Educational Evaluation and Policy Analysis
September 2017, Vol. 39, No. 3, pp. 429–447
DOI: 10.3102/0162373717691611

More Than Content: The Persistent Cross-Subject Effects of English Language Arts Teachers' Instruction

Benjamin Master
RAND Corporation
Susanna Loeb
Stanford University
James Wyckoff
University of Virginia

Evidence that teachers' short-term instructional effects persist over time and predict substantial long-run impacts on students' lives provides much of the impetus for a wide range of educational reforms focused on identifying and responding to differences in teachers' value-added to student learning. However, relatively little research has examined how the particular types of knowledge or skills that teachers impart to students contribute to their longer-term success. In this article, we investigate the persistence of teachers' value-added effects on student learning over multiple school years and across subject areas. We find that, in comparison with math teachers, English language arts (ELA) teachers' impacts on same-subject standardized achievement scores are smaller in the year of instruction, but that teacher-induced gains to ELA achievement appear to reflect more broadly applicable skills that persist in supporting student learning in the long run across disciplines. Our results highlight important variation in the quality of teacher-induced learning for long-run success, distinct from the variation across teachers in more typically measured short-term learning effects.

Keywords: *achievement, educational policy, middle schools, regression analyses, school/teacher effectiveness, teacher assessment*

Introduction

Individual K–12 teachers vary substantially in terms of their effects on students' academic performance (Kane & Staiger, 2008; Nye, Konstantopoulos, & Hedges, 2004; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004). Moreover, teachers' short-term effects on tested achievement have been shown to predict meaningful long-run outcomes in students' lives, including their college attendance and future earnings (Chetty, Friedman, & Rockoff, 2011). This link between teachers' short-term instructional effects and corresponding long-run impacts provides much of the impetus for a wide range of educational reforms focused on identifying and responding to differences in teachers' effects on students' tested achievement. However, relatively little research has examined the mechanisms by which the particular knowledge or skills that teachers impart to students contribute to their longer-term success. Accordingly, policymakers frequently focus on the relative size of teachers' short-term "value-added" impacts within a subject area but attend less to the types of learning that teachers impart to students.

The distinction between the magnitude and type of learning that teachers impart may be an important one. Recent research indicates that the type of outcome used to measure teachers' short-term impact is a key factor when assessing teachers' effects on long-term outcomes. For example, researchers have frequently observed that reading or English language arts (ELA) teachers have smaller short-term impacts on students' relative achievement levels than math teachers (Kane & Staiger, 2008; Nye et al., 2004; Rockoff, 2004). However, although Chetty and colleagues (2011) similarly find that ELA teachers' effects on achievement are smaller than that of math teachers, they also find that an English teacher who raises students' reading test scores by 1 unit has an impact on long-term life outcomes approximately 1.7 times that of a teacher who does the same in math. Accounting for both of these differences, the authors find that, overall, similarly ranked math and ELA teachers predict long-run effects on students' lives to a similar degree. In a separate example, Jackson (2012) finds that ninth-grade teacher effects on both test scores and on noncognitive short-term outcomes (e.g., attendance, grade progression) independently predict teacher effects on longer term outcomes such as high school graduation and college and career aspirations. Jackson (2012) also observes that accounting for teacher effects on noncognitive outcomes is particularly important when assessing ELA teachers' long-term impacts; doing so triples the predictable variability of their effects on longer term outcomes. By contrast, accounting for the short-term noncognitive effects of Algebra teachers' instruction augments their long-term effect sizes by just 20%.

The previous research provides evidence that teachers in different subject areas impart different types of skills that vary in their long-term relevance for students and that different short-term outcomes are not necessarily interchangeable as measures of teachers' overall instructional impact (Jackson, 2012). If some types of knowledge have more generalizable or persistent impacts on students' lives, then understanding these differences across measures can lead to insights as to how teachers or other educational interventions improve students' long-run outcomes and to how best to measure teachers' contributions to student success. In particular, the available evidence suggests that ELA teachers impart meaningfully different skills than math teachers (Jackson, 2012) and that the teacher-induced learning reflected in ELA achievement gains are differentially relevant to students' future success than that of math achievement gains (Chetty et al., 2011).

In this article, we utilize data from two large urban school districts in different states to compare the generalizability and persistence of teachers' instructional effects on ELA and math achievement. Consistent with prior research, we find that ELA and math teachers' short-term effects persist at similar rates on future assessments within the same subject area. However, we observe that the learning induced by ELA teachers has substantially more long-term persistence across subject areas than the learning induced by math teachers. Our results indicate a channel through which ELA instruction influences students' long-run outcomes to a greater degree than is immediately apparent from students' short-term within-subject test score gains. Moreover, the results indicate that ELA teachers' effects on student achievement may be comparatively diffuse across subjects and over time, complicating efforts to measure instructional impacts using only within-subject measures of student learning gains during the year students spend with the teacher.

Background

Persistence of Teacher Value-Added Effects

One way to explore the relationship between teachers' impacts on short-term and longer-term student outcomes is to examine the persistence of value-added effects on academic achievement after a student leaves a teacher's classroom. Researchers have observed that only a relatively small portion of teachers' proximal year effects persists and continues to impact student achievement in a subsequent school year, with most estimates ranging from between one quarter and one third of the initial value-added effect size (Jacob, Lefgren, & Sims, 2010; Kinsler, 2012; Konstantopoulos & Chung, 2011; McCaffrey, Lockwood, Koretz, Louis, & Hamilton, 2004; Rothstein, 2010).[1] This persistent component decays less quickly in subsequent school years (Candelaria, 2016; Jacob et al., 2010). Value-added persistence rates have generally been observed to be similar in ELA and math.

Based on the limited duration of the majority of teachers' measured effects and on the broader applicability and slower decay that they observe in their persistent effects, Jacob and colleagues (2010) hypothesize that persistent learning effects represent a qualitatively different type of "long-term knowledge" that teachers can impart. In this framing, long-term knowledge represents a type of "transformational learning" that has relevance to student performance in a way that is distinct from content- or test-specific short-term knowledge. The consistency of long-term learning over time and across different grade-level assessments also suggests that teachers' effects on this long-term knowledge may be particularly relevant to students' long-run life outcomes. Researchers have generally observed only modest correlations between teachers' short- and long-term value-added effects (Candelaria, 2016; Mariano, McCaffrey, & Lockwood, 2010; Rothstein, 2010),[2] which indicates that there may be meaningful heterogeneity in the quality or types of knowledge that individual teachers impart to students.

No extant research that we are aware of has examined the persistence of teacher-induced learning effects across subjects. However, the persistence of long-term knowledge across subjects may offer valuable insight into the channels through which teachers influence students' longer-term outcomes. To the extent that learning in different subject areas persists at comparable rates across subjects, this would support the notion that different teachers' instructional effects, on average, include a comparable portion of generalizable knowledge that is useful in any academic context. Alternately, if learning of certain types is differentially persistent across subjects, then this would imply that teachers in some subject areas are providing systematically more (or less) generalizable knowledge, and likely having more (or less) future impact. ELA teaching may generate more generalizable knowledge if, for example, teacher-induced improvements in students' reading skills promote their ability to acquire knowledge in a wider array of subjects.

