Identifying Effective Teachers in the United States

Researchers:

Thomas Kane

Daniel McCaffrey

Trey Miller

Douglas Staiger

Sector(s): Education

Sample: 1,181 teachers

Target group: Students Teachers

Outcome of interest: Student learning Service provider performance

AEA RCT registration number: AEARCTR-0001487

Partner organization(s): Gates Foundation, Charlotte-Mecklenburg Schools, Dallas Independent School District, Denver Public Schools, Hillsborough County Public Schools, Measures of Effective Teaching Project (MET), RAND Corporation

Identifying which teachers are effective can help districts and schools improve teacher quality. Researchers evaluated how well different measures of teacher effectiveness predicted student performance in six school districts in the United States. Measures of teacher effectiveness based on student achievement in the previous year, classroom observations, and student surveys accurately identified which teachers produced higher average student test scores.

Policy issue

If school systems want to develop and retain highly effective teachers, they must be able to identify which teachers are most effective at helping students succeed academically. Classroom observations and student surveys are two means available to evaluate teachers. Additionally, “value-added” models assign teachers a score based on the achievement gains made by students in a class, while adjusting for student differences such as past test scores and demographic characteristics. All three methods could potentially identify higher effective teachers and predict student achievement in future years, but it is difficult to quantify the impact that teachers have on student achievement because many other factors can also affect grades and test scores. Some differences among students, such as level of motivation, are hard, if not impossible, to measure. If some teachers have classes with more exceptional students, they could appear to be more effective than they actually are. For this reason, there exists little rigorous evidence about how accurately different measures of teacher effectiveness actually predict student performance.

Context of the evaluation

The Measures of Effective Teaching (MET) project was a research partnership of academics, teachers, and education organizations funded by the Bill & Melinda Gates Foundation with the goal of finding better ways to identify strong teachers and improve the quality of teaching. The project compared several methods to measure teacher effectiveness, including classroom observations, student perception surveys, and student achievement gains.

As part of the MET project, researchers partnered with participating teachers and schools in six school districts in the United States: Charlotte-Mecklenburg Schools (NC); Dallas Independent Schools (TX); Denver Public Schools (CO); Hillsborough County Public Schools (FL); Memphis Public Schools (TN); and New York City Schools (NY). Within those districts, researchers studied teachers who taught math or English in grades four through eight.



A teacher engages a young student in a literacy activity in the United States.

Photo credit: Wirestock Creators, Shutterstock.com

Details of the intervention

Researchers evaluated measures of teacher effectiveness to see if they accurately predicted student achievement. Using data from 2009-2010, researchers calculated “value-added” scores for 1,181 teachers based on the gains that students made on state achievement tests. Researchers combined the value-added score with results from student surveys and classroom observations and controls for teacher experience and education to create a composite measure of teacher effectiveness.

During the following school year, researchers tested how well value-added scores, classroom observations, student surveys, and the composite scores predicted student test scores at the classroom level. Participating schools created rosters which listed students who would be in the same classroom but did not assign teachers. When there were two or more participating teachers in the same school, grade, and subject, researchers randomly assigned rosters to teachers. Researchers used each measure of teacher effectiveness to predict differences in average student test scores between classrooms. They compared the predicted scores to classes’ actual average test scores. Because classes assigned to more or less effective teachers, as predicted using the 2009-2010 data, were the same – on average – before randomization, the differences in scores at the end of 2010-2011 estimated the true effect of teachers on test scores.

Results and policy lessons

The measures of teacher effectiveness calculated using data from the 2009-2010 school year accurately identified stronger teachers.

Many students did not remain with the teacher that they were assigned because students transferred to other classes or schools, teachers changed classes, or principals decided not to honor the randomized assignments. To account for this, researchers used an instrumental variables strategy with random assignment used as the instrument. Researchers measured the impact of differences in assigned teachers' measured effectiveness and scaled up these estimates to infer the impact of actual teachers' measured effectiveness. Because, on average, a one-unit difference in the assigned teacher's measured effectiveness led to a 0.293 unit difference in the actual teacher's measured effectiveness, the impact of actually having a more effective teacher is found by multiplying the impact of being assigned a more effective teacher by about 3.4 (1/0.293).

Students assigned to teachers identified as more effective performed better on state tests in the 2010-2011 school year, and those assigned to less effective teachers performed worse. Additionally, the size of the effect of having a more effective teacher was consistent with what the composite model predicted: a one-unit increase in composite predicted teacher effectiveness corresponded to an average increase of 0.96 units in actual student achievement on state tests.

Students in classes with teachers identified as more effective also performed better on supplemental tests in reading and math. These tests included more cognitively challenging questions that required writing, analysis, and application of concepts. A one-unit change in predicted effectiveness corresponded to an average change of 0.66 units on these supplemental tests. While the results show that these measures can accurately predict differences between teachers within a school, further research is needed to determine if such measures can be used to compare teachers across different schools.

Kane, Thomas J., Daniel F. McCaffrey, Trey Miller, Douglas O Staiger. 2013. "Have we identified effective teachers? Validating Measures of Effective Teaching Using Random Assignment" MET Project Research Paper - Bill & Melinda Gates Foundation.