The Effect of Teacher Content Knowledge on Student Achievement: A Quantitative Case Analysis of Six Brazilian States

Raquel Guimarães[1]
Asha Sitaram[2]
Lucia Jardon[3]
Shimpei Taguchi[4]
Lenora Robinson[5]

Abstract

This study investigates the effect of teacher content knowledge, as measured by test performance, on the math achievement of fourth grade students in Brazil. Using longitudinal data from six Brazilian states that participated in the FUNDESCOLA program in 1999, we conduct the analysis using a value-added model that controls for student and teacher characteristics, class composition, and school fixed effects. Our evidence suggests that teachers with higher content knowledge have a greater impact on the math test scores of their students, an effect that is even larger at the school-level. Our results indicate the need for further research into this subject. They also suggest that the allocation of resources toward the improvement of teacher content knowledge, through approaches such as targeted professional development and more rigorous certification requirements, may be an important aspect of increasing student achievement.

Keywords: Educational Economics; Input Output Analysis; Human Capital

1. Introduction

A great deal of research shows that teachers can make a substantial difference in student learning (Rockoff, 2004; Boyd, Lankford, Loeb, Rockoff, & Wyckoff, 2008; Clotfelter, Ladd, &

---

[1] Doctoral Candidate at Cedeplar/UFMG/Brazil – raquel@cedeplar.ufmg.br
[2] Stanford University – asha.sitaram@alumni.stanford.edu
[3] University of Zurich – lucyjardon@gmail.com
[4] Stanford University – shimpeit@alumni.stanford.edu
[5] Harlem Village Academies – Lrobinson@harlemvillage.org

Vigdor, 2007a; Clotfelter, Ladd, & Vigdor, 2007b; Ladd, 2008; Nye, Konstantopoulos, & Hedges, 2004). Recent studies aimed at examining the relationship between diverse teacher characteristics and student achievement benefit from longitudinal designs and value-added models, by systematically controlling for student heterogeneity and for selection bias in the matching of students to teachers and schools (Ladd, 2008; Boyd et al., 2008; Clotfelter, Ladd, & Vigdor, 2007b). However, there is no consensus on the importance of individual teacher attributes for student learning gains. For instance, some authors find that experience, test scores, and regular licensure, are positively associated with student achievement (Clotfelter, Ladd, & Vigdor, 2007b), whereas other studies show no predictive value of teacher credentials on the variance of teacher effects (Aaronson, Barrow, & Sander, 2007; Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009; Nye et al., 2004; Rivkin et al., 2005). The effect of certification was found to be positive by some researchers (Carr, 2006; Cavaluzzo, 2004; Clotfelter, Ladd, & Vigdor, 2007a). Other studies suggest that the academic major of teachers may affect student achievement, given that specialization would require teachers to take a large proportion of courses within a particular subject area (Darling-Hammond, 1999; Harris and Sass, 2007; Monk, 1994).

The relevance of teacher test scores, as a proxy for teacher content knowledge, has not been deeply explored by the literature, perhaps as a result of the lack of adequate teacher assessments. However, certain studies and meta-analyses have found that the effect of teacher test scores on student achievement outcomes is positive and significant in some cases (Wayne and Youngs, 2003), but relatively small (Clotfelter, Ladd, and Vigdor, 2006). Another important measure of teacher effectiveness, teachers' pedagogical content knowledge, has been found to positively impact student achievement (Hill, Rowan, and Ball, 2005), though it is

difficult to dissociate pure content knowledge from the pedagogical aspects of this teacher characteristic.

While existing studies offer insight into the importance of some teacher characteristics, the scarcity of studies that specifically examine the effect of teacher content knowledge on student achievement highlights the need for and our interest in exploring this relationship within the Brazilian context. Hence, our study examines the effect of teacher content knowledge on student achievement in mathematics by employing teacher test score data. We use longitudinal data from the Plano de Desenvolvimento da Escola (PDE) program, or School Development Plan, from 1999 to 2003. This World Bank program provided funding for school improvement in six high-poverty Brazilian states: Rondônia, Pará, Pernambuco, Sergipe, Mato Grosso do Sul, and Goiás.

