APPLIED ECONOMICS
2020, VOL. 52, NO. 6, 568–582
https://doi.org/10.1080/00036846.2019.1650885

The impacts of teacher working conditions and human capital on student achievement: evidence from brazilian longitudinal data

Larissa da Silva Marioni[a], Ricardo Da Silva Freguglia[b] and Naercio A Menezes-Filho[c]

[a]Department of Economics, University of Southampton, Southampton, UK; [b]Department of Economics, Federal University of Juiz de Fora. Campus dos Martelos, Juiz de Fora-MG, Brazil; [c]Department of Economics, Insper Institute of Education and Research & School of Economics, Business Administration and Accounting (FEA-USP), São Paulo-SP, Brazil

ABSTRACT
We use a new longitudinal matched student-teacher data from Brazil to examine the impacts of teacher human capital and working conditions on student learning controlling for student, teacher, and student-teacher match fixed effects. We find that teacher working conditions impact students outcomes. Teachers employed in more than one school negatively affect student test scores, while those working more hours in the same school positively impact students outcomes. Additionally, teacher education, experience, and family income affect student outcomes in Portuguese but not in Mathematics. Students who change from a low to a high value-added teacher experience a significant increase in proficiency.

KEYWORDS
Education; working conditions; student-teacher match; fixed effects; GERES

JEL CLASSIFICATION
I21; I29

I. Introduction

Economists have long been concerned with the poor educational outcomes of students in developing countries, where most students in secondary schools perform poorly on international exams, such as the Program for International Student Assessment – PISA (2012). Although Chile, Brazil, and Peru registered some progress in the PISA tests from 2000–2012, they still need to do much more to decrease the gap concerning the performance of the Organization for Economic Cooperation and Development (OECD) countries. Moreover, Brazil is a very unequal country and the quality of its schools vary substantially across sectors (private versus public), across regions of the country and in rural versus urban areas. Primary schools are generally ran by the municipality, while secondary schools can be managed by the states or by the municipalities. Sometimes the same city can have both type of schools offering the same grade.

Schools can improve student performance if they have better inputs and adopt better management and teaching practices (Glewwe and Kremer 2006). It is, therefore, crucial to identify the factors that impact student learning (and the magnitude of their effects) to implement more effective policies. The main interest of the economics of education literature since the Coleman Report is to identify the parameters of the education production function (Coleman et al. 1966). Despite the importance of the family background (emphasized by Coleman), the recent availability of better data has allowed researchers to disentangle more precisely the effects of schools and teachers. Quality of education is now in the center stage of the developing economics literature.[1]

This paper examines whether teacher human capital and working conditions impact student performance even after controlling for the teacher, student, and match fixed effects. We focus on teacher characteristics such as education, training, family income, experience, tenure, multiple jobs, and hours worked. Moreover, because we observe the same teachers over time teaching for different sets of students, we can disentangle the impact of teacher characteristics from unobserved teacher quality, which is fixed over time. We also control for student fixed effects that might be correlated with teacher characteristics through sorting.

CONTACT Naercio A Menezes-Filho NaercioAMF@insper.edu.br Department of Economics, Insper Institute of Education and Research & School of Economics, Business Administration and Accounting (FEA-USP). Rua Quatá, 300, São Paulo-SP, 06704-700, Brazil

[1]Several studies show that test scores are strongly related to success in the labor market (Leibowitz 1974; Murnane, Willett, and Levy 1995; Neal and Johnson 1996; Keane and Wolpin 1997; Cameron and Heckman 1998; Hanushek and Rivkin 2006a; Cunha and Heckman 2007, 2009).

Finally, in contrast to previous studies in the literature, we also control for the fact that the productivity of above-average students may be higher when they are taught by above-average teachers. We control for this possibility by including match student-teacher fixed effects in the estimation to examine the net effect of teacher characteristics. To do so, we adapt the econometric methodology initially used to model firms and workers proposed by Woodcock (2015) to model teachers and students.

We use a new longitudinal data set that stems from the Longitudinal Study of School Generation project (*GERES – Estudo Longitudinal da Geração Escolar 2005*)[2], which follows approximately 10,000 students from four large Brazilian cities over four years. The use of longitudinal matched data to estimate teacher quality effects is a novelty in Latin America. The results show that teacher working hours positively impact student achievement in both Portuguese and mathematics exams. Teachers who teach at more than one school have a negative impact on Mathematics test scores. Teachers who hold a degree in pedagogy negatively impact their students' Portuguese test scores, while those who hold a master or doctorate positively impact learning.

Moreover, teachers with more than 15 years of experience in the profession have a positive effect on Portuguese test scores, while tenure at the same school negatively impacts student grades. Interestingly, however, teacher education, experience, and tenure do not have a significant impact on mathematics scores. We also show that a student who switches from a low value-added to a high value-added teacher improve her grade by 5.22 standard deviations in Portuguese and 6.41 standard deviations in Mathematics.

This paper contributes to several lines of the economics of education literature. First, it relates to the literature that evaluates the effects of teacher characteristics on student achievement. The literature emphasizes that teachers do matter for student learning (Rockoff 2004; Clotfelter, Ladd, and Vigdor 2006, 2007; Aaronson, Barrow, and Sander 2007; Kane and Staiger 2008; Chetty, Friedman, and Rockoff 2014b; Kirabo, Rockoff, and Staiger 2014; Kraft 2019). Parents, teachers, and educational managers alike emphasize the central role of teachers in contributing to the quality of the learning process. Teacher motivation and ability to transmit knowledge can make a difference in a classroom. However, how much teachers matter and whether their working conditions impact student outcomes continue to be a topic of debate among researchers. The impact of observed teacher human capital, such as experience, education, and salary, on student test scores remains highly controversial (Hanushek 2003; Hanushek and Rivkin 2006b; Croninger et al. 2007; Buddin and Zamarro 2009). But, there is evidence showing that students matched with a good teacher accumulate more human capital in the long run (Rockoff 2004; Hanushek and Rivkin 2010; Chetty, Friedman, and Rockoff 2014b).

The literature has traditionally used the value-added approach to identify teacher effects and it has investigated different model specifications (Todd and Wolpin 2003; Rockoff 2004; Rivkin, Hanushek, and Kain 2005; Aaronson, Barrow, and Sander 2007; Kane and Staiger 2008; Goldhaber, Cowan, and Walch 2013; Chetty, Friedman, and Rockoff 2014a; Kirabo, Rockoff, and Staiger 2014; Guarino, Reckase, and Wooldridge 2015). Rivkin, Hanushek, and Kain (2005) develop an estimator of the variance of teacher quality based on patterns of differences in performance gains within schools that avoids the problem of student sorting bias. It relies on student outcomes to evaluate the overall magnitude of the effect of the teacher, regardless of the identification and measurement of specific effects. This semi-parametric approach generates an estimate of the role of teacher quality in academic development and information on the degree to which specific factors explain the differences in teacher effectiveness.[3] We contribute to the literature by using a different estimation technique that uses matched student-teacher data in which a student-teacher match is observed over multiple years to estimate the impact of teacher human capital and working conditions on student outcomes.

