Are Value-Added Estimates Unbiased? An Assessment Using Multiple Districts

By Elias Walsh, Philip Gleason, Eric Isenberg, and Jeffrey Max

Background: Value-added estimates of teacher effectiveness have been used by school districts for information, bonus awards, and even retention decisions, so it is essential to know whether these measures are accurate. Chetty, Friedman, and Rockoff (2014) used a novel quasi-experimental approach based on teacher transitions to measure the degree of bias in value-added estimates, and tested the approach using data from a single large urban district. They found no statistically significant bias in value-added estimates. Similarly, Bacher-Hicks, Kane, and Staiger (2014) replicated these findings using data from Los Angeles Unified School District, as did Rothstein (2015) using data from North Carolina. We extend previous applications of the method by applying it to data in a broad-based sample of 23 districts from 13 states and examining several value-added models.

Research Questions: This study, sponsored by the Institute of Education Sciences (IES) at the U.S. Department of Education, examines whether value-added estimates are unbiased. In particular, we ask

1. To what degree are value-added estimates unbiased?

2. Are there differences in the level of bias based on different value-added models?

3. Do models that exhibit greater bias systematically penalize teachers of disadvantaged students?

Setting: We examined bias in value added in a geographically dispersed sample of school districts over a five-year period (2008–2009 to 2012–2013). While past studies in this literature have focused on teachers in a single district or single state, we use data on teachers in 23 districts located in 13 states in all four Census regions.

Population: We examined teachers of English/language arts and math in 23 study districts. Due to data limitations, we examined grades six to eight in 14 districts and grades four to eight in 9 districts.

Program: Our goal is to assess the degree of bias in value added, which might be used for teacher evaluation purposes, rather than to evaluate a particular program.

Research Design: We base our analysis on the quasi-experimental approach developed by Chetty, Friedman, and Rockoff (2014) to test for bias in value-added estimates. The approach examines changes in the average test scores of students in a given school (and grade and subject) when there is a change in the average value added of teachers in the school. When a teacher with low value added leaves a school and is replaced the next year by a teacher with high value added, student test scores in the school should increase if the value-added estimates accurately reflect the teachers’ effectiveness. If average student scores increase by the amount predicted (or “forecast”) by value added, the value-added estimates are said to have no forecast bias.
Following Chetty, Friedman, and Rockoff (2014), we use the following steps to conduct this test for bias. First, we estimate each teacher’s value added in each year of the study using data on their students’ test scores in that year as outcomes, and prior year test scores and other background characteristics of students and their peers as control variables. Next, for each two-year period during the study, we predict each teacher’s value added in those two years based on the teacher’s value-added estimates from outside of the period. We use these “leave-out” value-added estimates, rather than the single-year estimates, to avoid basing the value-added estimates on the same student test scores used to measure changes in student achievement within schools. Third, within each school-grade-year-subject combination, we calculate averages of the leave-out value-added estimates $\bar{VA}_{jgys}$ and test scores $\bar{y}_{jgys}$. Fourth, we calculate year-to-year differences in the averages between each pair of years to obtain $\Delta \bar{VA}_{jgys}$ and $\Delta \bar{y}_{jgys}$. Finally, we estimate the following regression: 

$$\Delta \bar{y}_{jgys} = \alpha + \beta \Delta \bar{VA}_{jgys} + e_{jgys}.$$ 

When $\hat{\beta} = 1$ the value-added estimates are unbiased, so $1 - \hat{\beta}$ provides the measure of forecast bias.

In some specifications, we additionally account for the possibility that teachers might transition in or out of schools based on changes in the characteristics of students in those schools or based on other factors that might cause changes in student test scores, such as new principals. We use two approaches to do this. First, we account for the characteristics of students by subtracting students’ predicted scores based on the coefficients on prior test scores and background characteristics in the value-added model from the test scores used to estimate the school averages. Second, we include fixed effects for school-year-subject combinations to account for school-wide shocks (such as new principals) that might affect certain cohorts.

**Data Collection and Analysis:** We collected the data that would be used in a typical value-added model, including standardized student test scores from state assessments in grades three to eight, a set of student characteristics, school enrollment data for students, and teacher-student-course links indicating the teacher responsible for teaching English/language arts and/or math to each student.

We compare bias in value-added models that include or exclude some common features such as accounting for measures of students’ peers in the same classroom or addressing measurement error in pre-test scores. We then examine how using a value-added model identified as more biased affects teachers of certain students, such as disadvantaged students. Finally, we describe some pitfalls to avoid when applying the method to small or moderate sized districts or short panels, and contribute to a discussion about potential threats to the validity of the bias estimates.

**Results:** The report will not be public until it has completed the entire IES review process. We expect that the report will be released in December 2016, and will share the results publicly once that has occurred.

**Conclusions:** Bias in value-added estimates is an important topic, and the results of this IES-sponsored report, to be made public soon, will inform policymakers about this issue.
References

