Variations in Teacher Preparation as Determinants of Variation in Student Performance

David Almonte

Emory University

Table of Contents

Abstract

Teacher preparation for the implementation of the Common Core State Standards (CCSS) has, to date, been the subject of relatively little research. This paper builds one existing literature by using quantitative methods, including OLS multiple regression analysis, to identify the relationship between methods of teacher preparation and gains in student performance on standardized tests. Results indicate that 1) teachers who perceive themselves to be more prepared for standards see greater year-to-year improvements in student test scores and 2) collaborative planning time dedicated to familiarity with standards improves scores more than time dedicated to curriculum alignment 3) teacher membership in professional learning communities may induce performance gains in students. Policymakers and administrators seeking to maximize student performance should thus consider such in the design of education policy.

Variations in Teacher Preparation as Determinants of Variation in Student Performance

The importance of effective education policy on outcomes such as global competitiveness, social inequality, and public health is rarely understated. To that end, significant efforts have been made in recent years to develop and implement improved, broadly consistent standards for K-12 education. The Common Core State Standards (CCSS) represent the culmination of years of investment and research into math pedagogy, with widespread involvement from both public sector and private sector stakeholders. As such, the rollout of CCSS can be understood using the analytic strategies and quantitative methods of policy evaluation and analysis.

Variation in public policy outcomes is unavoidable. Understanding the reasons for variation in programmatic effects should consequently be a key component of the evaluation process and should be undertaken to ensure that future policy implementations are optimally designed. Evaluation research typically looks to explain variation in program outcomes by addressing underlying differences in either program beneficiaries, program administrators, or program content (Weiss, 2013). While a strong body of research exists analyzing CCSS varying effects on varying student groups (Minor, 2016), and on the effects of differing content strategies (Rappaport, 2017) there are significant gaps in research understanding the effects of "the level of teacher knowledge and perceptions of CCSS." (Nadelson & Pluska, 2014) Specifically, existing research lacks insight into the effects of the specific strategies and tools teachers use to prepare for new standards. Nadelson and Pluska, who conduct the most rigorous assessment of teacher preparedness under CCSS to date, note that their "search of the literature failed to reveal any reports of educators' perceptions and knowledge of CCSS." Their research identifies teacher preparedness as a potential determinant of student performance, but they lack the quantitative methods necessary to link teacher preparation to student outcomes.

As such, we seek to build on the existing literature on teacher preparedness for CCSS by using quantitative methods to identify the association between variation in teacher preparation methods the variation in student performance gains under CCSS. An OLS multiple regression model is constructed to test whether there are changes in student performance associated with the implementation of different preparation mechanisms. Results may inform best practices for education policy and may guide teachers in how to best allocate their time preparing for standards.

Methods

We use data from the MRDC, a nonprofit education and social policy research organization which conducted an evaluation of Powerteach methods under CCSS from 2012 to 2016. The evaluation data contains observations for 103,670 students and 432 teachers across 58 schools in five districts. Student data contains relevant student background characteristics, including their school, gender, race, income, free lunch status, exposure to Powerteach, and baseline test scores. Teacher data includes school identifiers, alongside responses to the MRDC's questionnaire on teaching methods. Binary variables are available detailing the presence of key independent variables for teachers.

In order to conduct random assignment of middle schools to teach Powerteach methods, districts had to have multiple middle schools. Consequently, the evaluation underrepresents rural middle schools, and evaluation participants are largely non-representative of US national demographics. 50.5% of students in the sample are Hispanic, compared to 23% nationally: only 11.5% of sample students were white, relative to a national average of 51.5%. While this may pose a limitation to the external validity of this research, there isn't strong reason to expect there to be strong variation in the effect sizes across rural and urban settings.