Cross-Subject Instructional Effects

A growing body of evidence indicates that teachers can have meaningful effects on their students' contemporaneous performance across subject areas, although much of this evidence comes from studies at the high school level. For example, Aaronson, Barrow, and Sander (2007) observe that ninth-grade math and ELA teachers in Chicago Public Schools have sizable cross-subject effects in addition to their within-subject effects. Buddin and Zamarro (2009) observe the same trend in students in Grades 9 to 11 in the Los Angeles Unified School District. Koedel (2009) finds that 9th- to 11th-grade teachers in both ELA and math in the San Diego Unified School District impact student performance on students' reading achievement, although science and social studies teachers do not. Jackson (2012), however, does not find evidence of significant cross-subject effects of ninth-grade Algebra and English teachers in North Carolina.

Both theory and some empirical evidence indicate that ELA instructional effects may be especially generalizable across subjects. For example, students' reading and language skills have been shown to be important across a range of other subjects and may be particularly important for students of lower socioeconomic status or those with limited proficiency in English (Abedi & Lord, 2001; Chang, Singh, & Filer, 2009; O'Reilly & McNamara, 2007). Moreover, in a recent study of teachers' cross-subject effects, Yuan (2014) examines teachers in middle school grades in an anonymous urban school district and finds that ELA teachers contribute to student achievement in mathematics, ELA, science, and social studies, whereas, by comparison, mathematics teachers contribute significant cross-subject effects only in ELA achievement.

Teachers' contemporaneous cross-subject effects may stem from a variety of mechanisms. For example, teachers may directly collaborate in their planning or instruction of a shared group of students, reinforcing a common set of skills. Alternately, knowledge generated in one subject area may consist of generalizable skills that directly support students' learning in an otherwise unrelated subject area. Thus, it is difficult to disentangle the different ways in which teachers may be contributing to their peers' contemporaneous cross-subject performance. Research on cross-subject spillover to-date generally does not distinguish between contemporaneous effects of peer collaboration in instruction and differences in the generalizability of acquired student knowledge.

Contribution

In this study, we expand upon the prior research related to both teachers' persistent effects and to their cross-subject effects. Following a conceptual and methodological framework described by Jacob and colleagues (2010), we investigate the persistence of teachers' value-added effects, distinguishing between short-term, test-specific knowledge and longer-term learning that accumulates. We extend this framework by differentiating between long-term knowledge that is content-specific and generic knowledge that persists across subjects.

Our approach allows us to identify the persistent impacts of the subject-specific learning that prior-year teachers impart to students. As a consequence, we are better able to isolate the extent to which different types of teacher-induced learning generalize across subject areas, distinct from teacher spillover effects that stem from contemporaneous instructional collaboration. Finally, to better gauge the generalizability of our findings, we consider evidence from two large urban school districts that are located in different states and that utilize quite different standardized assessments of students' performance in ELA and math.

We specifically consider the following research questions:

Research Question 1: What is the rate of persistence of previously assessed ELA and math knowledge within and across subjects?

Research Question 2: What is the rate of persistence of teachers' value-added effects on student learning, within and across subjects?

Research Question 3: How does the relative magnitude of ELA and math teachers' effects on achievement differ when considering short-term, subject-specific effects versus persistent, multisubject effects?

The remainder of the article proceeds as follows. In the section "Data," we describe the data utilized in the study. The "Method" section details our methods for estimating value-added measures and value-added persistence within and across subjects. In the "Results" section, we detail our results, and we conclude in the final section "Conclusion and Discussion" with a discussion of conclusions and potential limitations.

Data

Administrative Data

To investigate the within- and across-subject persistence of teachers' value-added effects, we draw upon extensive administrative data about students, teachers, classrooms, and schools in two large urban school districts: New York City (NYC) and Miami-Dade County Public Schools (M-DCPS). In both NYC and M-DCPS, our available data include students in third through eighth grade from school years (SY) 2003–2004 through SY 2011–2012. Both district data sets include data on students' annual achievement test scores in ELA and math, and identification of students' primary teacher and classroom in each year and subject area. For the purposes of our analysis, we standardize students' achievement test scores within each grade, subject, and year in each district.

The content assessed by the annual New York State ELA and math student achievement tests has been aligned with the State's content standards in Grades 3 through 8 throughout this period of time, and exams in each subject area include a mix of multiple-choice and open-response questions. The ELA exams primarily assess students' comprehension of reading passages and writing ability, whereas math exams address a range of topics including number sense, algebra, probability, and geometry, with overlapping topics across grades. Similarly, student assessments provided by M-DCPS include scores on the statewide Florida Comprehensive Assessment Test (FCAT), in the subject areas of reading and math. FCAT exams in reading and math consist primarily of multiple-choice questions, but also include a small subset of short or extended open-response questions. The reading exams primarily assess students' comprehension of reading passages. Student writing ability is tested via a separate exam in Florida that is not administered across Grade Levels 3 to 8. As such, assessment scores from M-DCPS include much less data on student writing skills than assessment data from NYC. As in NYC, math and reading FCAT exams include content corresponding to the state's content standards.

In addition to achievement data, we have access to a rich set of demographic and behavioral characteristics that we utilize in our analyses. For both NYC and M-DCPS students, these characteristics include their race/ethnicity and home language, as well as their absences, suspensions, school transfers, free or reduced-price lunch status, disability or special education status, and English language learner status in each school year.

We estimate teacher value-added effects using a sample of student-year records in Grades 4 through 8 for whom current and prior year achievement data are available. However, because we are investigating the persistence of teachers' value-added effects on student achievement in the year(s) *after* they teach a student, the subsamples for our persistence analyses are restricted to student-year observations in which we can identify prior-year teachers with appropriate value-added scores, as detailed in the "Method" section below. In practice, these requirements reduce our analysis of teachers' persistent effects on student outcomes to student-year observations in Grades 5 through 8, from SY 2005–2006 through SY 2011–2012. This sample includes only students who are present in each district for at least three consecutive years. Finally, as explained in our "Method" section below, we only estimate cross-subject persistence for students whose lagged math and ELA teachers are not the same person, further restricting our sample. Summary statistics for each district's analytical sample are presented in Table 1.

The grade levels of students represented in our analyses vary by district and as a function of whether we are investigating 1- versus 2-year persistent effects. In M-DCPS, a sizable portion of students in Grades 4 and 5 are taught by different teachers in their ELA and math classrooms, and we can thus estimate persistent effects of students in the fifth through eighth grade. However, in NYC very few students are taught by different teachers in ELA and math until the sixth grade, and therefore we limit our analysis of persistent effects to students in the seventh or eighth grade. Moreover, when evaluating teachers' 2-year persistent effects, our sample is further restricted (to students in sixth through eighth grade in M-DCPS and in eighth grade in NYC). Table 2 provides a summary of our available sample sizes by grade in each district and for each type of analysis.

TABLE 1

Summary Statistics for Students, Teachers, and Schools in the Analytical Sample for Each District

	NYC	M-DCPS
A. Students		
% Free or reduced-price lunch	71.2	64.8
% Black	28.4	22.5
% Hispanic	36.8	65.4
% White	17.2	9.5
% Asian or Other	17.2	2.5
% Female	51.6	50.9
N of distinct students	283,771	132,102
B. Teachers		
N of distinct ELA teachers	5,863	4,254
N of distinct math teachers	5,036	3,649
C. Schools		
Average % of students eligible for free or reduced lunch	74.6 (21.4)	65.9 (22.2)
Average % of students Black	28.1 (27.9)	22.8 (28.9)
Average % of students Hispanic	40.0 (25.8)	65.4 (28.0)
N of distinct schools	531	354

Note. Analytical samples include students in Grades 5 through 8 in school years 2005–2006 through 2010–2011 for whom necessary historical math and reading teacher information is available to estimate 1-year persistence both within- and across subject areas. NYC = New York City; M-DCPS = Miami-Dade County Public Schools; ELA = English language arts.

