

Teacher Effectiveness and Pathways into Teaching in California

**Xiaoxia Newton,
University of California, Berkeley
April, 2010**

This paper reports results of teacher effectiveness estimates based on teachers' pathways into teaching and the teacher education programs they engaged in. The work reported here is embedded in a longitudinal study, which examines the relationships among teacher education, teaching practices, and pupil learning as part of the Teachers for a New Era (TNE) reform initiative at Stanford University.

Methodology

Study Context and Data Source

As part of the Teachers for a New Era (TNE) reform initiative at Stanford, a sample of approximately 250 secondary teachers of mathematics, science, history/social studies, and English language arts and roughly 3500 students taught by these teachers was examined. All were from a set of six high schools in the San Francisco Bay Area. Because California did not at this time have a state longitudinal data system, student and teacher data had to be secured from individual schools and districts' electronic data files.

Conceptualization of Teacher Effectiveness

The measurement of "value added" gains in achievement was based on the variation in pupils' test scores on the California Standards Tests (CSTs), controlling for prior-year scores, rather than on variation in year-to-year test score gains, because the CSTs are not vertically scaled and, therefore, do not yield interpretable gain scores. Although the CSTs do use IRT scaling to create scale scores, these scores are not vertically equated in California.

The study used ordinary least square (OLS) regression analyses to predict pupils' CSTs after taking into consideration prior year's achievement (CST scores in the same subject area) and key demographic background variables (i.e., race/ethnicity, gender, free/reduced lunch status, English language learner status, and parent education). The study also controlled for school fixed effects to take into account the unobserved differences among schools that may influence teachers' measured effectiveness (e.g. school leadership, resources, parental involvement). With these statistical controls, teacher's effectiveness is then measured by the average difference between actual scores and predicted scores for all students assigned to that teacher (i.e., the mean residual).

Linking Students to Teachers

For each of the schools in the study sample, we obtained student course enrollment files. Based on these course files, we linked individual pupils with teachers

from whom they took the English language arts or mathematics courses for both fall and spring semesters (i.e., during the entire academic year). Additionally, when a teacher was teaching different courses to different groups of students (e.g., algebra 1 and geometry; or regular English and honors English), we generated separate value-added estimates for the teacher (one for each course).

Because California students take different high school courses, each with its own end-of-course examination (e.g., algebra 1, geometry, algebra 2, etc.), and CST scale scores are not directly comparable across different course-specific tests, scale scores from each CST were converted to z scores prior to performing these regressions. We transformed raw scale scores into z-scores based on the sample mean and standard deviation of a particular grade (for English language arts, where students take grade-level tests each year) or of a particular subject test (for math, where students take subject-specific tests). In addition to enabling the pooling of prior-year scores across different CSTs, this linear transformation of raw scale scores also facilitated the presentation of study outcomes in a standardized metric.

Sample and Data

Tables 1 and 2 describe the sample and the types of data that formed the basis of the value added analysis for the current study.

Table 1: Teacher and Student Samples for the VAM Analyses

Sample	2005-06	2006-07
Mathematics teachers	57	46
ELA teachers	51	63
Science teachers	33	29
Students		
Grade 9	646	881
Grade 10	714	693
Grade 11	511	789

Note: Some teachers taught multiple courses. There were 13 such math teachers for year 2005-06 and 10 for year 2006-07; and there were 16 such ELA teachers in 2005-06 and 15 in 2006-07. These distinctive courses were counted as separate teacher records for the purpose of these analyses.

Table 2: List of Variables

Variables	Scale
<i>Outcome measures:</i> CST math or ELA	CST scale scores were transformed to Z scores
<i>Student prior achievement:</i> CST math or ELA in previous year On track status (for math)	CST scale scores were transformed to Z scores. Variable indicating that a student took a math course at the usual grade level it is offered in

Fast track status (for math)	the school Variable indicating that a student took a math course at an earlier grade level than it is usually offered in the school
<i>Student demographic background:</i> Race or ethnicity	Indicator variables for African American, Hispanic, Native American, Pacific Islander, or Asian
Gender	Indicator variable for female
English language learners	Indicator variable for English language learner
<i>Student social economic status proxies:</i> Parent educational level	Indicator variable for high school or above Ordinal measure (0-4) from less than high school to education beyond college
Meal program	Indicator variable for free or reduced lunch meal participation
<i>School differences:</i>	Dummy indicator variable for each school

In addition to student test scores and background variables, we also surveyed all teachers teaching core subject areas in our sampled schools. The survey gave further information about their preparation, professional learning opportunities, teaching context, and self-reported practices. For this study, teacher pathway and preparation program information were used.

Data Analysis

We conducted a series of parallel ordinary least square (OLS) linear regressions with school fixed effects, separately for math, for ELA, and for science, and for years 2006 and 2007, respectively. These OLS analyses generated residual (observed minus predicted) scores for each student. These residual scores for each student were aggregated to the teacher (or course within teacher) level. These aggregated residual gain scores serve as teacher effectiveness estimates. For this report, we combined the 2006 and 2007 teacher effectiveness estimates so as to maximize the sample size (i.e., observed effectiveness estimates) when comparing different programs and pathways. We then examined the average teacher effectiveness estimates along the following dimensions: (1) STEP graduates vs. non-STEP graduates by years of teaching (8 years or longer or fewer than 8 years); and (2) STEP graduates vs. other pathways to teaching.