[2]See Franco, Brooke, and Alves (2008) for a description of the data.
[3]Rockoff (2004) uses a random-effects meta-analysis approach to measure the variance in teacher fixed effects accounting for measurement error and identifying significant effects of teachers and experience on vocabulary and reading.

However, estimates of the impact of teacher working conditions are scant, as information about safety and commuting to the workplace is missing in most databases and reliance on teachers' self-reports has disadvantages. Teachers' perspectives about their working conditions may be affected by their performance on the job (Hanushek and Rivkin 2007). Moreover, the evidence on Latin American countries in this subject is even sparser, although the average quality of teachers seems to be low in the region (Bruns and Luque 2014). This paper provides new evidence for a vast Latin American country that teacher working conditions can be necessary for student performance.

In this paper, we adopt a model proposed by Woodcock (2015) for firms and workers and apply it to teachers and students. To the best of our knowledge, this is the first paper in the literature that uses a specification that controls for the match student-teacher in a developing country. This specification allows us to consistently estimate the effects of teacher human capital and working conditions on student outcomes. Other papers use a similar framework, but do not control for the match student-teacher fixed effects (Rockoff 2004; Hanushek and Rivkin 2010; Bruns and Luque 2014; Chetty, Friedman, and Rockoff 2014b). Failure to control for the student/teacher match based on unobserved characteristics may result in biased estimates if match fixed effects are correlated with teacher characteristics.

The rest of this paper is organized as follows: following this introduction, Section II presents the database and the descriptive statistics used. Section III presents the empirical strategy, while Section IV shows the results. Lastly, Section V presents the final considerations.

II. Data and descriptive statistics

The data used in this paper stems from the Longitudinal Study of School Generation project – *GERES* and covers the years 2005–2008. This project followed students in a sample of public and private schools during the first four years of elementary school testing their students with an exam designed to estimate proficiency levels in Mathematics and Portuguese language. The exams were low- stakes and not used to evaluate teacher or school performance. Schools signed a formal declaration stating their participation over the period, and after each wave, they received a report with the performance of students from different classes. Questionnaires were also administered to teachers, directors, parents, and students to assess family and school factors that influence the learning process. The project evaluated 303 public and private schools in five large Brazilian cities: Belo Horizonte, Rio de Janeiro, Campo Grande, Campinas, and Salvador.[4]

A complex probabilistic sample of students, classes, and schools was selected in each city, based on the 2003 School Census.[5] A series of cognitive tools, tests of Portuguese and Mathematics, contextual tools, and questionnaires for students, parents, teachers, and directors were then administered at the selected schools (Brooke and Bonamino 2011). The exams were prepared based on abilities matrices for reading and math, outlined by specialists, and used multiple-choice questions. A pre-test was conducted among students in public and private schools in Rio de Janeiro and Juiz de Fora. The analysis of pre-tested items was based on item response theory – IRT (see Lord 1980; Hambleton, Swaminathan, and Jane Rogers 1991; DeMars 2010). The most suitable items, those with high technical and pedagogical quality, were included in the final tests administered to students. To keep all schools in the project over the four years, the researchers provided a report after each wave that allowed to identify the performance bands of the students of the different classes. Professors and principals were actively engaged in the project, helping to maintain a suitable environment for the application of tests. Note, however, that the tests were applied by researchers linked to the *GERES* project. Thus, professors did not have access to the tests before the application.

[4]The municipalities selected were those that presented the most adequate conditions for developing the analysis, including the proximity to the universities that took part in the study (Brooke and Bonamino 2011).

[5]Schools that had only multi-grade classes for first grade were excluded from the sample. Private schools with fewer than 10 students enrolled in the second grade were also excluded, as were schools with fewer than 20 students. Furthermore, schools in rural zones and who had first grade students only at night classes were excluded. Thus, schools were selected based on their average socioeconomic level of the schools and their average size.

In each wave, students were tested in two main subjecs: the Portuguese language and Mathematics.[6] Two test versions – an easier test and a more difficult one – were constructed for each wave to minimize the possibility of measurement error estimated for students with different profiles. Different versions of the test (within and between waves) had similar items, enabling equalized scores from the IRT. For instance, two students answered similar questions even if they took different versions of the test in the same wave. Moreover, the same student may have answered the same question in two different waves.

In the first wave, the project took into account previous information about the schools to determine which version of the test students would take. From the second wave onwards, the version of the test was determined based on results from the previous year. From the fourth occasion onward, the tests began to include items from the Basic Education Evaluation System (SAEB) to compare the results of the GERES students and the results of the SAEB, which has national coverage. Besides, the tests lasted one hour each (Portuguese and Mathematics). After the first hour of testing, this was collected, and the students went for an interval. Then, also with sixty minutes, the second test began. The *GERES* scale of proficiency was built based on the assumptions of IRT and in consideration of the proximity of items to represent measures of difficulty (Brooke and Bonamino 2011). The levels are in ascending order and are cumulative; that is, students at the highest levels demonstrate mastery of not only skills at that level but also skills from lower levels. Additionally, students' proficiencies are recalculated each wave, by the same team, using the same methodology, which allows the comparison of the different waves and guarantees the reliability of the cognitive results obtained (Franco, Brooke, and Alves 2008).

The first wave, which was carried out as a diagnostic assessment, observed students in the 1st grade of primary education and took place at the beginning of the academic year, that is, in March 2005 in all cities.[7] The second wave was carried out at the end of the academic year, in November 2005, and the remaining waves were carried out in November 2006, 2007 and 2008.[8] Note that the tests were standardized across schools. This set up allows the project to follow the students over four years, that is, during the first cycle of primary education (Elementary Education I).[9] Not all registered students could be followed over time since some changed to schools that were not in the *GERES* sample. Students who failed a grade were kept in the sample if they remained at the same school or changed to another school included in the sample.[10] However, students who switched to a school that was not in the sample were excluded, and this could represent a selection bias.

The Elementary Education I curriculum has a joint national base, complemented in each school by a diversified part. The knowledge that forms part of the joint national basis to which all must have access, regardless of the region and the place in which they live, assures the unitary character of the national, states, the Federal District and the municipalities curricular guidelines, and the political-pedagogical projects of the schools. The curriculum of the joint national base of Elementary Education I must necessarily include, according to article 26 of the National Educational Bases and Guidelines Law (*Lei de Diretrizes e Bases da Educação* – LDB), the study of Portuguese language and Mathematics, as well as knowledge of the physical and natural world, and social and political reality, especially in Brazil. Also, it includes Arts, Physical Education (sports) and Religious Education.[11]

[6]Tests were administered to students present in class. Therefore, differences in the samples for Portuguese language and Mathematics may arise. Some of the loss of students over time may be due to school transfers.