Method

Teacher Value-Added Measures

To examine the persistence of teachers' effects, we first generate teacher value-added estimates of the effects of each teacher on tested student achievement in each year. We employ a value-added model that has been used by the NYC Department of Education in the past to evaluate their teachers' performance (Value-Added

TABLE 2

N *of Student-Year Observations, by Grade Level and Subject, in Analytical Samples Used for Estimation of 1- and 2-Year Persistence in Each District*

	NYC				M-DCPS			
	Grade 5	Grade 6	Grade 7	Grade 8	Grade 5	Grade 6	Grade 7	Grade 8
Math								
N eligible for 1-year persistence sample	—	—	153,141	179,278	41,214	48,274	60,244	61,523
N eligible for 2-year persistence sample	—	—	—	123,628	—	23,613	29,494	36,399
ELA								
N eligible for 1-year persistence sample	—	—	152,753	179,562	41,214	48,274	60,244	61,523
N eligible for 2-year persistence sample	—	—	—	123,618	—	24,083	29,530	36,650

Note. Eligible samples are restricted to students taught by different math and reading teachers with value-added scores in the prior school year or years. In NYC, students in Grades 5 and below are taught almost exclusively by dual-subject teachers. In M-DCPS, however, students in Grades 4 and 5 are taught by different subject-specific teachers in many schools. NYC = New York City; M-DCPS = Miami-Dade County Public Schools; ELA = English language arts.

Research Center, 2010). Conceptually, this model compares teachers to other "similarly circumstanced" teachers by first predicting students' achievement with both prior achievement measures (math and ELA) and a range of observable student, classroom, and school characteristics that may influence their achievement, and then attributing the remaining unexplained variation in student performance to individual teachers. The Appendix provides details of our value-added model specification.

As an additional specification check, we also duplicate our main results using an alternative value-added model that includes teacher fixed effects in combination with student and classroom control variables directly in our regression, rather than post hoc aggregation of student residuals. We provide details regarding this alternative model and corresponding results with respect to within- and across-subject persistence of teacher induced learning in the Appendix.

Estimating the Persistence of Teacher Value-Added Effects Within and Across Subjects

We estimate the persistence of teachers' value-added effects within and across subject areas using a modified version of an instrumental variables approach first described by Jacob et al. (2010). As previously discussed, these authors conceptualize students' tested knowledge as a combination of "short-term" knowledge that has no observed impact on future achievement, and "long-term" knowledge that is relevant to both contemporaneous and future achievement tests. In their formulation, observed student achievement Y in a given period represents a combination of that student's long-term knowledge from a prior period and all contemporaneous impacts (including teachers' effects) that influence both their long-term and short-term knowledge in the current period:

$$Y_t = \theta y_{l,t-1} + \eta_t^l + \eta_t^s. \tag{1}$$

Here, current achievement is a function of contemporaneous impacts η_t^l and η_t^s on long- and short-term knowledge, as well as prior long-term knowledge from the period $y_{l,t-1}$, which carries forward with some rate of decay $(1 - \theta)$.

In practice, achievement test scores do not directly reflect long-term knowledge, but rather the combination of long-term and short-term knowledge assessed in the prior period, Y_{t-1}. In light of this challenge, Jacob and colleagues (2010) use an instrumental variables approach to estimate the decay of prior long-term knowledge, using twice lagged achievement Y_{t-2} as an instrument for Y_{t-1} to predict current achievement, Y_t. This removes the short-term knowledge component of Y_{t-1}. They find that the vast majority of a student's previously assessed long-term knowledge in the same subject area persists between one year and the next, with coefficients very close to 1. This serves as a benchmark for estimating the proportion of teachers' value-added contributions to student learning that reflect this kind of persistent, "long-term" knowledge.

Following a similar approach, we estimate the proportion of teachers' total instructional effects that consist of long-term knowledge in a given subject by instrumenting each student's lagged knowledge (Y_{t-1}) with their lagged teachers' estimated contributions (value-added) to that knowledge to predict current achievement (Y_t). Because students' lagged test scores in a given subject area may have been influenced contemporaneously by either their same or alternate-subject teacher's value-added quality, we include in our estimation both the same and alternate-subject lagged teacher quality measures as instruments for both the same and alternate-subject test scores in the lagged year. This approach allows us to mitigate possible bias from contemporaneous spillover effects across the two subjects in the year of instruction. In addition, because teachers' value-added scores in any given year include estimation error that is correlated with other classroom-specific learning effects in that year, we calculate, for each student in each subject area and year, their lagged teacher's average value-added quality across all years other than the year in which they taught that student, expressed as $T_{ijt-1} = \sum_{y \neq t-1} M_{jy}$. Our second-stage equation for estimating the persistence of teacher value-added in a particular subject area is as follows:

$$Y_{icjt} = \theta Y_{it-1} + \theta^{alt} Y_{it-1}^{alt} + \beta X_{it} + \pi_{cjt} + \varepsilon_{ijt}. \qquad (2)$$

Here, our subscripts refer to each student i in time t, classroom c, and taught by teacher j. The values of T_{ijt-1} and T_{ijt-1}^{alt} for students' lagged teachers in the same and alternate subject areas serve as the two excluded instruments for students' prior test scores (Y_{it-1} and Y_{it-1}^{alt}) in both subjects in the first stage.

Because student assignment to teachers is nonrandom, the measured quality of a student's lagged teachers may be correlated with the quality of their current teacher in each subject. To minimize possible bias in our teacher persistence estimates due to persistent nonrandom patterns of teacher assignment, we include in Model (2) controls for both student covariates χ, and for contemporaneous classroom fixed effects π (which also incorporate school, year, and grade fixed effects) in the subject for which we are estimating persistent effects. In this formulation, the persistence of student learning in a subject is a function of variation in the measured quality of both of their lagged teachers, distinct from the effects of the student's school or their same-subject classroom assignment in the current year. In addition, to avoid possible bias due to the correlation between an individual teacher's quality across multiple subjects that they teach (Loeb & Candelaria, 2012), we exclude from our analyses any students taught by the same individual teacher in both ELA and math in the lagged year. As shown in Table 2, this restriction leads us to focus our analyses primarily on students initially taught in middle school grades where dual-subject teaching is less common.

Although we include controls for student characteristics and for contemporaneous school and classroom assignments, our estimates of the within-subject and cross-subject persistence of teacher-induced learning could still be influenced by two potential sources of bias. First, our estimates may be biased if schools systematically adjust the instructional inputs (other than classroom assignments) that students receive as a response to the quality of a prior-year teacher. This could occur, for example, if effective teachers raise students' achievement and this in turn leads schools to provide fewer instructional supports to those students. It seems unlikely, however, that schools would be sufficiently attuned to differential student achievement gains (rather than levels) to make this a major factor. Second, our estimates of cross-subject persistence could be biased if teacher value-added measures themselves are biased due to nonrandom within-school student sorting to instructional experiences of differing quality and if, additionally, this sorting bias systematically affects teacher value-added estimates in one subject area more than

TABLE 3

Estimates for the Persistence of Observed Knowledge and Long-Term Knowledge, Within Subjects

	NYC			M-DCPS		
	Observed knowledge	Long-term knowledge Model 1	Long-term knowledge Model 2	Observed knowledge	Long-term knowledge Model 1	Long-term knowledge Model 2
Predicting math with math knowledge						
Coefficient on lagged achievement	0.808 (0.001)	0.991 (0.001)	1.000 (0.002)	0.816 (0.001)	0.968 (0.002)	0.970 (0.003)
N of student-year observations		376,823			229,558	
Predicting ELA with ELA knowledge						
Coefficient on lagged achievement	0.663 (0.001)	0.974 (0.002)	0.949 (0.003)	0.777 (0.001)	0.971 (0.002)	0.990 (0.003)
N of student-year observations		376,715			244,260	

Note. Coefficient for Observed Knowledge from a regression of current achievement on prior achievement in the same subject. Coefficient for Long-Term Knowledge in Model 1 is from an instrumental variables (IV) regression of current achievement on prior achievement in either subject instrumented with twice-lagged achievement in the same subject. Model 2 further includes a control for lagged achievement in the alternate subject. First-stage *F* statistics for all IV regressions exceed 100,000. NYC = New York City; M-DCPS = Miami-Dade County Public Schools; ELA = English language arts.

another. We discuss the implications of this potential caveat to our results in our discussion section.

