

Use of value added in teacher policy measures

Eric A. Hanushek

Eric A. Hanushek is Paul and Jean Hanna Senior Fellow at the Hoover Institution of Stanford University, and an IRP affiliate.

I would like to offer a different take on the policy issues related to value-added estimates than that provided in Jesse Rothstein's article. I believe that the primary value of these estimates is in illustrating how much difference there is between teachers. When the estimates are made in low-stakes situations where there is little incentive to teach to the test, estimates of the variance in teacher quality are very precise. In this article, I discuss the implications of the results of these types of studies, and then explore the implications for teacher policy. I believe that where Rothstein's argument falters is that there are not currently any school systems that make teacher personnel decisions solely on the basis of value-added estimates, nor am I aware of any current proposals for such a system. For example, in regard to the District of Columbia policy described by Raj Chetty, only 18 percent of teachers in the system have value-added scores available, so this information is clearly only a relatively small part of what goes into making firing and bonus decisions.

Magnitudes of effects

Estimates of the average standard deviation in gains in student achievement over one year attributable to higher value-added teachers within a given school range from 0.13 to 0.17. Any between-school differences in teacher effectiveness would need to be added on top of this. Although Chetty has already discussed some of the implications of these differences, I will very briefly offer my own calculations.¹ Estimates of the effect of test scores on earnings indicate that a standard-deviation increase in scores translates into a 13 to 20 percent annual increase in earnings. Figure 1 illustrates the effect on student lifetime income by class size and teacher effectiveness, allowing for some depreciation in scores over time. This figure shows the estimated marginal effect, compared to an average teacher, of having a teacher in various percentiles. Calculations for individual students are multiplied by class size. So, for example, the present value at the beginning of high school for a 75th percentile teacher with a class of 30 students is \$430,000, while that for a 25th percentile teacher with the same class size is \$425,000. These numbers appear large enough to suggest that, although there may be some error in particular teacher personnel policies, having no personnel policy at all cannot be the correct answer.

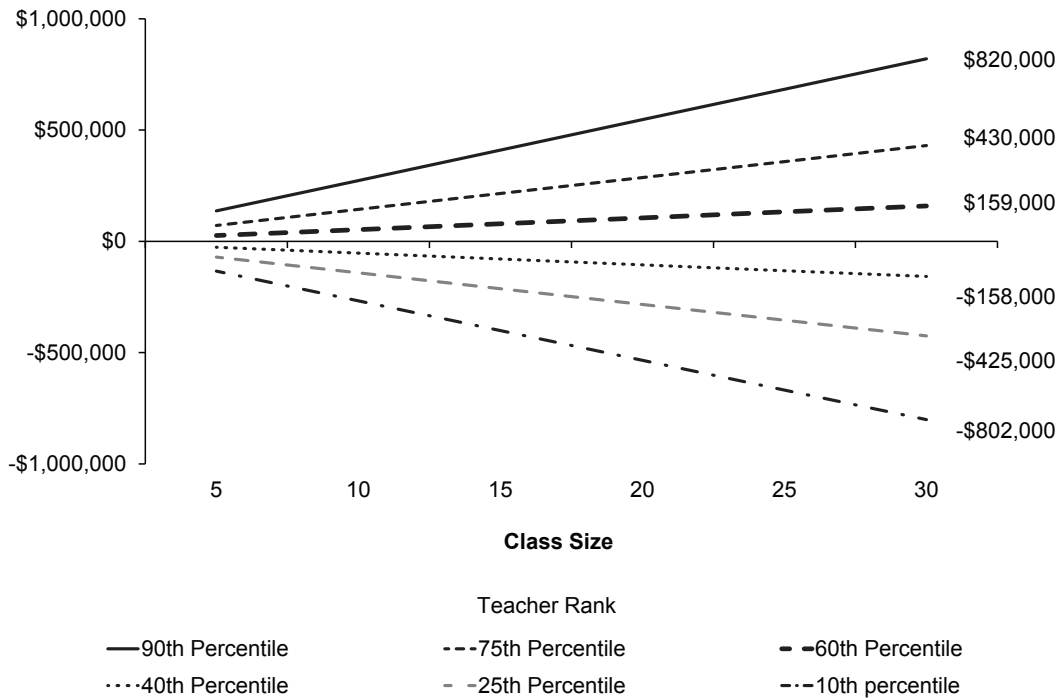


Figure 1. Effect on student lifetime incomes by class size and teacher effectiveness (compared to average teacher).

Source: Calculations by author relying on estimates of teacher quality using 0.2 standard deviations, and reflecting between-school calculations.

