

THE LONG-TERM IMPACTS OF TEACHERS: TEACHER VALUE-ADDED AND STUDENT OUTCOMES IN ADULTHOOD

Raj Chetty, Harvard University and NBER
John N. Friedman, Harvard University and NBER
Jonah E. Rockoff, Columbia University and NBER

Many policy makers advocate increasing the quality of teaching, but there is considerable debate about the best way to measure and improve teacher quality. One method is to evaluate teachers based on their impacts on students' test scores, commonly termed the "value-added" (VA) approach. A teacher's value-added is defined as the average test-score gain for his or her students, adjusted for differences across classrooms in student characteristics such as prior scores. School districts from Washington D.C. to Los Angeles have begun to use VA to evaluate teachers. Proponents argue that using VA can improve student achievement (e.g. Hanushek 2009), while critics argue that test score gains are poor proxies for a teacher's true quality (e.g. Baker et al. 2010).

The debate about VA stems primarily from two fundamental questions. First, does VA accurately measure teachers' impacts on scores or does it unfairly penalize teachers who may systematically be assigned lower achieving students? Second, do high VA teachers improve their students' long-term outcomes or are they simply better at teaching to the test? Researchers have not reached a consensus about the accuracy and long-term impacts of VA because of data and methodological limitations.

In Chetty, Friedman, and Rockoff (2014a,b), we address these questions by tracking one million children from an urban school district from 4th grade to adulthood. We evaluate the accuracy of standard VA measures using several methods, including natural experiments that arise from changes in teaching staff. We find that when a high VA teacher joins a school, test scores rise immediately in the grade taught by that teacher; when a high VA teacher leaves, test scores fall. Test scores change only in the subject taught by that teacher, and the size of the change in scores matches what we predict based on the teacher's VA. These results establish that VA accurately captures teachers' impacts on students' academic achievement and thereby reconcile the conflicting conclusions of Kane and Staiger (2008) and Rothstein (2010). These methods provide a simple yet powerful method to estimate the bias of value-added models in any district; interested readers can download computer code to implement these tests from this [link](#).

In the second part of our study, we analyze whether high VA teachers also improve students' long-term outcomes. We find that students assigned to higher VA teachers are more successful in many dimensions. They are more likely to attend college, earn higher salaries, live in better neighborhoods, and save more for retirement. They are also less likely to have children as teenagers. Teachers have large impacts in all the grades we analyze (4 to 8). Teachers' impacts on earnings are also similar in percentage terms for students from low and high income families.

Teachers' impacts on students are substantial in monetary terms. Replacing a teacher whose true VA is in the bottom 5% with one of average quality would generate cumulative earnings gains of \$50,000 per student or more than \$1.4 million for the average classroom; discounting at a 5% interest rate to age 12 yields a present value gain of more than \$250,000 per classroom.

Critics have correctly expressed concern that VA estimates based on data from few classes may be unreliable. However, we find that VA estimates based on even a few years of data are reliable enough that personnel changes would yield large gains. For example, replacing a teacher whose estimated VA based on three years of data is in the bottom 5% with one of average quality would generate lifetime earnings gains for students exceeding \$1 million per class. This is because very few of the teachers with low VA estimates ultimately turn out to be excellent teachers: we estimate that only 3% of teachers rated in the bottom 5% based on three years of data are actually above-average in terms of quality.

Overall, our study shows that great teachers create great value and that test score impacts are helpful in identifying such teachers. However, more work is needed to determine the best way to use VA for policy. For example, using VA in teacher evaluations could induce counterproductive responses that make VA a poorer measure of teacher quality, such as teaching to the test or cheating. There will be much to learn about these issues from school districts that start using VA to evaluate teachers. Nevertheless, it is clear that improving the quality of teaching – whether using value-added or other tools – is likely to have large economic and social returns.

Works Cited

Baker, Eva L., Paul E. Barton, Linda Darling-Hammond, Edward Haertel, Helen F. Ladd, Robert L. Linn, Diane Ravitch, Richard Rothstein, Richard J. Shavelson, and Lorrie A. Shepard. 2010. "Problems with the Use of Student Test Scores to Evaluate Teachers." Economic Policy Institute Briefing Paper #278.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014a. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." Forthcoming, *American Economic Review*.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014b. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." Forthcoming, *American Economic Review*.

Hanushek, Eric A. 2009. "Teacher Deselection." in *Creating a New Teaching Profession*, ed. Dan Goldhaber and Jane Hannaway, 165--80. Washington, DC: Urban Institute Press.

Kane, Thomas J., and Douglas O. Staiger. 2008. "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation," NBER Working Paper No. 14607.

Rothstein, Jesse. 2010. "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement," *Quarterly Journal of Economics* 125(1): 175-214.