Estimating the Persistence of "Long-Term Knowledge" Across Subjects

In addition to estimating the persistence of teacher-induced learning across subjects, we also investigate the persistence of previously assessed long-term knowledge (not just teacher-induced knowledge) in one subject on future achievement in an alternate subject. To do this, we modify the instrumental variables approach employed by Jacob and colleagues (2010) to estimate within-subject long-term knowledge persistence by simply swapping the dependent variable, students' same-subject achievement, with alternate-subject achievement. This approach measures the extent to which prior-year achievement in the first subject (Y_{t-1}), instrumented using twice-lagged achievement in that same subject (Y_{t-2}), is predictive of achievement test scores in the current year in an alternate subject (Y_t^{alt}). In other words, this model examines the predictive power of knowledge that has previously shown to be persistent in one subject area on future achievement outcomes in a different subject area. We consider variations of this model with and without an additional control for students' prior achievement levels in the alternate subject (Y_{t-1}^{alt}).

Results

Persistence of Long-Term Knowledge Within and Across Subjects

We begin to address our first research question by investigating the persistence of long-term knowledge within subjects, via a two-stage regression predicting current achievement outcomes using prior achievement, instrumented with twice-lagged achievement, as previously discussed. We compare these results to persistence estimates from an ordinary least squares (OLS) model in which we predict current test scores with observed prior achievement. These results are shown in Table 3. Consistent with prior research and with our intuitive understanding of long-term knowledge, we find that nearly all of previously assessed long-term knowledge

TABLE 4

Estimates for the Persistence of Observed Knowledge and Long-Term Knowledge, Across Subjects

	NYC			M-DCPS		
	Observed knowledge	Long-term knowledge Model 1	Long-term knowledge Model 2	Observed knowledge	Long-term knowledge Model 1	Long-term knowledge Model 2
Predicting math with ELA knowledge						
Coefficient on lagged achievement	0.565 (0.001)	0.859 (0.002)	0.339 (0.003)	0.711 (0.002)	0.852 (0.002)	0.362 (0.003)
N of student-year observations		376,823			229,558	
Predicting ELA with math knowledge						
Coefficient on lagged achievement	0.630 (0.001)	0.792 (0.002)	0.486 (0.002)	0.713 (0.002)	0.813 (0.002)	0.401 (0.003)
N of student-year observations		376,715			244,260	

Note. Coefficient for Observed Knowledge from a regression of current achievement on prior achievement in the alternate subject. Coefficient for Long-Term Knowledge in Model 1 is from an instrumental variables (IV) regression of current achievement on prior achievement in the alternate subject instrumented with twice-lagged achievement in the alternate subject. Model 2 further includes a control for lagged achievement in the same subject. First-stage *F* statistics for all IV regressions exceed 100,000. NYC = New York City; M-DCPS = Miami-Dade County Public Schools; ELA = English language arts.

(i.e., knowledge that is relevant across two prior school years) within a subject area also persists into a third year. Across NYC and M-DCPS, partial coefficients on long-term knowledge range in a fairly narrow band from 0.949 to 1.000, regardless of whether our models include an additional control for prior achievement in the alternate subject. The coefficients on long-term knowledge persistence are in contrast to those for students' observed prior-year test scores, which reflect a mix of short- and long-term knowledge (including measurement error) and persist at substantially lower rates. Lower observed knowledge persistence in ELA may correspond to a higher degree of measurement error in ELA assessments.

Next, in Table 4, we examine how previously assessed long-term knowledge in one subject area predicts current performance in the alternate subject. Here, as previously, we consider a simple OLS regression across subjects, as well as a long-term knowledge "Model 1" that, as previously discussed, predicts current achievement in a subject with instrumented prior achievement in the alternate subject. We also consider an alternate long-term knowledge "Model 2" that includes a control for prior-year achievement in the subject of interest.

Our results differ by model specification. Results for the OLS regression continue to be less than 1 and cross-subject coefficients are smaller than the same-subject coefficients reported in Table 3. Results from long-term knowledge Model 1 indicate that knowledge that is relevant over time and across assessments in one subject area persists at a substantial, but not perfect, rate into the alternate subject. Coefficients when predicting math with long-term ELA knowledge range from 0.852 to 0.859, whereas coefficients for predicting ELA with long-term math knowledge are somewhat smaller, ranging from 0.792 to 0.813. In other words, in contrast to our within-subject results, we observe that not all of the long-term knowledge that is relevant within a subject is relevant across subject areas. Including controls for prior achievement in the subject of interest (long-term knowledge Model 2) substantially reduces these coefficients, with a larger reduction in the adjusted persistence of long-term ELA knowledge (0.339 to 0.362)

TABLE 5

Estimates for the 1-year Persistence of Teachers' Value-Added Effects Within and Across Subjects

	NYC		M-DCPS	
	Persistent impacts on math scores	Persistent impacts on ELA scores	Persistent impacts on math scores	Persistent impacts on ELA scores
Same-subject teacher persistence coefficient	0.289 (0.023)	0.257 (0.087)	0.411 (0.027)	0.389 (0.076)
Across-subject teacher persistence coefficient	0.179 (0.069)	0.013 (0.026)	0.178 (0.056)	0.020 (0.027)
First-stage F statistic	51.098	45.183	86.264	74.517
N of student-year observations	329,989	329,855	209,722	209,183

Note. Coefficients for teacher value-added effects from regressions of current achievement in a given subject area on prior achievement in both subject areas instrumented with the prior year teachers' value-added quality in both subject areas. Models include controls for current student characteristics and classroom fixed effects from the test subject of the dependent variable. One-year persistence coefficients reflect the portion of initial teacher effects that continue to impact achievement in the following school year. NYC = New York City; M-DCPS = Miami-Dade County Public Schools; ELA = English language arts.

versus long-term math knowledge (0.401 to 0.486). Collectively, these results indicate that much, but not all, of the within-subject long-term knowledge that students possess reflects foundational skills, abilities, or other human capital that are relevant to student performance across multiple academic subjects. Controls for prior achievement in the same subject partial out much of this generalizable knowledge, and controls for prior math achievement reduce coefficients more, perhaps because they include less measurement error than ELA measures with respect to these generalizable skills.

Overall, our results are consistent with the idea that students' foundational knowledge and skills in any subject are predictive of performance across multiple subjects. However, it does not necessarily follow that the cross-subject effects of teacher-induced learning are the same for math and ELA teachers. Some teacher-induced learning is reflected in long-term knowledge, but our approach thus far to isolating this persistent knowledge does not distinguish between learning induced by teachers and other types of knowledge.