[7]In Brazil, the K–12 system (basic education) is divided into two levels: primary education and high school. Primary education is compulsory and it is divided into Elementary Education I (1st to 5th grade) and Elementary Education II (6th to 9th grade). High school corresponds to grades 10–12. In this paper, our focus is on students from 1st to 5th grade, i.e. during the first cycle of primary education (Elementary Education I).

[8]As the students from Salvador did not participate in all waves, they were excluded from our sample.

[9]The regular age to attend Elementary Education I in Brazil is between 6 and 7 years old in the first year of elementary school and between 9 and 10 years old in the fourth year of elementary school. That means, the GERES project followed students that were 6–7 years old in 2005 (when the longitudinal panel started).

[10]The majority of students are attending at a regular age and 13% are repeating students and thus, older than the others.

[11]To test whether schools are roughly the same in our sample, we decomposed the R^2 and we found that schools explain only 1.66% of the variance for Portuguese scores, and only 2.61% for Mathematics.

Moreover, note that the LDB does not determine that schools should have different teachers for each module in elementary school. Therefore, schools hire only one teacher for Portuguese and Mathematics. Elementary school teachers are graduated in Pedagogy. That is because the content associated with Elementary Education I (Theoretical-Methodological Foundations in Portuguese and Theoretical-Methodological Foundations in Mathematics) is taught during undergrad in Pedagogy. Thus, the teachers are adequately trained to perform their functions.

This study followed students from an unbalanced panel of schools. Table 1 shows the number of students and schools in each wave. The total observations for proficiency tests amounted to 54,591 for Portuguese language and 54,533 for Mathematics over the five waves.

Table 2 shows the number of students who changed teacher between waves. Because we control for the teacher, student, and match fixed effects, our identification strategy requires that we observe teachers matched to the same student over time with varying working conditions and human capital. The Table shows that in every wave, a significant number of students remain with the same teacher. This happens because some students repeat the same grade (failure) and also because some of the teachers teach in consecutive grades.

Table 1. The GERES sample.

		Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
Portuguese	Students	9,447	10,070	12,607	10,263	12,204
	Schools	199	200	240	204	213
	Teachers	427	428	812	401	474
Math	Students	9,464	10,064	12,623	10,237	12,145
	Schools	199	200	240	204	213
	Teachers	427	428	813	401	474

Source: GERES (2005–2008).
Entries are the number of students, schools, and teachers in the sample of GERES Schools.

Table 3 describes and summarizes the teachers' working conditions and characteristics used in this paper. The Table shows that the majority of teachers have a college degree in pedagogy. Most of them do not have a post-graduate degree, and among those who do, most have a single specialization. More than half of the teachers have more than 15 years of experience teaching, and nearly half of teachers had another job. A small share of teachers had more than 15 years of tenure at the same school, and the majority has been working at the same school for approximately 5–10 years, between 21 hours and 25 hours per week. Teachers' family earnings are evenly spread across the income brackets.

III. Empirical strategy

Our econometric framework aims at identifying the impact of teacher characteristics controlling for unobserved heterogeneity of students, teachers and the student-teacher match. Failure to control for these fixed effects may bias the estimates of the effect of teacher human capital and working conditions if these are correlated with any such unobserved characteristics. Fixed effects capture the specific abilities of both students and teachers, while student-teacher match effects control for the fact that the productivity of above-average students may be higher when they are taught by above-average teachers.

Suppose that the student overall ability can be decomposed into general and specific abilities. The general component is controlled for by the student fixed effect, while the specific component represents the student skill that depends on the teacher ability. Therefore, by controlling for the student-teacher match, it is possible to capture not only the student general ability (which shapes her performance regardless of the teacher) but also the productivity gain that arises when student i match with a given teacher j. Although student general ability is constant

Table 2. Students who changed teachers.

		Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Total
	Did not change	9,447	9,843	3,494	2,746	3,228	28,758
Portuguese	Changed	0	227	9,113	7,517	8,976	25,833
	Total	9,447	10,070	12,607	10,263	12,204	54,591
Math	Did not change	9,464	9,835	3,494	2,720	3,188	28,701
	Changed	0	229	9,129	7,517	8,957	25,832
	Total	9,464	10,064	12,623	10,237	12,145	54,533

Source: *GERES* (2005–2008).
Entries are the number of students that changed teachers from one wave to the next. Waves 1 and 2 were carried out in the same year.

Table 3. Descriptive statistics.

Variables	Type	Categories	Portuguese Frequency	Mathematics Frequency
Other job	Categorical	Yes	45.19	45.26
		No	54.81	54.74
Number of worked hours at school	Categorical	Up–20 h/week	22.74	22.81
		21–25 h/week	30.61	30.64
		26–30 h/week	11.48	11.46
		31–40 h/week	20.84	20.76
		More than 40 h/week	14.33	14.33
Education	Categorical	High School	10.66	10.58
		Higher Education – Pedagogy	57.03	57.12
		Higher Education – Others	32.31	32.30
Qualification	Categorical	Did not made/completed postgraduate	52.20	52.18
		Training	8.57	8.54
		Specialization	37.33	37.38
		Master's degree/Doctorate	1.90	1.90
Experience	Categorical	Up–4 years	7.28	7.30
		5–10 years	14.89	14.91
		11–15 years	18.96	18.90
		More than 15 years	58.87	58.89
Tenure	Categorical	Up–1 year	19.66	19.70
		1–2 years	14.04	14.04
		3–4 years	16.74	16.76
		5–10 years	22.28	22.34
		11–15 years	15.06	14.96
		More than 15 years	12.33	12.21
Teacher's family earnings	Categorical	Up–R$1900	33.03	33.07
		R$1901–R$3100	35.17	35.11
		More than R$3100	31.80	31.83

Source: *GERES (2005–2008).*
The Portuguese and Math proficiencies frequencies are different due to a different number of students in each proficiency.

over time, her specific component varies across teachers.

Our specification is based on Woodcock (2015), which was originally designed to model the behavior of firms and workers. A pair is represented as a student (i) and teacher (j), so (i, j) determines the match between them. Every match is treated as a single link, even if the bond is separated over time. For a student who failed and repeated a grade with the same teacher, this tie in different years is considered the same match. However, if a student failed and repeated the grade with a different teacher, a new match is registered. The total number of matches in the sample is 34,979 in Portuguese and 34,991 in Mathematics. The model is as follows:

$$Y_{ijt} = \mathbf{x}'_{ijt}\boldsymbol{\beta} + \gamma_i + \theta_j + \delta_{ij} + \varepsilon_{ijt} \tag{1}$$

where Y_{ijt} is the test score of student i with teacher j at time t;[12] $\mathbf{x}'_{ijt}$ is the vector $(1 \times K)$ of observed control variables plus the grade dummies and wave dummies, which vary over time for both student i and teacher j; γ_i is the student fixed effect; θ_j is her teacher fixed effect; δ_{ij} is the student-teacher match fixed effect, and finally ε_{ijt} is the random error term.