Measures

The 58 schools involved in the MRDC’s evaluation reported data on students including their scores on standardized tests in a baseline year and the year after. We use test scores as an instrument for measuring student performance, since test scores are highly suitable for quantitative analysis. Other potential measures of student performance, including grades, were unavailable. To account for variation in the administration of exams across the 4 evaluation states, scores are reported in standard deviations from the state’s mean score. Improvements in student performance are identified by subtracting scores on a follow up examination from scores on the baseline. The difference in scores is also normally distributed (Figure 1).

Teachers preparedness metrics are drawn from responses to the Common Core Teacher Survey, self-reported questionnaire assessed to all teachers involved in the MRDC evaluation. Some questionnaires were incomplete, meaning that key dependent variables often had incomplete response sets. Teachers who failed to completely respond to all relevant questions were excluded from analysis. Fortunately, we do not expect a relationship between non-response and any key variables; such a relationship would bias our regression effect estimates.

Independent Variables

Because students have likely had different teachers across both years of the evaluation and because student data lacks teacher identifier information, it is impossible to link students to particular teachers. Consequently, there is no direct way to identify, for each student, the preparation methods their teachers used. Teachers and students are, however linked by their school code. We thus average teacher responses for each school, and link average schoolwide teacher responses to each student as independent variables. This approach acts as a proxy for teacher direct effects. While this approach renders our effect estimates relatively imprecise, the signs of regression coefficients are expected accurately represent the direction of relationships.

Perceived Prepared to Teach CCSS

Teachers were asked to respond to the question “Do you feel prepared to teach CCSS Standards in Math.” Responses were coded 1 through 4, with one indicating the teacher felt “extremely prepared” and 2 through 4 indicating the teacher felt “somewhat,” “not prepared at all,” or unsure. This variable was recoded into a binary variable, with 1 representing teachers who felt extremely prepared and 0 representing all other teachers. A response of 1 was selected as a cut-point because very few (only 35) teachers reported either feeling unprepared or unsure of their level of preparedness. Using this cut point consequently minimizes the variance of our estimated ß.

Curriculum Adjustment Planning time

104 teachers were selected to participate in "Collaborative planning time dedicated to aligning curriculum to the CCSS Standards in Math." While we assume all teachers have dedicated some amount of time to aligning curriculum to standards, selection into *collaborative planning time* specifically may be a proxy for increased curriculum planning time and may hint at the benefits and costs to a collaborative approach to curriculum adjustment.

Standards familiarization planning time.

81 Teachers were selected to participate in "Collaborative planning time dedicated to understanding and deconstructing the CCSS Standards in Math." Selection is coded as a binary variable. Notably, there is an overlap between participation in standards familiarization and

curriculum adjustment planning time; some teachers do both, some do either while some do neither. Overlap is necessary to avoid issues with perfect multicollinearity.

Content-focused Trainings

236 teachers reported participating in Content-focused trainings. Content-focused trainings are among several kinds of instructional tools available to teachers and focus specifically on building teacher familiarity with the unique content of CCSS. The questionaire question is a simple yes/no checkbox, meaning replies are binary. There is no measure for the amount of content trainings available.

Job-Embedded Trainings

124 teachers were selected to participate in job embedded training, which “refers to teacher learning that is grounded in day-to-day teaching practice” (Croft & Coggshal, 2010). Critically, job embedded training is defined as training that takes place in the classroom, either with or without students. Job-embedded training hours are unavailable, meaning a teacher who engages in one our of job-embedded training appears in the data in the same way one who engages in such training regularly.

Professional Learning Community

186 teachers reported participating in a Professional Learning Community (PLC) focused on CCSS in math. PLCs are groups of individuals who “collaborate toward continued improvement in meeting learner needs through a shared curricular-focused vision” (Reichstetter, 2008). We expect teacher membership in PLCs to positively influence teaching ability, and by extension student performance.