The Persistence of Teacher-Induced Learning Within and Across Subjects

To address our second research question related to teacher-induced learning, we examine the extent to which teachers' value-added effects on student learning persist over time (i.e., consist of long-term knowledge) within and across each subject area. In Table 5, we show results from a two-stage regression in which we predict current achievement in either subject area with students' prior achievement in both the same and the alternate subject, instrumented by the prior-year estimated value-added quality of both their same-subject and alternate-subject teachers. In NYC, we find that 26% to 29% of teachers' within-subject value-added effects persist into a subsequent school year, with similar results across math and ELA teachers. One-year within-subject persistence rates are higher in M-DCPS in both subjects, but are also fairly similar across math (0.411) and ELA (0.389). Standard errors are larger when estimating persistent impacts of ELA instruction than when estimating persistent impacts of math instruction.

The differences that we observe in the magnitude of general value-added persistence across our two districts may stem from a range of factors, including differences in the year-over-year alignment and content covered by state tests in each district. Overall, however, our estimates of same-subject teacher persistence in both districts are in a comparable range to those reported in prior research, and align with previous research showing roughly equivalent persistence rates of teacher-induced learning in ELA and math.

TABLE 6

Estimates for 1-Year and 2-Year Persistence of Teachers' Value-Added Effects Within and Across Subjects, for a Common Sample of Students and Teachers

	NYC		M-DCPS	
	Persistent impacts on math	Persistent impacts on ELA	Persistent impacts on math	Persistent impacts on ELA
1-year				
Same-subject teacher persistence coefficient	0.336	0.278	0.417	0.516
	(0.034)	(0.049)	(0.034)	(0.099)
Across-subject teacher persistence coefficient	0.095	−0.017	0.223	−0.009
	(0.038)	(0.041)	(0.071)	(0.037)
First-stage *F* statistic	178.473	187.112	53.285	41.144
2-year				
Same-subject teacher persistence coefficient	0.134	0.173	0.250	0.240
	(0.032)	(0.041)	(0.031)	(0.076)
Across-subject teacher persistence coefficient	0.073	0.001	0.062	0.044
	(0.034)	(0.034)	(0.065)	(0.033)
First-stage *F* statistic	258.642	239.696	82.950	82.507
N of student-year observations	122,455	122,420	87,628	88,063

Note. Coefficients for teacher value-added effects from regressions of current achievement in a given subject area on prior achievement in both subject areas instrumented with the prior year teachers' value-added quality in both subject areas. Models include controls for current student characteristics and classroom fixed effects from the test subject of the dependent variable. NYC = New York City; M-DCPS = Miami-Dade County Public Schools; ELA = English language arts.

In contrast, across both NYC and M-DCPS, we observe stark differences in the rate of cross-subject persistence between ELA and math teachers, with ELA teacher effects persisting at a much higher rate across subjects. For example, in NYC, ELA teachers' persistence coefficient across subject is 0.179, which is 70% of their persistence rate within the same subject (0.257). In contrast, NYC math teachers' persistence coefficient across subjects is 0.013, less than 5% of their same-subject persistence (0.289).

The differential cross-subject persistence of ELA teachers' instruction is also apparent in M-DCPS. In this district, ELA teachers' estimated cross-subject persistence rate (0.178) is approximately 46% of their within-subject persistence (0.389). In contrast, M-DCPS math teachers' cross-subject persistence rate (0.020) is less than 5% of their estimated within-subject persistence (0.411).

Next, in Table 6, we show the same- and cross-subject persistence of teacher-induced learning 2 years following instruction. Because our available samples for estimating 2-year persistence are smaller and reflect a different composition of students by grade level, we also include here results for estimates of 1-year persistence of instruction for the same teachers teaching the same students. For example, we report on the 1-year persistence of a student's teacher from Grade 6 on Grade 7 achievement, alongside the 2-year persistence of that same student's teacher from Grade 6 on Grade 8 achievement. Presenting both 1-year and 2-year persistent estimates for this common sample allows us to distinguish between any trend in results that relate to our sample rather than from differences in patterns of 2-year versus 1-year cross-subject persistence.[3]

In this restricted sample, 1-year persistence rates are somewhat higher than in our full sample, particularly in M-DCPS. However, we continue to observe no cross-subject 1-year persistence of math teacher effects on ELA in either district, with point estimates of −0.017 in NYC and −0.009 in M-DCPS. In contrast, ELA teachers' 1-year persistent effects are substantial, with point estimates of 0.095 and 0.223 in NYC and M-DCPS, respectively.

In line with our results for 1-year persistence, we observe differential cross-subject persistence of teacher-induced learning in ELA 2 years after instruction, particularly in NYC where our estimates are more precise. In NYC, the estimated 2-year within-subject ELA persistence rate of teacher-induced learning (0.073) is approximately 42% of the within-subject 2-year persistence estimate (0.173). In contrast, we observe very little 2-year cross-subject persistence for math teachers' instruction (0.001), an estimate that is less than 1% of math teachers' within-subject 2-year persistence estimate (0.134). In M-DCPS, our 2-year persistent results are inconclusive with respect to differences in persistence across subjects. ELA teachers' 2-year cross-subject persistence in this case (0.062) is 26% of their within-subject persistence (0.240), whereas math teachers' 2-year cross-subject persistence estimate (0.044) is around 18% of their 2-year persistence estimates within the same subject (0.250). However, in M-DCPS, the estimates for 2-year persistence are at odds with the measured (lack of) persistence of those same teachers' 1-year effects across subjects, and the relatively large standard errors around our 2-year estimates in this district limit our ability to draw clear conclusions from the results.

Overall, our analyses of the persistence of teacher-induced learning indicates that ELA teachers' impacts on students' assessed ELA skills are substantially more persistent in math than math teachers' impacts on future-year ELA performance. This pattern is largely consistent for both 1-year and 2-year persistence measures, particularly in NYC where our point estimates are more precise. Learning due to ELA instruction appears to impart long-term knowledge and skills that are reflected not only in short-term ELA scores, but also in future test scores in both subject areas.

Comparing the Relative Magnitude of Teacher-Induced Learning in ELA and Math

The previous results suggest that instructional effects captured by students' ELA test scores are qualitatively different than those captured by math test scores. This difference manifests as larger cross-subject persistent effects for ELA learning. We are unable to assess the extent to which ELA learning plays out differently with respect to nonacademic outcomes, but in light of our results it is worth reevaluating the relative magnitude of ELA teachers' instructional effects on academic achievement when accounting for persistent cross-subject effects. To do this, we consider the relative impact of ELA and math teachers both with respect to longer-term (i.e., persistent) knowledge in particular, and for the combination of short- and longer-term knowledge. In Table 7, we provide estimates of the magnitude of teachers' effects on long-term knowledge across multiple subjects and on the combination of short-term within-subject and longer-term, multisubject knowledge.

