *Why staying one chapter ahead doesn't really work: subject-specific pedagogy*. Vol. 88. 6. National Center for Research on Teacher Education, Michigan State University.

Menezes-Filho, N., and E. Pazello. 2007. Do teachers' wages matter for proficiency? Evidence from a funding reform in Brazil. *Economics of Education Review* 26, no. 6: 660–672. doi:10.1016/j.econedurev.2007.08.003.

Monk, D. 1989. The education production function: its evolving role in policy analysis. *Educational Evaluation and Policy Analysis* 11, no. 1: 31 –45. doi:10.3102/01623737011001031.

Muijs, D., J. Campbell, L. Kyriakides, and W. Robinson. 2005. Making the case for differentiated teacher effectiveness: an overview of research in four key areas. *School Effectiveness and School Improvement* 16, no. 1: 51–70. doi:10.1080/09243450500113985.

Nye, B., S. Konstantopoulos, and L. Hedges. 2004. How large are teacher effects? *Educational Evaluation and Policy Analysis* 26, no. 3: 237 –257. doi:10.3102/01623737026003237.

OECD. 2005. *Teachers matter: attracting, developing and retaining effective teachers*. Education and training policy. OECD Publishing.

___. 2011. The PISA 2009 profiles by country/economy. http://stats.oecd.org/PISA2009Profiles/.

Oliveira, D. A. 2004. A reestruturação do trabalho docente: precarização e flexibilização. *Educação & Sociedade* 25, no. 89: 1127–1144.

Pereira, D. 2006. Factors associated with school performance in mathematics and portuguese at school: a longitudinal perspective [In Portuguese.]. Doctorate dissertation, Belo Horizonte: Universidade Federal de Minas Gerais.

Raudenbush, S. 2004. What are value-added models estimating and what does this imply for statistical practice? *Journal of Educational and Behavioral Statistics* 29, no. 1: 121 –129. doi:10.3102/10769986029001121.

Reise, S., A. Ainsworth, and M. Haviland. 2005. Item Response Theory fundamentals, applications, and promise in psychological research. *Current Directions in Psychological Science* 14, no. 2: 95–101. doi:10.1111/j.0963-7214.2005.00342.x.

Rios-Neto, E., and R. Guimaraes. 2010. The demography of education in Brazil: inequality of educational opportunities based on grade progression probability (1986-2008). *Vienna Yearbook of Population Research* 8,: 283–312. doi:10.1553/populationyearbook2010s283.

Rivkin, S., E. Hanushek, and J. Kain. 2005. Teachers, schools, and academic achievement. *Econometrica* 73, no. 2: 417–458. doi:10.1111/j.1468-0262.2005.00584.x.

Rizopoulos, D. 2006. Ltm: an R package for latent variable modelling and item response theory analyses. *Journal of Statistical Software* 17, no. 5: 1–25.

Rogosa, D. 1988. Myths about longitudinal research. In *Methodological Issues in Aging Research*, ed. K. Schaie, R. Campbell, W. Meredith, and S. Rawlings, 171–209. New York, NY, US: Springer Publishing Co.

Rothstein, J. 2007. *Do value-added models add value? Tracking, fixed effects, and causal inference*. Ceps working paper.

___. 2009. *Student sorting and bias in value added estimation: selection on observables and unobservables*. *National Bureau of Economic Research Working Paper Series*. No. 14666.

___. 2010. Teacher quality in educational production: tracking, decay, and student achievement. *The Quarterly Journal of Economics* 125, no. 1: 175 –214. doi:10.1162/qjec.2010.125.1.175.

Samejima, F. 1969. Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement* 34, no. 4, Pt. 2: 100.

Sijtsma, K. 2009. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika* 74, no. 1: 107–120. doi:10.1007/s11336-008-9101-0.

Soares, J. F. 2006. Measuring cognitive achievement gaps and inequalities: the case of Brazil. *International Journal of Educational Research* 45, no. 3: 176–187. doi:10.1016/j.ijer.2006.11.009.

___. 2007. Improvement of cognitive performance of elementary school students [In Portuguese.]. *Cadernos De Pesquisa* 37,: 135 – 160.

Stout, W. 2007. Skills diagnosis using IRT-based continuous latent trait models. *Journal of Educational Measurement* 44, no. 4: 313–324. doi:10.1111/j.1745-3984.2007.00041.x.

Stuart, E., and D. Rubin. 2008. Best practices in quasi-experimental designs. In *Best Practices in Quantitative Methods*, ed. J. Osborne, 155–176. Sage Publications.

Todd, P., and K. Wolpin. 2003. On the specification and estimation of the production function for cognitive achievement. *The Economic Journal* 113, no. 485: F3–F33. doi:10.1111/1468-0297.00097.

De la Torre, J., and J. Douglas. 2004. Higher-order latent trait models for cognitive diagnosis. *Psychometrika* 69, no. 3: 333–353. doi:10.1007/BF02295640.

UNESCO. 2006. *Teachers and educational quality: monitoring global needs for 2015*. UNESCO Inst for Statistics.

Yen, W. 2005. The choice of scale for educational measurement: an IRT perspective. *Journal of Educational Measurement* 23, no. 4: 299–325. doi:10.1111/j.1745-3984.1986.tb00252.x.