School districts have needed to lay off teachers in substantial numbers only quite recently, as a result of the recent recession. The standard policy for determining layoffs is to use teacher seniority. A recent simulation comparing this policy to one that used a measure of effectiveness found some differences between the two approaches.² Since seniority-based layoffs generally mean that those with lower salaries are more likely to lose their jobs, more layoffs are required to achieve a given budget reduction. In this simulation, a system based on value-added results in about 25 percent fewer layoffs than one based on seniority. In addition, the typical teacher laid off using a value-added system is less effective than the typical seniority-based layoff, by 26 percent of a standard deviation.

Another mental exercise is to imagine ranking all teachers in the United States based on effectiveness, and look at the performance gains that would result from deselecting some percentage of the lowest-ranked teachers, and replacing them with an average teacher. In this case, unlike the one-year effects that Rothstein estimated, I am looking at lifetime effects. I find that, depending on whether a high or low estimate of teacher effectiveness is used, a deselection rate of between 5 and 8 percent would result in achievement levels similar to that of Canada, a country that currently ranks 0.42 standard deviations above the United States. According to calculations I have made along with Ludger Woessman, such an increase in achievement is worth \$72 trillion in GDP.³ Larger estimates of the variation in teacher effectiveness result in even higher estimates. Although the precise value can certainly be argued, it is clear to me that the value of having policies based on teacher effectiveness is enormously higher than having no policy at all, and that policies based on teacher effectiveness in fact represent the future of the U.S. economy.

Use of value-added measures in teacher personnel policy

There has been a great deal of discussion about errors in estimating value added, and whether it is acceptable to, for example, have a 5 percent error rate in determining which teachers contribute the most to student achievement. I believe that the current state of having no policy translates to a 100 percent error rate, and that we should be striving not for perfection, but for a policy that improves teacher effectiveness overall.

Rothstein discussed some of the implications of making teacher-retention decisions based on imperfect value-added scores. If the rate of dismissal and replacement is somewhere between 5 and 8 percent, that translates to 2 to 3 individuals in a school of 30 teachers. I have found in all of my dealings with teachers, administrators, parents, and staff in numerous schools, that there is very little uncertainty about who the 2 to 3 least-effective teachers in any given school are. I believe that an evaluation process that allowed decisions based on this type of common knowledge would not necessarily need to depend on value-added data that might not be available in a timely manner, and that the evidence

suggests that such a policy would likely result in substantial gains in student achievement.

As has been mentioned, both the *Los Angeles Times* and the *New York Times* have recently published teacher value-added scores for their respective school districts. This was extremely controversial, and the aftershocks are still being felt. I was one of the few researchers to support the idea of publishing value-added scores, not because I think that personnel policy should be done through newspapers, but because within a week of these publications, unions and school officials were meeting to discuss teacher-evaluation policy. This is an issue that had been on the agenda forever with no progress. It seems that providing a strict value-added ranking as one (extreme) option prompts people to develop better personnel systems that incorporate other teacher-evaluation tools, and this is exactly what is needed.

Issues and areas for further study

One could ask whether the currently available achievement tests are really up to the task of providing reliable value-added scores. I would say certainly not, and that value-added measures should never be the sole basis for personnel decisions. Rothstein also raised the possibility that value-added measures can become less reliable when used for consequential purposes. While this and the accompanying loss in reliability and validity is certainly possible, I believe such problems can be dealt with in feasible ways.

On the question of whether value-added measure can be used to rate principals, I agree with Rothstein that a parallel system is required. There are some indications that reliable value-added measures can be constructed. Preliminary estimates from work that I have been involved in suggest that principal quality is extremely important and that a one standard deviation increase in principal quality results in an increase of approximately 0.05 standard deviations in average student growth.⁴ While this effect is much smaller than that seen for teachers within a given school, principals affect all students in a school, so an increase in principal quality will have effects much greater than a similar increase in the quality of a single teacher. ■

¹E. A. Hanushek, "The Economic Value of Higher Teacher Quality," *Economics of Education Review*, 30, No. 3 (June 2011): 466-479.

²D. Boyd, H. Lankford, S. Loeb, and J. Wyckoff, "Teacher Layoffs: An Empirical Illustration of Seniority versus Measures of Effectiveness," *Education Finance and Policy*, 6, No. 3 (Summer 2011): 439-454. The simulation was conducted using fourth- and fifth-grade math and language arts achievement scores for students in New York City public schools.

³E. A. Hanushek and L. Woessmann, *The high cost of low educational performance: The long-run economic impact of improving PISA outcomes* (Paris: Organisation for Economic Cooperation and Development, 2010).

⁴G. F. Branch, E. A. Hanushek, and S. G. Rivkin, "Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals," NBER Working Paper No. 17803, National Bureau of Economic Research, 2012.