In the first column of Table 7, we report our unshrunken estimates of the impact of a one standard deviation difference in ELA and math teachers' effectiveness on adjusted student achievement in the same-subject in the year of instruction. Our results, which range from 0.235 to 0.283, are consistent with estimates from prior research on the raw standard deviation of teachers' value-added effects (Rockoff, 2004), and in both districts our estimates are consistent with a common finding of ELA value-added effects that are somewhat smaller than (approximately 84% to 88% of) math value-added effects. When considering same and cross-subject effects that persistent into a second school year, however, the situation is reversed. For this multisubject longer-term knowledge, ELA teachers' estimated value-added effects are somewhat larger than (approximately 110% to 126% of) math value-added effects. Accordingly, when we estimate teachers' combined same-year and persistent impact on student test scores and account for cross-subject persistence, ELA teachers' relative magnitude of impact on achievement moves substantially closer to that of (approximately 92% to 97% of) math teachers' effects. Note that this sizable improvement in ELA teachers' estimated magnitude of impact comes from accounting for just cross-subject persistence in the first year following instruction, and thus likely represents a lower bound for ELA teachers' overall contributions. It does not account for other factors that may also differ across subjects, including possible cross-over instructional

TABLE 7

Estimated Relative Magnitude of Impacts of ELA and Math Instruction on Student Achievement Scores in ELA and Math Over 2 Years, Accounting for Cross-Subject Effects in Year 2

	Same-year test score impact	Same-subject 1-year persistence	Cross-subject 1-year persistence	Total 1-year persistent impact	Total impact
NYC					
ELA instruction	0.248	0.064	0.044	0.108	0.356
Math instruction	0.283	0.082	0.004	0.086	0.368
Ratio of ELA to math impacts	87.6%	78.0%	1,100.0%	125.6%	96.7%
M-DCPS					
Reading instruction	0.235	0.091	0.042	0.133	0.368
Math instruction	0.281	0.115	0.006	0.121	0.401
Ratio of ELA to math impacts	83.6%	79.1%	700.0%	109.9%	91.7%

Note. Same-year impacts are estimates of unshrunken true teacher value-added effects in the year of instruction. Estimated same- and cross-subject persistence effects are a function of the point estimates for 1-year persistence rates from Table 5, multiplied by the estimated size of teacher effects in the year of instruction. Total persistent effects combine same-and cross-subject 1-year persistence impacts. Total impacts combine within-subject impacts in the year of instruction (i.e., same-year), and multisubject persistent impacts in the year following instruction. ELA = English language arts; NYC = New York City; M-DCPS = Miami-Dade County Public Schools.

effects in the year of instruction (which we do not measure), any cross-subject effects on additional academic subjects such as science or social studies, or any cross-subject effects of ELA instruction on math (or the reverse) that are not mediated by ELA (or math) academic knowledge.

Conclusion and Discussion

This study offers new insight into K–12 teachers' instructional effects, contributing to our understanding of how teachers' efforts benefit students over time and on different assessments. We identify important differences in the type of learning that teachers of different subjects impart, as evidenced by a much higher rate of cross-subject persistence associated with teacher-induced learning in ELA than in math. These results may help to explain why ELA teachers' effects on student achievement have been shown to predict students' long-run outcomes to a similar degree as math teachers' effects, in spite of smaller contemporaneous effects on within-subject standardized achievement scores (Chetty et al., 2011). The knowledge that ELA teachers impart and that is captured by ELA achievement scores may be differentially important to student learning more broadly. We hypothesize that due to their broader applicability, a given increase in ELA skills may yield greater long-term benefits to students' life outcomes than a comparable increase in math skills. In line with this theory, the magnitude of ELA teachers' estimated effects on measured achievement is substantially increased, relative to that of math teachers, when we examine longer-term knowledge over multiple subjects rather than only short-term, within-subject knowledge.

Although our study focuses on teachers' contributions to student learning, our findings may also have broader implications for other interventions that aim to develop students' skills in reading and language arts. The results suggest that a variety of educational interventions that focus on ELA skill development may yield long-term, generalizable benefits that are not fully captured by observed gains on ELA assessments in the year of instruction. Investments in developing students' ELA skills, whether through classroom instruction or other means, may yield larger benefits to students than are generally recognized or immediately apparent. For example, across our two samples we estimate that between 32% and 41% of ELA teachers' persistent effects on achievement consist of cross-subject impacts on math achievement. If this pattern of cross-subject impacts holds with respect to the overall academic gains from ELA-focused interventions, researchers may, by focusing only on ELA test scores, fail to identify a sizable portion of the overall benefits of developing students' skills in reading and language arts.

To the extent that teachers' instructional effects influence student achievement over time and across subject areas, educators and policymakers may miss valuable information if they rely only on short-term within-subject student learning to evaluate teachers' "value added" to student achievement. Previous research has shown that teacher value-added rankings are sensitive to modeling choices around the inclusion of cross-subject spillover effects or the persistent effects of prior-year teachers (Kinsler, 2012; Yuan, 2014) and that individual teachers' short- and longer-term academic effects are only moderately correlated (Mariano et al., 2010; Rothstein, 2010). We find that ELA teachers' contributions to student learning are particularly diffuse, and thus a larger portion of their instructional impact may typically go undetected or be ascribed to teachers in other subject areas or school years. In the context of teacher performance evaluations that are linked to student achievement gains, our results reinforce the value of attending to team-level indicators of teachers' contributions to student achievement, or to estimating value-added models that simultaneously account for multiple different subject-teachers' contributions (Yuan, 2014).

We can only speculate as to whether impacts on proximal versus future-year test scores are more relevant to students' lifelong outcomes, but nevertheless it is notable that persistent ELA instructional effects across subject areas are larger than teacher-induced persistent math effects. More persistent and generalizable effects on test scores may indicate a deeper learning that yields larger lifelong benefits for students, though we cannot rule out the possibility that short- and longer-term value-added reflect student learning of equivalent value.

Our findings do illustrate that different types of student learning play out differently in students' academic performance over time. The results corroborate prior theory and research regarding the importance of students' reading skills across a variety of contexts (Abedi & Lord, 2001; Chang et al., 2009; O'Reilly & McNamara, 2007). They also highlight the potential importance of considering multiple measures of teachers' contributions to student learning, as different measures may pick up on qualitatively different instructional contributions (Jackson, 2012).

This study also contributes to a growing body of evidence regarding teachers' cross-subject spillover effects. Although previous research has identified substantial contemporaneous spillover effects across different teachers' classrooms (Aaronson et al., 2007; Buddin & Zamarro, 2009; Koedel, 2009; Yuan, 2014), these effects may reflect multiple contributing factors, including direct instructional collaboration with shared students in the year of instruction. Our method allows us to better isolate the persistent spillover effects of different types of previously observed teacher-induced student learning, and we demonstrate that ELA learning in particular is more broadly relevant to performance in another subject area. Our study is also among the first to explore the cross-subject effects of teachers in elementary and middle school grade levels, rather than in high school contexts. Consistent with Yuan's (2014) findings for teachers in middle grades, we find greater spillover from ELA teachers' instruction, relative to that of math teachers.

Our data and analyses have some key limitations that are typical of research on teacher value-added effects. First, we are unable to examine instructional effects in subject areas other than ELA and math. As a result, we are unable to examine whether the cross-subject generalizability of ELA learning extends to subject areas other than math. Although we hypothesize that this might be the case, future research in contexts where multisubject testing occurs across grade levels is needed to test this. Second, our analysis is limited by the fact that students are not randomly assigned to teachers in our sample. When estimating persistence, our results could be biased if in fact teacher value-added measures include bias from systematic patterns of nonrandom student assignments.

In particular, our findings could have been impacted if ELA teachers' value-added estimates are systematically influenced by nonrandom patterns in the overall quality of instruction their students receive (e.g., through tracking that might afford some students systematically better learning experiences), whereas math teachers' value-added estimates are not influenced to the same degree by this nonrandom sorting. If this were the case, ELA teachers' measured value-added might correspond not just to their own instructional impact, but also to the impact of students' overall instructional experience, whereas math teachers' measured value-added would better reflect just their own instructional impact. That kind of lopsided bias in ELA value-added measures could lead to the larger apparent cross-subject effects of ELA teacher-induced learning that we observe. We are unable to verify whether this may be a factor in our results. That said, the best available evidence on teacher value-added measures in middle-school grades does not indicate any systematic bias in teacher value-added measures of the kind we estimate, even in the absence of random student assignment (Kane, McCaffrey, Miller, & Staiger, 2013). Additional research leveraging data from extant random-assignment studies could help to clarify whether the potential bias described here is present in value-added measures in middle school settings.