In utilizing a value-added model that controls for student and teacher characteristics, class composition, and school fixed effects, we measure the association between teacher content knowledge and student achievement in mathematics for fourth graders in Brazil. We employ teacher test scores as a proxy because it offers the most current assessment of teacher content knowledge, when compared to other variables (e.g., academic major).

The findings of this study could have important implications for policy and practice, as administrators and policy makers seek the most effective ways to invest in teachers and schools. Furthermore, this study is one of the first papers to examine the relationship between teacher content knowledge and student achievement in a developing country context.

2. Data and methods

The treatment in our analysis is teacher content knowledge, as measured by teachers' standardized test scores. These tests evaluated teachers' knowledge of the mathematical concepts that should be taught to students. The outcome variable that we observe is fourth grade student mathematics achievement, as measured by student standardized test scores in mathematics at the end of the academic year. We treated this variable by controlling for the following observable variables:

- Student initial standardized test score;

- Student gender;

- Student socio-economic status (SES);

- Student's mother's education level;

- Teacher's college degree status (dummy);

- Teacher's experience (in years);

- Teacher's certification (dummy);

- Class composition in terms of the students' scores;

- School fixed effects (as a proxy for resources and materials available at each school)

Each control variable impacts both the treatment and the outcome. For example, student SES likely affects the treatment because, theoretically, higher SES students' parents will demand better schools and teachers for their children. In addition, student SES affects student achievement outcomes because higher SES students also, theoretically, have greater access to

resources outside of the classroom, which will impact their test performance (Rothstein, 2004). Prior to our analysis, we expected that most of these observable variables would have a positive correlation with both the treatment and the outcomes.

We hypothesized that the treatment variable—teacher content knowledge—is positively associated with student learning. There are various theoretical assumptions that support this hypothesis, namely that teachers who dominate their subject matter are able to create better opportunities to enhance students' learning and interest and are more capable of addressing diverse learning needs in their classrooms (Ball, 2000). Furthermore, in the case of mathematics, teaching core tasks generally involves significant mathematical reasoning in the context of practice, and, hence, teachers with strong content knowledge tend to positively influence student outcomes (Ball, 2000).

2.1. Data

The PDE longitudinal study involved six rounds of data collection, following students from the fourth grade in 1999 to eighth grade in 2003. We utilized the sample from 1999, which consists of assessment data at the beginning of the school year (April) and at the end of the school year (November). The sample includes public schools from six high-poverty Brazilian states that participated in the PDE program, namely Rondônia, Pará, Pernambuco, Sergipe, Mato Grosso do Sul and Goiás. Schools in the sample were required to meet the following requirements: (i) all of their elementary classes were held during the day; and (iii) they had at least 200 students.

In this paper, our focus is on mathematics scores for both teachers and students from participating states. As part of the PDE program, fourth grade students and teachers were both

tested on Portuguese and mathematics in April and November of 1999. The tests were composed mainly of items from the National Evaluation System of Basic Education (SAEB). Hence, teachers were evaluated according to national standards that define the mathematical content to be taught to students. In Brazil, national assessment data from 2007 found that roughly 40 percent of fourth grade students did not achieve the basic skill level in mathematics (Soares, Fonseca, Alvares & Guimarães, 2012), emphasizing the need to identify ways to support teachers in increasing student achievement in this subject area.

The PDE survey also includes information on the socioeconomic characteristics of students, teachers and schools. The total number of students surveyed by the PDE in 1999 was 6,619. These students were distributed among 126 schools. The average number of students per school is 52.53 and the average number of fourth grade teachers per school is 2.21.