Results

Figure 1 displays the average teacher effectiveness estimates of STEP graduates versus others by years of teaching experience (i.e., more than 8 vs. 8 or fewer). As shown in Figure 1, the average teacher effectiveness estimates for STEP graduates who had taught more than eight years were about .30 standard deviations above the mean

effectiveness of other similarly experienced teachers. The mean effectiveness estimates for STEP graduates who had taught eight or fewer years were about .14 standard deviations above those of non-STEP graduates with similar levels of experience. Whereas STEP graduates appear to experience returns to experience beyond 8 years (that is, greater effectiveness with more years of experience), the reverse was true for non-STEP teachers, for whom the less experienced cohort appeared more effective than the highly experienced group. This may be a function of improvements in the preparation of teachers generally over recent years, which has been a goal of state policy.

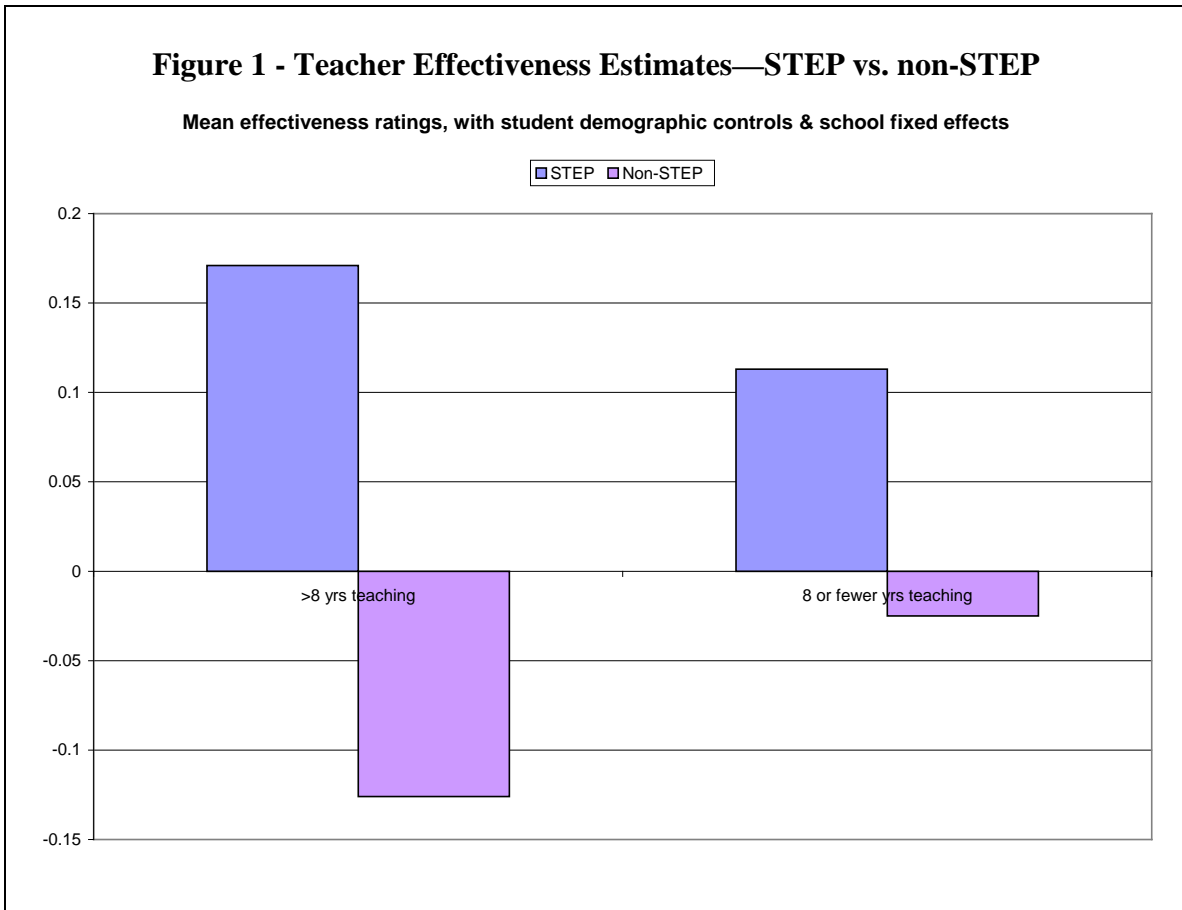


Figure 2 displays the average teacher effectiveness estimates for alumni from different teacher education programs and pathways. As shown in Figure 2, graduates from STEP produced higher value-added achievement gains for their students than those of the other teacher education program groups and teachers from intern / alternative programs. Specifically, STEP graduates as a group had an average teacher effectiveness estimate that was a little over .08 standard deviations above the mean effectiveness. The effectiveness estimates for other teacher education programs and pathways ranged from .05 standard deviations above the mean effectiveness score to about .09 standard deviations below.

To summarize, the descriptive analysis showed that graduates from different teacher education programs and pathways exhibited different effectiveness in students’

learning outcomes on the standardized tests. These results should be interpreted with caution, since our sample size is moderate and the sample sizes for individual programs were of varying sizes (ranging from 12 to 79). In addition, these results should be interpreted as descriptive, not causal (i.e., as signals of effectiveness). Further analysis of larger samples is needed. In-depth case studies of purposively selected group of teachers who exhibited different effectiveness estimates – another aspect of the TNE study -- will also help to triangulate the quantitative analysis and illuminate relationships between teachers’ preparation, their practices, and their students’ outcomes. .

Figure 2 - Estimates of High School Student Value-Added Achievement for Graduates of Teacher Education Programs / Pathways