In practice, we estimate the match fixed effects models, by taking deviations from the match mean, as in Equation 2:

$$Y_{ijt} - \bar{Y}_{ij} = \left(\mathbf{x}'_{ijt} - \overline{\mathbf{x}'}_{ij}\right)\boldsymbol{\beta} + \left(\varepsilon_{ijt} - \bar{\varepsilon}_{ij}\right) \tag{2}$$

Equation 2 uses only the variation in teacher characteristics fixing the teacher/student match over time. We do not include student variables in the regressions because they do not vary over time (the student questionnaire was administered only in wave 1).

For the sake of comparison, we also estimate the pooled ordinary least squares – POLS (Equation 3), student fixed effects (Equation 4) and student-plus-teacher fixed effect models (Equation 5):

$$Y_{ijt} = \mathbf{x}'_{ijt}\boldsymbol{\beta} + \varepsilon_{ijt} \tag{3}$$

$$Y_{ijt} = \mathbf{x}'_{ijt}\boldsymbol{\beta} + \gamma_i + \varepsilon_{ijt} \tag{4}$$

[12]We normalized the test scores across all students and waves. As previous explained, proficiencies are recalculated each wave, allowing us to compare different waves.

$$Y_{ijt} = \mathbf{x}'_{ijt}\boldsymbol{\beta} + \gamma_i + \theta_j + \varepsilon_{ijt} \quad (5)$$

To identify the model conditional on all fixed effects the variables need to vary within students over time.[13] Table 4 presents the within-student variation in student and teacher characteristics. The variation described in Column 1 (within the student and across teachers) allows us to estimate the POLS and student fixed effects specifications. To estimate the match student-teacher specification – Equation 2 – we use the within-match variation described in Column 2. Lastly, the Table also presents the percentage of teachers who changed their characteristics over time in Column 3. The within-student variation in Column 1 is larger than the within-match variation, which leads to more precise estimations in the POLS and student fixed effects model. Despite the small within-match variation in teacher characteristics, the estimator is identified, but it can be associated with large standard errors.

IV. Results

To compare the results of different models, regressions are estimated for each discipline using POLS, student fixed effects only (FEi), student and teacher fixed effects (FEi+FEj), and student-teacher and match student-teacher (FEij) using unbalanced panels.[14] Tables 5 and 6 present the results of these estimates for Portuguese and Mathematics, respectively. Student test scores are normalized so that the estimated coefficients can be interpreted directly in units of a standard deviation.

Table 4. Variation in teachers' characteristics.

Variables	(1)	(2)	(3)
Dependent variable			
Portuguese proficiency	0.31	0.07	-
Math proficiency	0.37	0.04	-
Teacher variables			
Education	0.42	0.02	2.21
Qualification	0.39	0.03	2.10
Experience	0.40	0.01	2.56
Tenure	0.35	0.01	5.83
Other job	0.50	0.03	6.88
Number of worked hours at school	0.33	0.02	5.01
Teacher's family earnings	0.32	0.02	8.28

Source: *GERES* (2005–2008).
Columns 1 and 2 present the within variation for students and match, respectively. Column 3 shows the percentage of teachers that change characteristics over time.

Table 5 shows the results for the Portuguese language. Students whose teachers have another job present higher test scores if we do not control for fixed effects (Column 1). When student fixed effects are controlled for the coefficient is no longer significant at the usual levels (Column 2). This suggests that student unobserved characteristics affect their test scores, and the positive effect found in the POLS estimation is overestimated. When the fixed effect of the teacher is also controlled for (Column 3), the coefficient becomes negative but remains not significant. This result may be related to teachers' intrinsic commitment. Teachers who have another job may be less focused, which may negatively affect their students' test scores. When student-teacher match fixed effects are controlled for (Column 4), the estimated coefficients are also negative and not significant. However, the coefficients are smaller than those in Column 3. This corroborates the hypothesis that holding another job may result in less focus in the classroom, which may hurt students' outcome.

Regarding teacher workload, the results of the POLS specification suggest that there is a negative association between teachers' working hours and students' test scores. The inclusion of students' fixed effects, however, changes the sign of the correlation. As more fixed effects are controlled for, the coefficients become larger. When teachers and students fixed effects are controlled for (Column 3), teachers work full time at school (31–40 hours) have an impact of 23.8% of a standard deviation of test scores relative to part-time teachers.

Note that in Brazil is common for people to attend night classes. Therefore, it is possible that a teacher will finish her high education degree or any type of qualification while teaching at the same time. Because of this possibility, we can still identify the impact of this teacher human capital after controlling for fixed effects. Hence, looking at the results of teacher education in Column 1, where no fixed effects are included,

[13]For more details, see Cameron and Trivedi (2010).
[14]We estimated all models without grade dummies and obtained similar results.

Table 5. Portuguese language regressions (continue).

PORTUGUESE	(1)	(2)	(3)	(4)
Variable	POLS	FEi	FEi+FEj	FEij
Other job				
No (omitted)				
Yes	0.0657***	0.00514	−0.0180	−0.0293
	(0.00867)	(0.00624)	(0.0234)	(0.0293)
Number of worked hours at school				
Less than 20 h/week (omitted)				
21–25 h/week	−0.0122	0.00321	0.188***	0.223***
	(0.00907)	(0.00831)	(0.0414)	(0.0479)
26–30 h/week	−0.233***	0.0111	0.0532	0.0747
	(0.0122)	(0.0109)	(0.0709)	(0.0995)
31–40 h/week	−0.121***	0.0263***	0.230***	0.308***
	(0.00980)	(0.00864)	(0.0464)	(0.0571)
More than 40 h/week	−0.129***	0.0354***	0.229***	0.238***
	(0.0110)	(0.00986)	(0.0449)	(0.0546)
Education				
High School (omitted)				
Higher Education – Pedagogy	0.0497***	−0.0205**	−0.0641	−0.247**
	(0.0115)	(0.00931)	(0.0717)	(0.113)
Higher Education – Others	0.128***	−0.00663	0.0527	−0.0389
	(0.0120)	(0.00961)	(0.0683)	(0.0970)
Qualification				
Do not made/complete postgraduate (omitted)				
Training	0.00351	−0.00145	0.117***	0.0705
	(0.0122)	(0.00936)	(0.0419)	(0.0591)
Specialization	0.0559***	−0.00634	−0.0704*	−0.117**
	(0.00722)	(0.00593)	(0.0398)	(0.0497)
Master's degree/Doctorate	0.514***	0.000161	0.109	0.0810
	(0.0214)	(0.0159)	(0.0917)	(0.133)
Experience				
Less than 4 years (omitted)				
5–10 years	−0.118***	0.0111	0.0202	0.0944
	(0.0154)	(0.0119)	(0.0668)	(0.0776)
11–15 years	−0.173***	−0.00670	0.0819	0.146
	(0.0149)	(0.0116)	(0.0826)	(0.104)
More than 15 years	−0.140***	0.00130	0.136	0.235**
	(0.0141)	(0.0111)	(0.0934)	(0.113)
Tenure				
Less than 1 year (omitted)				
1–2 years	0.0349***	−0.0123	−0.119***	−0.130***
	(0.0117)	(0.00896)	(0.0334)	(0.0433)
3–4 years	0.0929***	−0.0194**	−0.103**	−0.194***
	(0.0113)	(0.00888)	(0.0450)	(0.0625)
5–10 years	0.175***	−0.0262***	−0.210***	−0.375***
	(0.0105)	(0.00869)	(0.0619)	(0.0848)
11–15 years	0.168***	−0.0215**	−0.209**	−0.378***
	(0.0118)	(0.00982)	(0.0896)	(0.128)
More than 15 years	0.290***	−0.0143	−0.126	−0.302**
	(0.0133)	(0.0105)	(0.0891)	(0.121)
Teacher's family earnings				
Less than R$1900 (omitted)				
R$1901–R$3100	0.0435***	0.00544	−0.0507	−0.0784*
	(0.00799)	(0.00671)	(0.0310)	(0.0425)
More than R$3100	0.157***	0.0331***	0.00648	0.0110
	(0.00847)	(0.00754)	(0.0477)	(0.0685)
Repeating students	−0.224***	−0.0266***	−0.00889	−0.00488
	(0.00931)	(0.0101)	(0.0101)	(0.0140)
Grade dummies	Yes	Yes	Yes	Yes
Wave dummies	Yes	Yes	Yes	Yes
Constant	−1.124***	−1.053***	−1.051***	−0.957***
	(0.0212)	(0.0167)	(0.182)	(0.167)
Observations	54,539	54,539	54,539	54,539
Hausman (χ^2)		2026.13		536.53
R^2	0.430	0.753	0.784	0.595