2.2. Method: Value-added model with school fixed effects

According to Ladd (2008), mainly three methodological issues arise when estimating teacher effectiveness. First, it is difficult to isolate the effect of teacher quality from other inputs, particularly those deriving from contextual factors. Second, it is not always possible to measure all of the intervening determinants of student achievement—such as parental motivation and SES—or even control for them without measurement error. Lastly, and most importantly, causal inference in this framework is compromised since students are, in general, not randomly assigned to teachers. Therefore, teacher effects may be confounded by unobservable attributes of students, such as their ability and motivation. If the non-random matching of students to teachers is neglected, the estimate of teacher effects will be biased upwards if more motivated or academically capable students are allocated to more effective teachers, or downwards, if the opposite is true.

In order to solve this puzzle, value-added models (VAM) are employed to assess the effect of individual teachers on raising student achievement (Chetty, Friedman, & Rockoff, 2012; Ladd, 2008; McCaffrey, Lockwood, Koretz, Louis, & Hamilton, 2004). Intuitively, the value-added approach attempts to isolate teacher effects from other factors such as prior student performance and socioeconomic status, by focusing on learning gains during one school year (Ladd, 2008; McCaffrey et al., 2004).

Although the VAM has become a promising approach to estimating teacher effects, it does not mitigate bias due to nonrandom allocation of students to teachers. To solve this problem, researchers combine the value-added estimation with a quasi-experimental research design that accounts for time-invariant characteristics of students, both observable and unobservable—the fixed effects design (Clotfelter, Ladd, & Vigdor, 2007b). For instance, the use of school fixed effects mitigates the bias due to nonrandom distribution of teachers among schools.

This study advances the Brazilian literature on teacher effectiveness. It employs a value-added model with school fixed effects as an identification strategy to derive the effect of teacher content knowledge on student achievement for public school students evaluated by the PDE program in November of 1999. As there is no policy of within-school tracking in Brazil, we assume that the assignment to treatment (teachers with high content knowledge) is ignorable within schools, but not ignorable among schools.

Our model is described as follows. The dataset contains information for students who were present in two evaluations: at the beginning of the school year (April 1999) and at the end of the school year (November 1999). The outcome of interest is the standardized mathematics test score of students at the end of the school year.

7

This study uses the following value-added model with linear specification:

$$A_{ijt} = \beta_0 + \beta_1 A_{ijt-1} + \beta_2 \overline{CK}_{ijt} + BX_{ij} + \theta_j + \epsilon_{ij} \ (1)$$

Where:

- A_{ijt}: standardized achievement of student i in school j at the end of the year

- A_{it-1}: standardized achievement of student i in school j at the beginning of the year

- $\overline{CK}_t$: standardized average content knowledge of the teacher of student i in school j

- X: a matrix of time-invariant covariates of student i in school j

- θ_j: school j fixed effects

- ϵ_{ij}: individual disturbance error

Therefore, β_2 is the estimate of the effect of teacher content knowledge on student achievement. This corresponds to the average difference in score between a student who had a teacher with a high-level of content knowledge and a student who had a teacher with a low-level of content knowledge in the same school, everything else held constant. Assuming that the assignment of students to teachers was random within schools, $\hat{\beta}_2$ will be an unbiased estimate of the average within-school effect of teacher content knowledge on student achievement in mathematics for PDE students in 1999. The violation of the assumption that teachers are randomly assigned to classrooms within schools will result in biased estimates, as described previously.

In order to test for model specifications, we estimate five models:

8

- Model 1: baseline score and teacher content knowledge in November/1999

- Model 2: Model 1 + student's characteristics (sex, SES, mother's education)

- Model 3: Model 2 + teacher's characteristics (certification, experience, college degree)

- Model 4: Model 3 + classroom composition (average final grade in mathematics by classroom)

- Model 5: Model 4 + school fixed effects.

In the PDE data, there are three teacher test score categories: April test scores, November test scores, and the average of these two scores. Therefore, there are three potential treatment conditions: the content knowledge of the teacher at the beginning of the school year, the content knowledge at the end of the school year, and the average of the beginning and end-year test results. From the 1999 sample, 210 teachers took April's test, 232 took November's test, and 178 teachers took both tests. We chose to run the model using the treatment variable with the most data points, namely the November test scores. As we assumed that teachers' subject matter knowledge does not significantly change within a given year, choosing to use test scores at only one point in time is appropriate. However, we ran three sets of models using: (i) teacher's initial test score; (ii) teacher's final test score; (iii) the average score of the teacher. Using specifications (i) and (iii), we did not find a significant effect of teacher content knowledge on student achievement. However, we chose to examine specification (iii), which uses the November teacher test score as the treatment variable, because it had the most observations; as such, we believed that it was the most accurate measure of teacher content knowledge.