The consistency of our findings across two large urban school districts that utilize different standardized tests and content standards provides credible evidence that our findings are likely to generalize across a wide range of school settings where standardized math and ELA exams are the norm. That said, we would expect the persistence of teacher-induced learning to vary as a function of the overlap in assessed content between one year's test administration and the next. Similarly, cross-subject effects may differ in magnitude as a function of the overlap in content between different subject area tests, particularly if some non-ELA tests require student reading and writing skills to a greater degree than others. More work is needed that evaluates competing measures of teachers' instructional impact and that documents key differences in the types of student learning that are captured by different assessments. Such research could inform which types of assessments do a better job of reflecting teachers' full instructional effects.

K–12 teachers have substantial impacts on students' academic achievement, and the best evidence to-date indicates that these impacts are predictive of meaningful long-run effects on students' life outcomes (Chetty et al., 2011). This prior research suggests that value-added measures are useful indicators of teachers' contributions, at least in low-stakes settings. However, particularly in light of the frequent use of value-added measures in school accountability and personnel management systems, it is important to understand the mechanisms by which short-term instructional effects translate into longer-term student outcomes. Without a clearer understanding of those mechanisms, we may be unable to differentiate between short-term effects that are transient or context-dependent, versus those that have broader benefits for students. The findings from this study highlight the potential for meaningful variation in the type of learning that different teachers impart, as evidenced by substantial differences in how different instructional effects play out over time and across subject areas. Additional research is needed to illuminate the processes by which different types of learning yield broad and persistent benefits for students.

Appendix

Value-Added Models

Following Value-Added Research Center (2010), we compute teacher-by-year value-added scores in three stages. In the first stage, we regress posttest Y_t of student i in classroom c with teacher j in school s at time t on their same-subject pretest Y_{t-1}, other-subject pretest Y_{t-1}^{alt}, a vector of student-level time varying and time invariant variables X, and a set of indicator variables representing individual classroom fixed effects π, which can be expressed as

$$Y_{icjst} = \lambda Y_{it-1} + \lambda^{alt} Y_{it-1}^{alt} + \beta X_{it} + \pi_{cjst} + \varepsilon_{icjst}. \qquad (3)$$

Our student-level characteristics include students' gender, race, an indicator for whether the student's home language is English, student eligibility for free or for reduced price lunch, student disability status, English language learner status, an indicator for whether the student switched schools in the prior year, and the number of prior-year absences for the student. Because the effects of characteristics may vary across grade levels, we also include interactions of each student characteristic with each individual grade level. Our goal in this stage is to estimate the coefficients λ for students' pretests and β for student-level characteristics on students' posttest scores.

We estimate the first-stage regression using an errors-in-variables approach (following Fuller, 2009) that accounts for measurement error in pretests Y_{it-1} and Y_{it-1}^{alt}. This removes the variance in the pretests that is attributable to measurement error. To facilitate this approach, we rely on reliability information as reported in the technical manuals for the state assessments in New York and Florida.

In the second stage, we use the estimated coefficients λ and β from our first stage to compute a new left-hand side variable q_{icjst}, where $q_{icjst} = Y_{icjst} - \lambda Y_{it-1} - \lambda^{alt} Y_{it-1}^{alt} - \beta X_{icjst}$. q_{icjst} is, then, the difference between the student's actual score and what we would predict it to be given background characteristics and prior performance. We then regress q_{icjst} on a vector C of classroom-level characteristics, time-varying school-level characteristics K, and individual year and grade dummy indicators:

$$q_{icjst} = \gamma C_{cjst} + \eta K_{st} + \alpha_t + \rho_g + w_{icjst}. \tag{4}$$

Classroom-level characteristics include the racial and home language composition of the classroom, class size, the percent of students who are free or reduced price lunch eligible, percent of students who are English language learners, the class average number of prior year absences, the class average prior year test scores in the same and alternate subject, and the standard deviation of classroom test scores in each subject. As we did for the student covariates, we include interactions of each classroom characteristic with each grade-level indicator. School characteristics include total enrollment, the percent of Black, White, and Hispanic students in the school, and a control for the percent of students eligible for free or reduced price lunch. When running this regression, we specify a classroom random effect to take into account that errors are correlated within classrooms. From this regression, we obtain an estimate of w_{icjst}, that represents the residual test score variation for each student in each year that is not explained by our observable student, classroom, or school characteristics.

In our third stage, we estimate individual teacher value-added measures in each year, τ_{jt}, by attributing all remaining variation in students' posttest scores to a combination of the individual teacher effects and error. This measure can be expressed as:

$$w_{icjst} = \tau_{jt} + \varepsilon_{icjst}. \tag{5}$$

We obtain estimates of the error term ε_{icjst} by subtracting each teacher's mean effect, τ_{jt}, from the estimates of w_{icjst}. Finally, we standardize our teacher-by-year effect estimates across our sample to have a mean of zero and a standard deviation of one.[4] We include in our analysis of persistence only teacher-by-year effects that are based on at least five students and fewer than 100 students.

Specification Check Using an Alternative Value-Added Model

Our primary value-added model specification, as described above, aggregates teacher residuals via a two-stage process. This approach is commonly used both in practice and in the research literature (Chetty et al., 2011; Kane & Staiger, 2008). However, a potential drawback is that our estimates of the standard errors are somewhat biased and our estimates of the effects of classroom characteristics are identified across teachers and thus remove differences across teachers who teach in classrooms with different characteristics. The *felsdvregdm* program produces more accurate standard errors and uses this alternative estimation of classroom-level effects. Although the aggregated residuals model is easier to estimate and more commonly used in practice, the *felsdvregdm* model has been more common in research studies. Thus, as a specification check, we produce alternative value-added measures that include teacher fixed effects directly in our regression according to the following specification:

$$\begin{aligned} Y_{icjst} = {} & \lambda Y_{it-1} + \lambda^{alt} Y_{it-1}^{alt} + \beta X_{it} + \\ & \beta C_{cjst} + \tau_{jst} + \varepsilon_{icjst}. \end{aligned} \tag{6}$$

As in our primary model, we regress posttest Y_t of student i in classroom c with teacher j in school s at time t on their same-subject pretest Y_{t-1}, other-subject pretest Y_{t-1}^{alt}, and a vector of student-level time varying and time invariant variables X. We also include a vector of classroom aggregated characteristics and a set of indicator variables representing individual teacher-by-year fixed effects τ. Our student- and classroom-level covariate controls are the same as those described previously in our primary value-added modeling approach.

We implement this model using the Stata package *felsdvregdm*, which facilitates the inclusion of teacher fixed effects with sum-to-zero constraints (Mihaly, McCaffrey, Lockwood, & Sass, 2010). We estimate individual teacher effects separately by school year, and observe correlations between teacher-by-year value-added measures across the two modeling approaches of approximately 0.87 in both math and ELA. Moreover, when we use value-added measures from this fixed effects model in our analysis of teachers' within- and across-subject persistence in the year following instruction, we observe a very similar pattern of results.[5]

In brief, we find that using this value-added modeling approach yields results that are consistent with our finding of differential cross-subject persistence stemming from ELA instruction. In NYC, ELA teachers' persistent effects on math achievement are approximately 43% of the estimates for math teachers' persistent effects on math achievement. In contrast, math teachers' persistent effects on ELA achievement are much smaller, at approximately 7% of the estimates for ELA teachers' persistent effects on ELA achievement. Similarly, in M-DCPS, our estimates of ELA teachers' persistent effects on math achievement are approximately 32% of the estimates for math teachers' persistent effects on math achievement. In contrast, math teachers' persistent effects on ELA achievement are approximately 8% of the estimates for ELA teachers' persistent effects on ELA. These results indicate that our findings are not an artifact of our preferred approach to estimating teacher value-added effects.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by a grant from the National Center for the Analysis of Longitudinal Data in Education Research (CALDER). CALDER is supported by the Institute of Education Sciences, U.S. Department of Education Grant R305A060018.