Source: GERES database (2005–2008).
Robust standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1.

Table 6. Math regressions.

MATH				
	(1)	(2)	(3)	(4)
Variable	POLS	FEi	FEi+FEj	FEij
Other job				
No (omitted)				
Yes	0.0626***	−0.00466	−0.0732***	−0.0569*
	(0.00851)	(0.00716)	(0.0280)	(0.0337)
Number of worked hours at school				
Less than 20 h/week (omitted)				
21–25 h/week	0.00273	−0.00915	0.0299	0.0899
	(0.00858)	(0.00980)	(0.0492)	(0.0587)
26–30 h/week	−0.139***	0.0520***	−0.121	−0.0830
	(0.0118)	(0.0127)	(0.0814)	(0.114)
31–40 h/week	−0.113***	−0.0245**	0.227***	0.308***
	(0.00939)	(0.0101)	(0.0581)	(0.0724)
More than 40 h/week	−0.105***	0.0227**	0.246***	0.271***
	(0.0103)	(0.0114)	(0.0551)	(0.0668)
Education				
High School (omitted)				
Higher Education – Pedagogy	0.0732***	0.0184*	−0.0908	0.0560
	(0.0108)	(0.0110)	(0.0803)	(0.136)
Higher Education – Others	0.121***	0.0221**	−0.166**	0.0358
	(0.0114)	(0.0112)	(0.0807)	(0.124)
Qualification				
Do not made/complete postgraduate (omitted)				
Training	0.00734	0.0264**	0.0524	0.0101
	(0.0118)	(0.0106)	(0.0488)	(0.0701)
Specialization	0.0464***	0.0161**	0.0171	−0.0358
	(0.00686)	(0.00690)	(0.0449)	(0.0550)
Master's degree/Doctorate	0.493***	0.0288	0.625***	0.556***
	(0.0216)	(0.0197)	(0.112)	(0.144)
Experience				
Less than 4 years (omitted)				
5–10 years	−0.0945***	−0.0263*	−0.0908	−0.128
	(0.0144)	(0.0136)	(0.0904)	(0.100)
11–15 years	−0.146***	−0.0337**	−0.0961	−0.113
	(0.0139)	(0.0133)	(0.106)	(0.127)
More than 15 years	−0.138***	−0.0460***	−0.0786	−0.127
	(0.0133)	(0.0126)	(0.116)	(0.136)
Tenure				
Less than 1 year (omitted)				
1–2 years	0.0493***	0.0148	0.00135	−0.0506
	(0.0111)	(0.0105)	(0.0407)	(0.0526)
3–4 years	0.0984***	0.0232**	0.0140	−0.00869
	(0.0106)	(0.0101)	(0.0557)	(0.0724)
5–10 years	0.153***	−0.00499	−0.120	−0.142
	(0.0101)	(0.00969)	(0.0740)	(0.0988)
11–15 years	0.195***	0.0413***	−0.121	−0.0951
	(0.0117)	(0.0111)	(0.103)	(0.147)
More than 15 years	0.273***	0.0221*	0.0823	0.0755
	(0.0123)	(0.0120)	(0.103)	(0.148)
Teacher's family earnings				
Less than R$1900 (omitted)				
R$1901–R$3100	0.0464***	−0.0163**	−0.0532*	−0.0598
	(0.00757)	(0.00769)	(0.0321)	(0.0422)
More than R$3100	0.123***	−0.0268***	0.0217	−0.0437
	(0.00799)	(0.00877)	(0.0527)	(0.0742)
Repeating students	−0.220***	0.0308***	0.0105	0.0231**
	(0.00916)	(0.00925)	(0.00872)	(0.00953)
Grade dummies	Yes	Yes	Yes	Yes
Wave dummies	Yes	Yes	Yes	Yes
Constant	−1.036***	−0.987***	−0.606***	−0.915***
	(0.0187)	(0.0192)	(0.206)	(0.213)
Observations	54,481	54,481	54,481	54,481
Hausman (χ^2)		1827.38		737.6
R^2	0.480	0.773	0.824	0.637

Source: GERES database (2005–2008).
Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

students whose teachers have only a high school education have lower grades than students whose teachers have higher education. As one observes in Column 2, when student fixed effects are controlled for, students whose teachers have a college degree in pedagogy have lower grades than those whose teachers have only a high school education. Factors such as student effort may be overestimating the results so that when we control for these factors, the results change. Teachers with college degrees in majors other than pedagogy have no significant impact on their students relative to teachers with only a high school education. Notably, the results in Column 3 show that, when the students and teachers fixed effects are controlled for, there is no significant difference in test scores between students whose teachers hold a college degree and students whose teachers hold a college degree in pedagogy. However, when the student-teacher match fixed effect is controlled for (Column 4), having classes with a teacher qualified in pedagogy decreases student scores by 24.7% of the standard deviation. Curi (2005) shows that high school students who choose to study pedagogy in higher education tend to be the least prepared of all high school students and that they have the lowest scores on the national high school exam (ENEM). This may explain why students of teachers with a pedagogy degree fare worse than those of teachers with only a high school education.

Regarding academic degree, the results of Column 1 show that having specialization and having a master's/doctorate has a positive and statistically significant impact on students' test scores. However, when the fixed effects are included in the model, these results change. After controlling for the match fixed effects (Column 4), only the category of specialization is significant in the analysis. Nevertheless, it attracts a negative coefficient, which suggests that students whose teachers hold specialized degrees scored lower in Portuguese than students of teachers who did not complete postgraduate education.