Model 1 is the baseline, which only accounts for the score in April/1999 and the final measure of teacher content knowledge, controlling for students' initial mathematics test scores. By estimating more than one specification, we wanted to infer whether the effect of teacher content knowledge on student achievement would be robust to the inclusion of other variables in the model. In Model 2, we include student characteristics, such as gender, socioeconomic status (SES), and parents' education, as measured by mother's literacy; in Model 3, we include teacher characteristics that may be confounded with the teacher content knowledge measure, such as certification, teaching experience, and teachers' education level. In Model 4, we account for the effect of classroom composition by controlling for the average final grade in math for students at the classroom-level. Model 5 is our preferred specification because it accounts for all the potential confounders described above as well as for the nonrandom assignment of teachers to schools.

3. Findings

Table 1 shows the results of the OLS regression analyses on students' final standardized mathematics test scores, using the five specifications described in Section 2. Considering the effect of confounding variables, such as student characteristics, teacher attributes, and classroom composition, we can assume that Model 4 and 5 are more accurate in predicting student test scores than the other three models. The main difference between these two specifications is that Model 5 employs school fixed effects, and, hence, this model provides the most unbiased estimator of the effect of teacher content knowledge on student achievement. This is true because we assume that the assignment of students to teachers based on their content knowledge is ignorable within schools, but not ignorable overall.

Table 1: OLS regression analyses on students' score from the final standardized mathematics test

	Model 1	Model 2	Model 3	Model 4	Model 5
Treatment Variable					
Teachers' content knowledge	0,015	0.017+	0.019+	0.019+	0.039*
	(0.009)	(0.010)	(0.011)	(0.011)	(0.015)
Control Variables					
Standardized initial math test score	0.690**	0.690**	0.686**	0.685**	0.673**
	(0.010)	(0.010)	(0.011)	(0.011)	(0.011)
Female student (=1)		-0.086**	-0.088**	-0.083**	-0.089**
		(0.020)	(0.021)	(0.022)	(0.021)
Standardized student SES		0.019+	0.018+	0.019+	-0.019+
		(0.010)	(0.011)	(0.011)	(0.012)
Student's mother is literate (=1)		0,062	0,067	0.071+	-0,001
		(0.038)	(0.041)	(0.042)	(0.040)
Certified teacher (=1)			0.045*	0.049*	-0.060+
			(0.023)	(0.023)	(0.035)
Teaching experience more than 5 years (=1)			-0.077*	-0.089**	0,004
			(0.032)	(0.033)	(0.048)
Teacher graduated college (=1)			-0,005	-0,009	-0.116**
			(0.022)	(0.023)	(0.037)
Classroom average score				-0,016	0,01
				(0.013)	(0.021)
Constant	0,003	-0,006	0,029	0,145	0,071
	(0.009)	(0.038)	(0.051)	(0.103)	(0.155)
Number of Cases	5,613	5,151	4,647	4,416	4,416

R-Squared	0,48	0,48	0,48	0,47	0,45
F-test	2597,4	962,5	529,9	440	397,1

Source: PDE Data - April and November, 1999

Obs.: ** 99% significant; * 95% significant; + 90% significant

Our treatment variable, teacher test scores, which measures subject matter knowledge in mathematics, shows a significant association with student test scores in Model 4 at the 90 percent significance level and in Model 5 at the 95 percent significance level. In Model 4, a one standard deviation increase in teacher content knowledge, as measured by test scores, is associated with a 0.02SD increase in student test scores, holding other control variables constant. However, this effect size increases to 0.04SD once we employ school fixed effects in Model 5. As we argued, Model 5 provides the most unbiased estimate of teacher content knowledge. The literature suggests that effect sizes of the magnitude of 0.1SD are large. While our estimated effect size is not as large, we consider the impact of teacher content knowledge on student achievement to be important and relevant for educational policy. In Section 4, the findings will be discussed in the context of similar research.

