EVAAS Value-Added Scores for Teachers with High Need Students
By Dara Shifrer, PhD

BACKGROUND

Policymakers tend to emphasize differences in the teachers rather than differences in the students at low-performing schools, with ‘low-performing’ typically defined on the basis of average test scores. Because many teacher characteristics (e.g., certification status, educational attainment) do not clearly associate with student achievement, there is an increased reliance on differences in student achievement gains themselves as an indicator of differences in teacher quality. Houston Independent School District (HISD) contracts with EVAAS (Education Value-Added Assessment System) to obtain individual value-added scores for all of its core teachers of grades 3-8 (core referencing teachers who teach Language Arts, Reading, Math, Science and/or Social Studies). EVAAS, the most widely implemented value-added methodology across the nation, ostensibly isolates the effect of the teacher from all other factors that influence student achievement gains. Critics argue value-added methodology cannot sufficiently account for the nonrandom sorting of teachers across schools and students. EVAAS does not provide transparent or specific descriptions of their methodology, and there has been very little external statistical evaluation of EVAAS scores.

An earlier study from the Houston Education Research Consortium (HERC) showed teachers in HISD’s highest need schools received lower value-added scores on average than teachers in HISD’s lower need schools. It was unclear whether this was an indication that HISD’s highest need schools staffed lower quality teachers or whether teachers at the highest need schools were being penalized because of issues with EVAAS’ methodology. This study used longitudinal data on teachers in the HISD between 2007-08 and 2012-13, as well as data on their students and schools, to investigate the degree to which value-added scores capture the impact of teachers on student achievement gains rather than the influence of factors outside of teachers’ control.

RESEARCH QUESTIONS

1. Is variation in teachers’ value-added scores better explained by the qualities of teachers or by the qualities of teachers’ students?
2. Are the value-added scores of teachers who teach at both low- and high-performing schools consistent over time?
RESULTS

Baseline findings show correlations between teachers’ value-added scores in different subjects are low, but it is possible teachers are more efficacious in certain subjects. There are substantial discrepancies in teachers’ effectiveness ratings depending on whether they’re rated on the basis of their value-added scores or on the basis of their administrators’ observations of them; this could indicate issues with observational methods rather than EVAAS methodology.

Descriptive findings lay an important foundation. Ranking schools on the basis of average test scores, the teachers in HISD’s lowest performing schools have much lower average value-added scores than the teachers in HISD’s highest performing schools. It is important to note that value-added methodology focuses on achievement gains rather than baseline achievement levels, with the expectation that students with lower average test scores can and should experience gains comparable to those experienced by students with higher average test scores at the baseline. Descriptive differences also showed comparable proportions of teachers in HISD’s lowest performing schools and in HISD’s highest performing schools attained a Master’s degree or higher, completed a degree in a core subject, and were certified in a core subject. There were more inexperienced teachers in HISD’s lowest performing schools than in HISD’s highest performing schools, with respectively 30% and 20% of teachers with 3 years of experience or less. In contrast to differences in teachers, the differences in the students in HISD’s highest and lowest performing schools were marked. On average, 50% of the students in the lowest performing schools are living in poverty in contrast to only 25% of the students in the highest performing schools. Whereas only 5% of students in the lowest performing schools are classified as not economically disadvantaged, the same proportion is 40% in the highest performing schools. Students in the lowest performing schools are much more likely to be black rather than white, indicating social processes related to segregation and isolation that impact school climate and achievement.

Turning to the first research question, analyses focused on a single year used multiple techniques to show variation in teachers’ value-added scores within the same school was explained to a larger degree by the qualities of the students they taught than by their own qualities. More troubling, fixed effects models that tracked teachers over time showed teachers’ average value-added scores were partially a function of the performance level of their schools, with the same teacher’s value-added scores increasing with a switch to a higher performing school and decreasing with a move to a lower performing school.

DISCUSSION AND RECOMMENDATIONS

The accuracy and consistency of these scores is important for the teaching profession and for the correct direction of policy reform aimed at educational disparities. A large body of academic research argues there simply is not sufficient data or statistical techniques to use value-added scores as a central means of determining teacher quality (if at all). In HISD in 2014-15, teachers’ final evaluations coupled results from value-added scores and administrator observations. Value-added scores in other districts comprise an even smaller part of teacher evaluations, with some districts even incorporating student evaluations. HISD might also consider shifting some of the time and money invested into value-added scores to increased support and mentoring of current teachers, particularly as it is unclear whether there are sufficient higher quality teachers available to replace district teachers deemed low quality.

Conclusion

HERC’s first study focused on value-added scores and the ASPIRE program found teachers generally responded positively to receiving financial awards but that teachers in HISD’s highest need schools were less likely to receive an award (i.e., had lower value-added scores on average). This study built on these findings to examine the origins of differences in value-added scores, specifically investigating whether teachers’ value-added scores varied as a function of their students or schools. Results from this study indicate factors outside teachers’ control, such as student background and peer effects, partially influence teachers’ EVAAS value-added scores. This is troubling given the fact that these scores can impact teachers’ professional status, salaries, and job stability. Moreover, with evidence that differences in achievement are much more closely related to differences across homes than differences across schools and teachers, policy
interventions like these may reinforce misunderstandings as to the origins of educational disparities. They also jeopardize the recruitment and retention of quality teachers to schools who need them most.

This is an abbreviated version of a much longer research study written for peer review. For additional information on the findings presented here, or to obtain the full peer-review version of this research brief, contact the Houston Education Research Consortium at 713-348-2802 or email herc@rice.edu.