Concerning experience, without controlling for ability (Column 1) the coefficients are statistically significant and negative. These initial results corroborate initial findings in the literature that show that teachers at the beginning of their career have more impact on students' learning (for example, Rockoff 2004; Rivkin, Hanushek, and Kain 2005; Croninger et al. 2007; Buddin and Zamarro 2009). After student fixed effects are controlled for, we do not find a significant impact, however (Column 2). The same is true after controlling for teacher fixed effects (Column 3). However, after student-teacher match fixed effects (Column 4) are controlled for, students whose teachers have more than 15 years of teaching experience score higher (23.5% of the standard deviation) than students whose teachers have less than 1 year of experience. This finding suggests that more experienced teachers may better exploit their students' skills, leading them to achieve higher test scores. More recent results in the literature suggest that teacher effectiveness increase over the years as described, for instance by Wiswall (2013).

Experience, education, and qualifications are frequently used to analyze teacher performance in the classroom because these characteristics are expected to improve teacher skills and thus improve student learning (Hanushek and Rivkin 2007). Regarding teacher tenure at the same job, one sees a divergence in the coefficients across different specifications. In the POLS specification (Column 1), the coefficient for long tenure is positive, although it becomes not significant or negative when fixed effects are controlled for (Columns 2 and 3). Column 4 shows that when the student-teacher match fixed effect is controlled for, longer teacher tenure negatively affects student test scores relative to shorter teacher tenure.

The effect of teacher family income and student proficiency also changes significantly across specifications. The results of Column 1 show that teachers with lower family income tend to be less effective than teachers with higher family income. However, the results of Column 4, for which effects of a student-teacher match are controlled, show a U-shaped relationship between teacher family income and student test scores, even after we control for effects of student, teacher and student-teacher match. It is essential to highlight that in our preferred specifications (columns 3 and 4), we controlled for teacher's fixed effects and thus, teachers' earnings are not correlated to their ability. Therefore, it is plausible to assume that teacher family income is a proxy for teacher pay.

Grade repetition is also a major issue in Brazil that generates considerable discussion among education specialists and politicians. The results of Column 1 depict a negative relationship between grade repeating and test scores, suggesting that students who have failed in the past continue to fare worse academically than their non-repeating counterparts. This difference persists even after we control for student fixed effects. When we control for effects of both student and teacher fixed effects simultaneously, the coefficient loses statistical significance. Similarly, when we control for all three fixed effects (student, teacher and student-teacher match), the coefficient remains negative but insignificant.

Grade and wave effects are also included in all regressions to control for the fact that learning accumulates across grades and over time, as the results in Table 4 confirm. The grade and wave dummies are identified separately because grade repetition is quite common in Brazil, consequently, many students repeated the same grade in different waves.

Table 6 presents the results for Mathematics, which differ from the results for the Portuguese language in important ways. Regarding working conditions, the impact of having another job is overestimated when no fixed effect is controlled for, as Column 1 shows a positive and significant impact. However, after controlling for student fixed effects, there is no significant impact of having another job on test scores. In Column 3, when student and teacher fixed effects are controlled for, the coefficient changes its sign and becomes significant again. This shows that characteristics such as teacher motivation may negatively impact students' test scores. In the match specification (Column 4), the results corroborate those presented in Column 3, the coefficient is negative and significant, but smaller in magnitude.

Teacher workload also affects students' test scores. The results for Mathematics are very similar to those found for the Portuguese language. In Column 1, when unobserved effects are not controlled for, the impact of teacher workload is negative. However, comparing to other specifications, its impact is underestimated. After student fixed effects are controlled for, the impact becomes higher than the previous specification (POLS), although there are still coefficients with a negative sign. Subsequently, when teacher fixed effects are considered in addition to the student fixed effects, the results indicate that more hours worked per week by teachers positively impacts student test scores. Controlling for effects of a student-teacher match, as shown in Column 4, the results are similar to those shown for specification (3), but the coefficients in Column 4 are greater than those in Column 3. Therefore, teachers working for 31–40 hours a week (full time) can lead to an improvement in students' test scores by approximately 30.8% of the standard deviation. Teachers who work full time may have more time to prepare for classes, and they may be able to be more committed to the school and student learning.

The other results show that without controlling for any fixed effects (Column 1), students whose teachers have a high school education had lower test scores than students whose teachers have a college degree, either in pedagogy or in other majors. When the student fixed effects are controlled for, the result is similar (Column 2), but the coefficient is smaller. When all three fixed effects are controlled for (Column 4), teacher education does not seem to matter for the students' Mathematics test scores. Those results are similar to the ones found by Buddin and Zamarro (2009) for Los Angeles.

Academic qualifications do appear to be important for Mathematics test scores. Having a master's degree/doctorate appears to be important for Mathematics learning. Students whose teachers hold a master's/doctorate degree score higher by approximately 55.6% of the standard deviation relative to the mean value, even after student-teacher match fixed effects are controlled for. Note that when we control for student fixed effects (Column 2), the influence of this factor is not significant, then student's ability is overestimated in the results. When we control for teachers and students fixed effects (Column 3), holding a master's degree/doctorate again has a significant and positive impact on student test scores, but the coefficient is larger than the specification in Column 1. Finally, when the student fixed factors such as skills of students and teachers

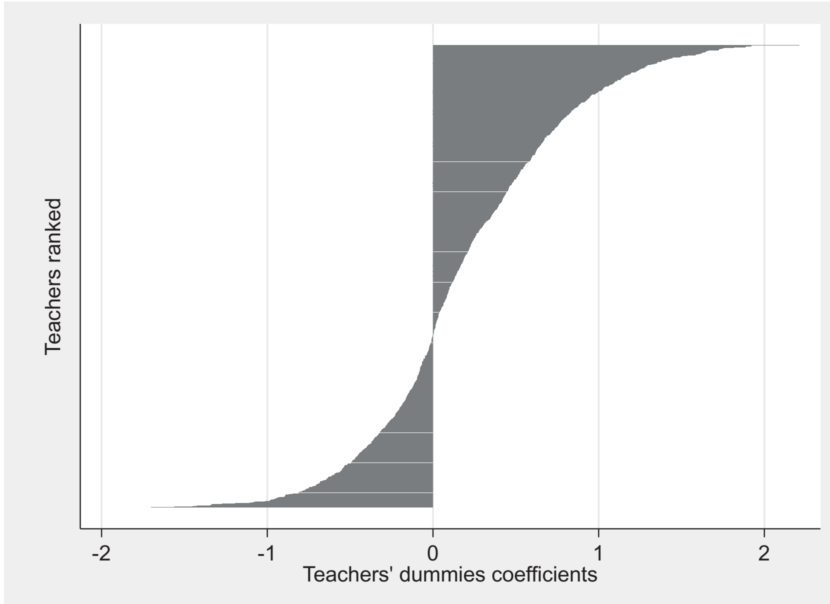

Figure 1. Teachers' ranking – Portuguese.