Results

OLS multiple regression analysis of schoolwide survey response rates on year-to-year changes in student standardized test scores reveals statistically significant relationships for each of the 6 chosen measures of teacher preparedness (Table 1). Schools where all teachers dedicated collaborative planning time to CCSS Standards familiarization could expect scores to rise by an average 0.230 standard deviations ($ß = 0.230$, $p < .001$). Schools where teachers reported "Feeling strongly prepared to teach CCSS" saw improved test scores ($ß = 0.228$, $p < .001$). Teacher participation in professional learning communities was also associated with improved scores ($ß = 0.189$, $p < .001$). Negative predictors of student performance included the dedication of collaborative planning time to "curriculum adjustment" ($ß = -0.330$, $p < .001$), as well as the occurrence of CCSS content trainings ($ß = -0.052$, $p < .001$) and Job Embedded training on CCSS ($ß = 0.088$, $p < .001$), though, while both Content training and Job Embedded training are statistically significant, their effects are nominally minimal. CCSS Teacher preparedness metrics in the model explain only 1.35% of the total variation in test scores ($R^2 = 0.01348$).

Discussion

The positive correspondence between teacher perceptions of their preparedness for Common Core standards and student performance has multiple implications. First, there might be reason to anticipate increasing returns to CCSS implementation over time as teachers build familiarity with standards and instructional methods. As a corollary, evaluations of CCSS implementation conducted in the immediate aftermath of standards changes likely overrepresent underprepared teachers and should be discounted. Such a relationship may also have implications for the assignment of teachers to students: administrators who prioritize closing or reducing the education gap consider assigning underperforming and at-risk students to better prepared teachers. Lastly, this result gives credence to the idea that teacher self-reporting of their preparedness can be a useful predictor of their ability.

The implications of this research are strongest for the allocation of collaborative planning time. The sign of the parameter estimate for curriculum adjustment is negative, while the same sign for standards familiarization is positive, implying there is strong reason to believe that collaborative planning time is better spent on improving teacher's familiarity with CCSS. Alternatively, it may be the case that teachers who would underperform anyway resort to "teaching to the test," which reflects in the data as curriculum adjustment. Teachers who focus on building familiarity with content and standards may be able to compensate for less aligned lesson plans and may generally rely less on strictly proscribed lesson planning materials.

The effect sizes of both Content-focused trainings and Job Embedded trainings on growth in student performance are statistically significant, but nominally insignificant. It likely the case that lower performing teachers are more likely to use either form of training, which may lead to higher scores than would appear in the absence of either form of training. The effect of both may be positive, but offsetting. More detailed data collection on the allocation of teacher time may be useful for future researchers in assessing the efficacy of training tools.

Lastly, this paper confirms the intuition that membership in professional learning communities can supplement teacher resources and improve teaching performance (Reichstetter, 2008). More research into the causal mechanisms behind professional learning community involvments' effects on student performance may be necessary to conduct cost benefit analyses.

References

Croft, A., & Coggshal, J. (2010). *Job-Embedded Professional Development:.* National Compehnsive Center for Teacher Quality.

Minor, E. (2016). *Racial Differences in Mathematics Test Scores for Advanced Mathematics Students.* The University of North Carolina Press.

Nadelson, L., & Pluska, H. (2014). Educators' Perceptions and Knowledge of the Common Core State Standards. *Issues in Teacher Education, 23*(2).

Rappaport, S. (2017). *Group Work Is Not Cooperative Learning: An Evaluation of PowerTeaching in Middle Schools* . MDRC.

Reichstetter, R. (2008). Defining A Professional Learning Community: A literaturee Review. *Resarch Alert, 6*(5).

Weiss, M. (2013). *A Conceptual Framework for Studying the Sources of Variation in Program Effects.* MDRC.

Tables

Table 1

CCSS Survey Predictors of Standardized Tests Scores

Variable	ß	SE	95% CI
Constant	0.032*	0.014	[0.0039, 0.0592]
Perceived Prepared to teach CCSS	0.228***	0.013	[0.204, 0.253]
Curriculum Adjustment planning time	-0.330***	0.022	[-0.372, -0.287]
Standards Familiarization planning time	0.230***	0.017	[0.195, 0.264]
Content Trainings	-0.052***	0.014	[-0.080, -0.0245]
Job Embedded Training	-0.088***	0.017	[-0.121, -0.054]
Professional Learning Community	0.189***	0.012	[0.166, 0.212]
R^2	0.013		
F	188.7***		

Note: *N* = 82852. CI = confidence interval. The table reports results from a linear regression of teacher survey responses on student test scores. For each question, the coefficient ß represents the expected one-year change in student test scores (measured in standard deviations) when all of a school's teachers respond affirmatively to the question.