Notes

1. There is, however, some variation in estimated persistence across schooling contexts, achievement measures, and methods of estimation. For example, using a mix of experimental and nonexperimental data, Kane and Staiger (2008) estimate persistence of value-added effects that are close to 50% after 1 year, whereas Lockwood, McCaffrey, Mariano, and Setodji (2007) estimate persistence parameters that are less than 20%.

2. Across different value-added models and data sets, researchers have identified correlations ranging from 0.3 to 0.6 between teachers' proximal and future-year effects.

3. Our second-stage equation for estimating 2-year persistence of teacher value-added effects corresponds to our approach for estimating 1-year persistence, and is as follows: $Y_{icjt} = \theta Y_{it-2} + \theta^{alt} Y_{it-2}^{alt} + \beta X_{it} + \pi_{cjt} + \varepsilon_{ijt}$. In this case, the value-added measures T_{ijt-2} and T_{ijt-2}^{alt} for students' twice-lagged teachers in the same and alternate subject areas serve as the two excluded instruments for students' twice-lagged test scores in those subjects.

4. Prior to standardizing, the standard deviation of our (unshrunken) ELA teacher-by-year value-added measures is 0.23, whereas the standard deviation for math teachers' value-added measures is 0.27.

5. In the interest of space, we do not include a full table of our results here, but are happy to provide them upon request.

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.

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, *14*, 219–234. doi:10.1207/S15324818AME1403_2

Buddin, R., & Zamarro, G. (2009). *Teacher effectiveness in urban high schools* (WR-693-IES). Santa Monica, CA: RAND Corporation. Retrieved from http://www.rand.org/content/dam/rand/pubs/working_papers/2009/RAND_WR693.pdf

Candelaria, C. (2016). *The short and long-term effects of teachers on student achievement* (Working paper). Vanderbilt University, Nashville, TN.

Chang, M., Singh, K., & Filer, K. (2009). Language factors associated with achievement grouping in math classrooms: A cross-sectional and longitudinal study. *School Effectiveness and School Improvement*, *20* , 27–45. doi:10.1080/09243450802605704

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood* (No. w17699). Cambridge, MA: National Bureau of Economic Research.

Fuller, W. A. (2009). *Measurement error models* (Vol. 305). New York, NY: John Wiley.

Jackson, C. K. (2012). *Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina* (No. w18624). Cambridge, MA: National Bureau of Economic Research.

Jacob, B. A., Lefgren, L., & Sims, D. P. (2010). The persistence of teacher-induced learning. *Journal of Human Resources*, *45*, 915–943.

Kane, T. J., McCaffrey, D. F., Miller, T., & Staiger, D. O. (2013). *Have we identified effective teachers? Validating measures of effective teaching using random assignment*. Seattle, WA: Bill & Melinda Gates.

Kane, T. J., & Staiger, D. O. (2008). *Estimating teacher impacts on student achievement: An experimental evaluation* (No. w14607). Cambridge, MA: National Bureau of Economic Research.

Kinsler, J. (2012). Beyond levels and growth estimating teacher value-added and its persistence. *Journal of Human Resources*, *47*, 722–753.

Koedel, C. (2009). An empirical analysis of teacher spillover effects in secondary school. *Economics of Education Review*, *28*, 682–692. doi:10.1016/j.econedurev.2009.02.003

Konstantopoulos, S., & Chung, V. (2011). The persistence of teacher effects in elementary grades. *American Educational Research Journal*, *48*, 361–386.

Lockwood, J. R., McCaffrey, D. F., Mariano, L. T., & Setodji, C. (2007). Bayesian methods for scalable multivariate value-added assessment. *Journal of Educational and Behavioral Statistics*, *32*, 125–150.

Loeb, S., & Candelaria, C. (2012). *How stable are value-added estimates across years, subjects, and student groups?* Carnegie Knowledge Network. Retrieved from http://www.carnegieknowledgenetwork.org/briefs/value-added/value-added-stability/

Mariano, L. T., McCaffrey, D. F., & Lockwood, J. R. (2010). A model for teacher effects from longitudinal data without assuming vertical scaling. *Journal of Educational and Behavioral Statistics*, *35*, 253–279.

McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, *29*, 67–101.

Mihaly, K., McCaffrey, D. F., Lockwood, J. R., & Sass, T. R. (2010). Centering and reference groups for estimates of fixed effects: Modifications to felsdvreg. *Stata Journal*, *10*, 82–103.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, *26*, 237–257.

O'Reilly, T., & McNamara, D. (2007). The impact of science knowledge, reading skill, and reading strategy knowledge on more traditional "high-stakes" measures of high school students' science achievement. *American Educational Research Journal*, *44*, 161–196. doi:10.3102/0002831206298171

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*, 417–458.

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

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics*, *125*, 175–214.

Value-Added Research Center. (2010). *NYC teacher data initiative: Technical report on the NYC value-added model*. Madison: Wisconsin Center for Education Research. Retrieved from http://schoolsstg.nycenet.edu/NR/rdonlyres/A62750A4-B5F5-43C7-B9A3-F2B55CDF8949/87046/TDINYCTechnicalReportFinal072010.pdf

Yuan, K. (2015). A value-added study of teacher spillover effects across four core subjects in middle schools. *Education Policy Analysis Archives*, *23*. doi:10.3102/0162373716665698

Authors

BENJAMIN MASTER is an associate policy researcher at the RAND Corporation. His research is in the area of educational policy with a particular focus on school leadership and human capital management in schools. His work explores promising investments in human capital development, including interventions that support teacher and principal effectiveness and the impacts of educational interventions on students' lives. Address: 1200 South Hayes Street, Arlington, VA 22202. Email: bmaster@rand.org

SUSANNA LOEB is the Barnett Family Professor of Education at Stanford University. She specializes in education policy, looking particularly at policies and practices that support teachers and school leaders. Her work spans the range of age levels, including early education, K–12, and higher education. Her recent work focuses particularly on information barriers to teaching improvement and parenting. She is a member of the National Board for Education Sciences, codirector of Policy Analysis for California Education, a senior fellow at the Stanford Institute for Economic Policy Research, and a faculty research fellow at the National Bureau of Economic Research. Address: 520 Galvez Mall #524, CERAS building, Stanford, CA 94305-3084. Email: sloeb@stanford.edu

JAMES WYCKOFF is the Curry Memorial Professor of Education and Policy, and director of the EdPolicyWorks, a research center at the University of Virginia. His research focuses on providing evidence for improving teacher quality through a variety of mechanisms including teacher preparation, recruitment, assessment, and retention. Currently he is working with colleagues to examine the effects of teacher evaluation in the District of Columbia and the ways in which principals improve teacher quality in New York City. Address: PO Box 400277, Charlottesville, VA 22904. Email: Jhw4n@virginia.edu

Manuscript received August 24, 2015
First revision received July 20, 2016
Second revision received November 18, 2016
Accepted January 2, 2017