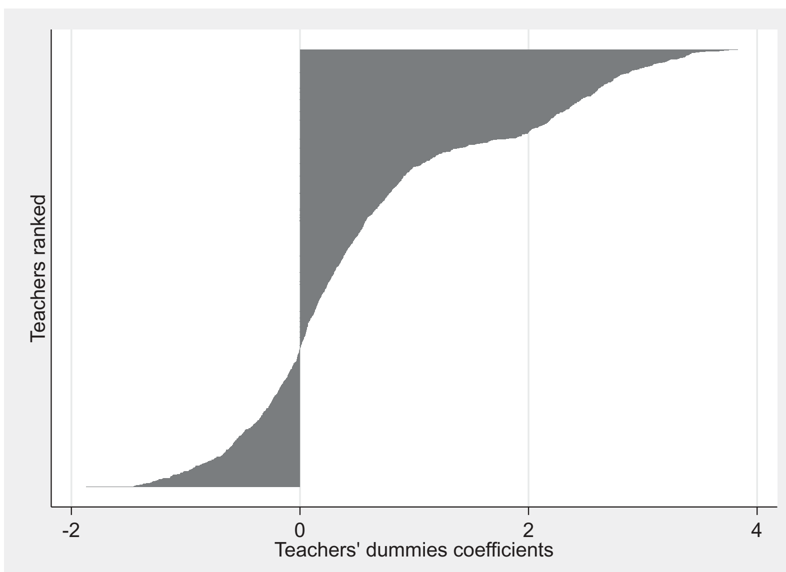

Figure 2. Teachers' ranking – Math.

and their interactions are controlled for, the impact of holding a master's/doctorate on student test scores is still positive and significant, albeit smaller than in specification (3).

Teaching experience does not seem to impact Mathematics proficiency after student-teacher match fixed effects are controlled for. We observe the same for tenure. Students who repeated a grade had higher test scores than those who did not, which we observe once we control for the student-teacher match fixed effect. This finding suggests that students who repeat are trying harder to earn better grades. Finally, as expected, test scores increase with grade progression and across waves.

To summarize the effect of teachers unobserved ability on student test scores Figures 1 and 2 are constructed with the coefficients of the teachers' dummies from the regression including student and teacher fixed effects only – $\underline{FEi}$ + $\underline{FEj}$ from Column 3 of Tables 5 and 6. The student-teacher match fixed effects specification is more complete and that is why we considered it as our main specifications so far. It does not allow us to recover teacher fixed effects, however. Thus, we go back one step to the specification that includes student and teacher fixed effects only to construct Figures 1 and 2 based on teacher fixed effects that are captured by the teachers' dummies. These two figures depict the difference in proficiency between students with below-average teachers (those negatively affecting students' proficiency) and students with above-average teachers (those positively affecting students' proficiency).

In terms of Portuguese scores, students whose teachers who are one standard deviation below the mean teacher ability score −35 points lower than average. In contrast, students whose teachers are one standard deviation above the mean teacher ability score 30 points higher than average. A fictitious student who moves from a teacher with the largest negative effect to a teacher with the largest positive effect would, therefore, increase in proficiency by 5.22 standard deviations, or approximately 170.35 points. This means that this student would initially only master basic reading skills and, after the change, would master more complex skills.[15] For example, a student who previously could only recognize letters of the alphabet, conduct inferences and recognize the mood effect of the use of ellipsis direction and onomatopoeia in a comic strip could also identify the purpose of unfamiliar text, establish relationships between texts of the same genre, recognize differences in information, establish relations between parts of a text through lexical substitution and cause-consequence in a short poem and a medium-length text, and identify patterns in a fable and verbally explicate them.

As for Mathematics, students whose teachers are one standard deviation below the mean teacher ability, score 73 points lower than average. In contrast, students whose teachers are one standard deviation above the average, score 65 points higher than average. A student moving from a teacher with the largest negative effect to a teacher with the largest positive effect would score the equivalent of 6.41 standard deviations higher than average or approximately 444.57 points. This student's knowledge and skills would therefore grow from mastering only basic Mathematics abilities to mastering more complex abilities, such as the ability to solve numerical problems involving different operations, including equalization with a change in quantity; comparison of proportional terms by subtracting rational numbers in decimal form and operations involving the rectangular configuration of multiplication. Skills such as problem-solving involving fundamental operations would also be broadened and consolidated. Finally, the student could solve problems incorporating operative actions with natural, rational numbers in the form of decimal or percentage; measures of length, mass, volume and their applications, such as calculating the perimeter and area; and measuring time and value.

V. Final comments

Teachers play a crucial role for students who are developing their abilities in language and Mathematics. Besides, most of the public policy discussions, directly and indirectly, involve teacher working conditions and student performance on standardized exams. However, there is no consensus on how to distinguish the impact of the specific features of a good teacher or her relative importance from other relevant factors.

This paper investigates the impact of observed and unobserved teacher human capital and working conditions on student learning, specifically for Portuguese language and Mathematics proficiency levels. Our empirical strategy relies on panel data of public and private schools for four large Brazilian cities. We take advantage of the panel to control for three types of unobserved heterogeneity: student, teacher, and the student-teacher match fixed effects in our analysis. Our results show that for the Portuguese language, some of the teacher working conditions have a significant impact on student proficiency. Teachers who have a second job have no impact on their students' performance relative to teachers who only have

[15]The interpretation of the GERES proficiency scale is based on a range of skills development, from basic knowledge and abilities to more complex ones in Portuguese, and in Mathematics as well. See Brooke and Bonamino (2011) to a detailed explanation about the interpretation of the proficiency levels in the lower grades of elementary school.

one job. For Mathematics, however, the effect is negative.

Working more hours positively impact student achievement in both disciplines. Teachers who work more hours may have more time to prepare for class, or they may be more committed to school and student learning. Additionally, for Portuguese teachers who hold a degree in pedagogy impact their students negatively and those with more than 15 years of experience impact their students positively. Conversely, more tenured teachers have a negative impact on their students' grades compared to those teachers that have less than one year tenure. For Mathematics, teachers who hold a master's/doctorate positively impact their students' scores rather than those teachers who did not complete a graduate program. Contrary, education, experience and tenure did not have a significant impact on Mathematics grades after controlling for all unobserved students, teachers, and the student-teacher match characteristics.

Moreover, we simulated an exercise in which we changed a student from the least skilled to the most skilled teacher. The results show that in the Portuguese language, this student grade could improve by 5.22 standard deviations and for Mathematics, a student making the same movement would improve by 6.41 standard deviations.

Overall, the results show that teacher human capital and working condition do matter for student achievement, and that is it is essential to control for heterogeneity at the three levels as we demonstrated in this paper.

Acknowledgments

The authors acknowledge the financial support of FAPEMIG, CNPq, and CAPES-INEP (Observatório da Educação 2010). We are also grateful to CAED/UFJF for providing us with the GERES database, and data technical support of ECONS - Economics Research Laboratory from UFJF. Any errors are our own.