*p < .05, ** p < .01, *** p < .001.

Figures

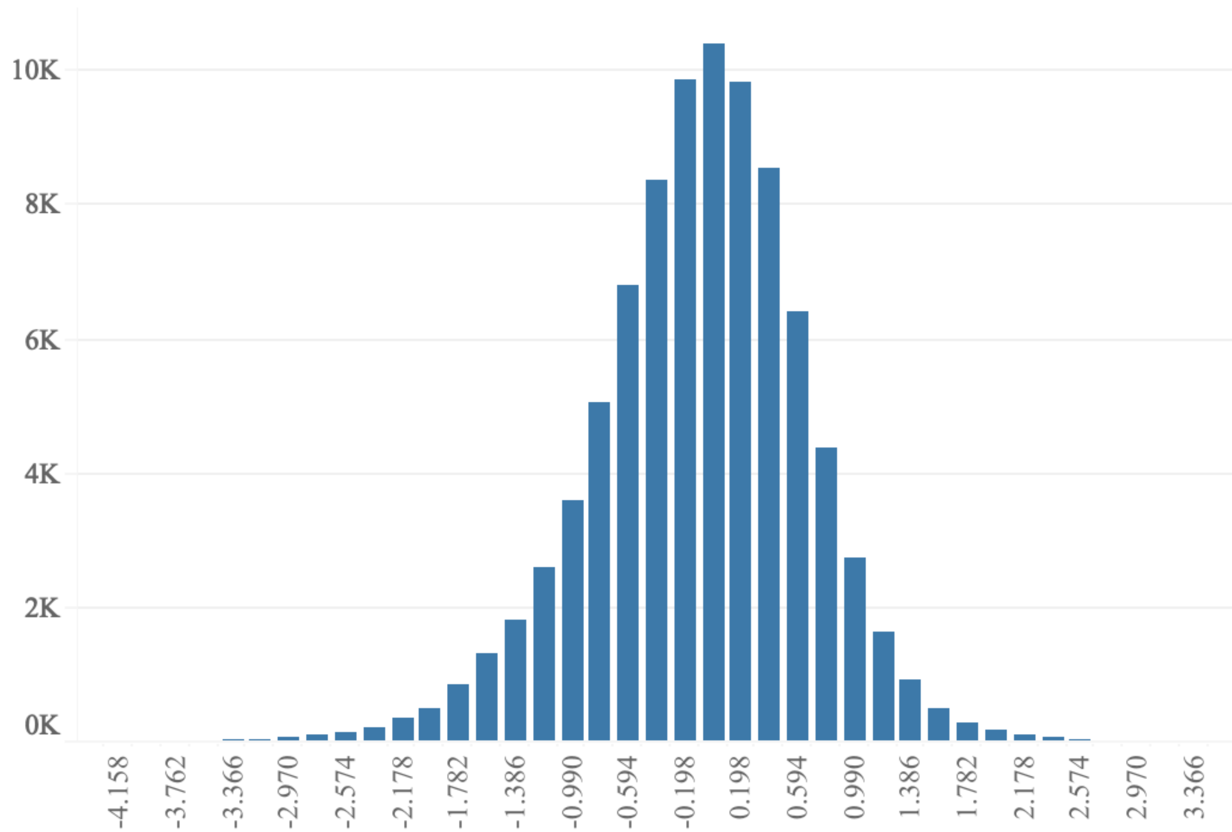

Figure 1. The distribution of the difference in baseline standardized test scores and a follow up for the MDRC evaluation of Powerteach methods (Rappaport 2017).