In addition to the significant and relevant impact of teacher content knowledge, which is of primary importance to this study, our analyses yield interesting results with respect to one of the control variables. We find that student's socioeconomic status (SES) has a significant association with student test scores at the 90 percent significance level in all specifications that include this variable. However, the direction of the association is different in the models that include and exclude school fixed effects. In Model 4, student's SES is positively associated with her/his test scores, whereas in Model 5, which includes school fixed effects, this association is negative. This suggests that, overall, student's SES has a positive association with test scores.

12

However, when considering its impact within schools, this parameter has a negative association with student achievement.

4. Discussion

The importance of teacher quality for student achievement has been accepted in the general literature (Ladd, 2008). However, the definition of teacher quality and the factors related to it are still an area of discussion and contention. Various teacher characteristics have been studied in order to identify their relative importance for student achievement, yet the results have been inconclusive.

In our study, we use teacher content knowledge as a proxy for teacher quality. In line with existing literature, the study demonstrates the importance of content knowledge for the achievement of students in the Brazilian provinces we studied. Teachers with higher content knowledge, as measured by their performance on end-of-year standardized tests, have a positive impact on the standardized mathematics test scores of students. The effect size found in Model 4, 0.02 SD, is consistent with the effect size of 0.023SD found in the study by Clotfelter, Ladd, and Vigdor (2006). The effect of teacher content knowledge is even larger at the school level. In the case of school fixed effects, our study finds an effect size more than twice as high as that in Clotfelter, Ladd, and Vigdor (2006), 0.040SD as compared to 0.017SD (school fixed effects only) and 0.012SD (school fixed effects and lagged achievement control). This result suggests that if greater resources are allocated towards increasing the content knowledge possessed by teachers, student achievement—as measured by standardized test scores—will be positively affected. Hence, in the Brazilian content, if high quality teachers—as defined by their high content knowledge—are placed in with lower performing schools and regions, they may positively influence the academic achievement of children in those schools/regions.

One limitation of this study is the possibility that the content knowledge variable captures other dimensions not accounted for by the control variables. One example of this is teacher motivation, as it is likely related to teacher content knowledge and other teacher characteristics, but for which it is very difficult to control. Perhaps, the most important dimension not captured by this study is pedagogical knowledge, which we could not include due to the limitations of our dataset. If pedagogical knowledge is effectively applied, it may have an effect in the same direction as content knowledge and estimates of the effect of content knowledge would be biased upward. However, if the pedagogical approach of a teacher has an effect in the contrary direction, then estimates of the effect of content knowledge would be biased downward, and we would be underestimating the effect of content knowledge.

As discussed earlier, another limitation of this study is our inability to determine whether there exists a non-random process of allocating teachers to students based on teacher content knowledge. If non-random allocation does occur, our estimates could be biased in either direction, depending on the way teachers are distributed.

Despite these limitations, this study suggests that the effect of teacher content knowledge on student achievement in Brazil is important and should be the subject of further study. The mechanisms by which teacher content knowledge can be increased and by which individuals with higher content knowledge can be attracted into the teaching profession are unclear. Exploring potential mechanisms is a critical area for future research.

Using a value-added model, we also confirm that student characteristics like socioeconomic status are relevant for the academic achievement of students; however an important observation is that when we used fixed effects at the school level, student SES had a negative association with achievement. This finding might be a result of the strong relationship

between student socioeconomic background and the school context. That is, higher SES children go to the best schools while lower SES children are concentrated within the worst ones. However, the reasons for these findings are unclear and are an area for further exploration in the Brazilian context.

5. Conclusion

This study, using the PDE data within a value-added model, has shown that for low income states in Brazil, teacher characteristics are important for student achievement. In particular, our study finds that teacher quality, as measured by the end-of-year teacher test scores, has a significant and positive effect on students' mathematics scores in the Brazilian context.