Data availability statement

The data that support the findings of this study are available from the CAED/UFJF. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from the members of the project at https://laedpucrio.wordpress.com/projetos/o-projeto-geres/ with the permission of them.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES-INEP Observatório da Educação, 2010).;;;

ORCID

Ricardo Da Silva Freguglia http://orcid.org/0000-0002-1355-2272

References

Aaronson, D., L. Barrow, and W. Sander. 2007. "Teachers and Student Achievement in the Chicago Public High Schools." *Journal of Labor Economics* 25 (1): 95–135.

Brooke, N., and A. Bonamino. 2011. *GERES 2005: Razões E Resultados De Uma Pesquisa Longitudinal Sobre Eficácia Escolar*. Rio de Janeiro: Walprint gráfica e editora.

Bruns, B., and J. Luque. 2014. *Great Teachers: How to Raise Student Learning in Latin America and the Caribbean*. Washington, DC: The World Bank.

Buddin, R., and G. Zamarro. 2009. "Teacher Qualifications and Student Achievement in Urban Elementary Schools." *Journal of Urban Economics* 66 (2): 103–115.

Cameron, A. C., and P. K. Trivedi. 2010. *Microeconometrics Using Stata*. Vol. 2. College Station, TX: Stata press.

Cameron, S. V., and J. J. Heckman. 1998. "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males." *Journal of Political Economy* 106 (2): 262–333.

Chetty, R., J. N. Friedman, and J. E. Rockoff. 2014a. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104 (9): 2593–2632.

Chetty, R., J. N. Friedman, and J. E. Rockoff. 2014b. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review* 104 (9): 2633–2679.

Clotfelter, C. T., H. F. Ladd, and J. L. Vigdor. 2006. "Teacher-Student Matching and the Assessment of Teacher Effectiveness." *Journal of Human Resources* 41 (4): 778–820.

Clotfelter, C. T., H. F. Ladd, and J. L. Vigdor. 2007. “Teacher Credentials and Student Achievement: Longitudinal Analysis with Student Fixed Effects.” *Economics of Education Review* 26 (6): 673–682.

Coleman, J. S., E. Q. Campbell, C. J. Hobson, J. McPartland, A. M. Mood, F. D. Weinfeld, and R. L. York. 1966. *Equality of Educational Opportunity*, 1066–5684. Washington, DC: National Center for Educational Statistics.

Croninger, R. G., J. K. Rice, A. Rathbun, and M. Nishio. 2007. “Teacher Qualifications and Early Learning: Effects of Certification, Degree, and Experience on First-Grade Student Achievement.” *Economics of Education Review* 26 (3): 312–324.

Cunha, F., and J. Heckman. 2007. “The Technology of Skill Formation.” *American Economic Review* 97 (2): 31–47.

Cunha, F., and J. Heckman. 2009. “The Economics and Psychology of Inequality and Human Development.” *Journal of the European Economic Association* 7 (2–3): 320–364.

Curi, E. 2005. “A Formação Matemática De Professores Dos Anos Iniciais Do Ensino Fundamental Face Às Novas Demandas Brasileiras.” *Revista Iberoamericana De Educación* 37 (5): 1–10.

DeMars, C. 2010. *Item Response Theory*. Oxford, UK: Oxford University Press.

Franco, C., N. Brooke, and F. Alves. 2008. “Longitudinal Study of Quality and Equity in the Brazilian Elementary Education: GERES 2005.” *Ensaio: Avaliação E Políticas Públicas Em Educação* 16 (61): 625–637.

Glewwe, P., and M. Kremer. 2006. “Schools, Teachers, and Education Outcomes in Developing Countries.” *Handbook of the Economics of Education* 2: 945–1017.

Goldhaber, D., J. Cowan, and J. Walch. 2013. “Is a Good Elementary Teacher Always Good? Assessing Teacher Performance Estimates across Subjects.” *Economics of Education Review* 36: 216–228.

Guarino, C. M., M. D. Reckase, and J. M. Wooldridge. 2015. “Can Value-Added Measures of Teacher Performance Be Trusted?” *Education Finance and Policy* 10 (1): 117–156.

Hambleton, R. K., H. Swaminathan, and H. Jane Rogers. 1991. *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage.

Hanushek, E. A. 2003. “The Failure of Input-Based Schooling Policies.” *The Economic Journal* 113 (485): F64–F98.

Hanushek, E. A., and S. G. Rivkin. 2006a. *School Quality and the Black-White Achievement Gap*. No. w12651. Cambridge, MA: National Bureau of Economic Research.

Hanushek, E. A., and S. G. Rivkin. 2006b. “Teacher Quality.” *Handbook of the Economics of Education* 2: 1051–1078.

Hanushek, E. A., and S. G. Rivkin. 2007. “Pay, Working Conditions, and Teacher Quality.” *The Future of Children* 69–86.

Hanushek, E. A., and S. G. Rivkin. 2010. “Generalizations about Using Value-Added Measures of Teacher Quality.” *American Economic Review* 100 (2): 267–271.

Kane, T. J., and D. O. Staiger. 2008. *Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation*. No. w14607. Cambridge, MA: National Bureau of Economic Research.

Keane, M. P., and K. I. Wolpin. 1997. “The Career Decisions of Young Men.” *Journal of Political Economy* 105 (3): 473–522.

Kirabo, J. C., J. E. Rockoff, and D. O. Staiger. 2014. “Teacher Effects and Teacher-Related Policies.” *Annual Review of Economics* 6 (1): 801–825.

Kraft, M. A. 2019. “Teacher Effects on Complex Cognitive Skills and Social-Emotional Competencies.” *Journal of Human Resources* 54 (1): 1–36.

Leibowitz, A. 1974. “Home Investments in Children.” *Journal of Political Economy* 82 (2, Part 2): S111–S131.

Lord, F. M. 1980. *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, NJ: Lawrence Erbaum Associates.

Mathematics, Reading and Science, PISA, OECD Publishing, PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition, February 2014): Student Performance in Paris, doi:10.1787/9789264208780-en.

Murnane, R. J., J. B. Willett, and F. Levy. 1995. *The Growing Importance of Cognitive Skills in Wage Determination*. No. w5076. Cambridge, MA: National Bureau of Economic Research.

Neal, D. A., and W. R. Johnson. 1996. “The Role of Premarket Factors in Black-White Wage Differences.” *Journal of Political Economy* 104 (5): 869–895.

Rivkin, S. G., E. A. Hanushek, and J. F. Kain. 2005. “Teachers, Schools, and Academic Achievement.” *Econometrica* 73 (2): 417–458.

Rockoff, J. E. 2004. “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data.” *American Economic Review* 94 (2): 247–252.

Todd, P. E., and K. I. Wolpin. 2003. “On the Specification and Estimation of the Production Function for Cognitive Achievement.” *The Economic Journal* 113 (485): F3–F33.

Wiswall, M. 2013. “The Dynamics of Teacher Quality.” *Journal of Public Economics* 100: 61–78.

Woodcock, S. D. 2015. “Match Effects.” *Research in Economics* 69 (1): 100–121.