Despite our inability to separate teacher content knowledge from other teacher characteristics, such as motivation and pedagogical knowledge, our findings do highlight a need for further research on the topic. It is also critical that we explore the relationship between content knowledge and student achievement across a range of subject areas. The findings in this study combined with further research on the subject may also warrant an exploration into the best mechanisms by which to increase teacher content knowledge in mathematics and possibly other subjects as well. One suggestion could be to increase teacher content knowledge requirements in the certification process or to increase teacher salaries and resources in an attempt to attract individuals with higher academic achievement and content knowledge into the teaching profession. Once a targeted method for increasing teacher content knowledge is found, policy makers may wish to consider directing the allocation of resources toward these strategies in order to boost student achievement in mathematics and, perhaps, also in other content areas.

References

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

Ball, D.L. (2000). Bridging practices intertwining content and pedagogy in teaching and learning to teach. *Journal of Teacher Education*, 51 (3), 241-247. doi:10.1177/0022487100051003013.

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

Boyd, D., Lankford, H., Loeb, S., Rockoff, J., & Wyckoff, J. (2008). The narrowing gap in New York City teacher qualifications and its implications for student achievement in high poverty schools. *Journal of Policy Analysis and Management*, 27 (4), 793-818. doi:10.1002/pam.20377.

Carr, M. (2006). *The determinants of student achievement in Ohio's public schools*. Columbus, OH: The Buckeye Institute for Public Policy Solutions.

Cavalluzzo, L. (2004). *Is national board certification an effective signal of teacher quality?* CNA Corporation Working Paper.

Chetty, R., Friedman, J.N., & Rockoff, J.E. (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., Ladd, H., & Vigdor, J. (2007a). *How and why do teacher credentials matter for student achievement?* NBER Working Papers. National Bureau of Economic Research.

———. (2007b). *Teacher credentials and student achievement in high school: a cross-subject analysis with student fixed effects*. Working Paper. National Bureau of Economic Research. http://www.nber.org/papers/w13617.

———. (2006). Teacher-student matching and the assessment of teacher effectiveness. *The Journal of human resources*, 41 (4), 778-820.

Darling-Hammond, L. (1999). *Teacher quality and student achievement: a review of state policy evidence*. Seattle: University of Washington, Center for Teaching and Policy.

Harris, D., & Saas, T. (2007). *Teacher training and teacher productivity*. Working paper. Florida State University.

Hill, H., Rowan, B., & Ball, D. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42 (2), 371-406. doi:10.3102/00028312042002371.

Ladd, H. (2008). Teacher effects: what do we know? In G. Duncan & J. Spillane, *Teacher quality: broadening and deepening the debate* (pp. 3-26). Evanston, IL: Northwestern University.

McCaffrey, D., Lockwood, J., Koretz, D., Louis, T., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, 29 (1), 67-101. doi:10.3102/10769986029001067.

Monk, D.H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review*, 13 (2), 125-145. doi:10.1016/0272-7757(94)90003-5.

Nye, B., Konstantopoulos, S., & Hedges, L. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26 (3), 237-257. doi:10.3102/01623737026003237.

Rivkin, S., Hanushek, E., & Kain, J. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73 (2), 417-458. doi:10.1111/j.1468-0262.2005.00584.x.

Rockoff, J.E. (2004). The impact of individual teachers on student achievement: evidence from panel data. *The American Economic Review*, 94 (2), 247-252.

Rothstein, R. (2004). Class and schools: Using social, economic, and educational reform to close the achievement gap. Washington, DC: Economic Policy Institute.

Soares, J. F., Fonseca, I., Alvares, R., & Guimaraes, R. (2012). Exclusão intra-escolar nas escolas públicas brasileiras: um estudo com dados da Prova Brasil 2005, 2007 e 2009. *Debates ED*, 4, 1–77.

Wayne, A.J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: a review. *Review of Educational Research*, 73 (1), 89-122. doi:10.3102/00346543073001089.