



**Carnegie Foundation** for the Advancement of Teaching

# What Do We Know About the Use of Value-Added Measures for Principal Evaluation?

**SUSANNA LOEB**

STANFORD UNIVERSITY

**JASON A. GRISSOM**

VANDERBILT UNIVERSITY

CARNEGIE KNOWLEDGE NETWORK

**What We Know Series:**

Value-Added Methods and Applications

**ATIL**

ADVANCING TEACHING - IMPROVING LEARNING

## HIGHLIGHTS

- Value-added measures for principals have many of the same problems that value-added measures for teachers do, such as imprecision and questions about whether important outcomes are captured by the test on which the measures are based.
- While most measures of teachers' value-added and schools' value-added are based on a shared conception of the effects that teachers and schools have on their students, value-added measures for principals can vary in their underlying logic.
- The underlying logic on which the value-added measure is based matters a lot in practice.
- Evaluation models based on school *effectiveness*, which measure student test-score gains, tend not to be correlated at all with models based on school *improvement*, which measure *changes* in student test-score gains.
- The choice of model also changes the magnitude of the impact that principals appear to have on student outcomes.
- Estimates of principal effectiveness that are based on school effectiveness can be calculated for most principals. But estimates that are based on school effectiveness relative to the effectiveness of *other* principals who have served at the same school or estimates that are based on school *improvement* have stricter data requirements and, as a result, cover fewer principals.
- Models that assume that most of school effectiveness is attributable to the principal are more consistent with other measures of principal effectiveness, such as evaluations by the district. However, it is not clear whether these other measures are themselves accurate assessments.
- There is little empirical evidence on the advantages or disadvantages of using value-added measures to evaluate principals.

## INTRODUCTION<sup>1</sup>

Principals play a central role in how well a school performs. They are responsible for establishing school goals and developing strategies for meeting them. They lead their schools' instructional programs, recruit and retain teachers, maintain the school climate, and allocate resources. How well they execute these and other leadership functions is a key determinant of school outcomes.<sup>2</sup>

Recognizing this link between principals and school success, policymakers have developed new accountability policies aimed at boosting principal performance. In particular, policymakers increasingly are interested in evaluating school administrators based in part on student performance on standardized tests. Florida, for example, passed a bill in 2011 requiring that at least 50 percent of every school administrator's evaluation be based on student achievement growth as measured by state assessments and that these evaluations factor into principal compensation.

## WHAT DO WE KNOW ABOUT THE USE OF VALUE-ADDED MEASURES FOR PRINCIPAL EVALUATION?

Partly as a result of these laws, many districts are trying to create value-added measures for principals much like those they use for teachers. The idea is compelling, but the situations are not necessarily analogous. Estimating value-added for principals turns out to be even more complex than estimating value-added for teachers.

Three methods have been suggested for assessing a principal's value-added. One method attributes all aspects of school effectiveness (how well students perform relative to students at other schools with similar background characteristics and students with similar peers) to the principal; a second attributes to the principal only the difference between the effectiveness of that school under that principal and the effectiveness of the same school under other principals; and a third attributes school improvement (gains in school effectiveness) to the principal. Each method has distinct strengths, and each has significant drawbacks. There is now little empirical evidence to validate any of these methods as a way to accurately evaluate principals.

While substantial work has shaped our understanding of the many ways to use test scores to measure teacher effectiveness, far less research has focused on how to use similar measures to judge school administrators. The current state of our knowledge is detailed below.

**Using test scores.** When we use test scores to evaluate principals, three issues are particularly salient: understanding the mechanisms by which principals affect student learning, potential bias in the estimates of the effects, and reliability of the estimates of the effects. The importance of *mechanisms* stems from the uncertainty about how principals affect student learning and, thus, how student test scores should be used to measure it. *Potential bias* comes from misattributing factors outside of the principal's control to value-added measures. *Reliability*, or lack thereof, comes from imprecision in performance measures that results from random variations in test performance and idiosyncratic factors outside a principal's control.

How best to create measures of a principal's influence on learning depends crucially on the relationship between a principal's performance and student performance. Two issues are particularly germane here. The first is the time span over which a principal's decisions affect students. For instance, one might reasonably question how much of an impact principals have in their first year in a school, given the likelihood that most of the staff were there before the principal arrived and are accustomed to existing processes.

Consider a principal who is hired to lead a low-performing school. Suppose this principal excels from the start. How quickly would you expect that excellent performance to be reflected in student outcomes? The answer depends on the ways in which the principal has impact. If the effects are realized through better teacher assignments or incentives to students and teachers to exert more effort, they might be reflected in student performance immediately. If, on the other hand, a principal makes her mark through longer-term changes, such as hiring better teachers or creating environments that encourage effective teachers to stay, it may take years for her influence to be reflected in student outcomes. In practice, principals likely have both immediate and longer-term effects. The timing of principals' effects are important for how we should measure principal value-added and also point to the importance of the length of principal tenure in using value-added measurements to assess principals.

The second consideration is distinguishing the principal effect from characteristics of the school that lie outside of the principal's control. It may be that the vast majority of a school's effects on learning, aside from those associated with the characteristics of the students, is attributable to the principal's

performance. In this case, identifying the overall school effect (adjusted for characteristics of the students when they entered the school) is enough to identify the principal effect. That is, the principal effect is equal to the school effect.<sup>3</sup>

Alternatively, school factors outside of the principal's control may be important for school effectiveness. For example, what happens when principals have little control over faculty selection—when the district's central office does the hiring, or when it is tightly governed by collective bargaining agreements? One means for improving a school—hiring good people—will be largely outside a principal's control, though a principal could still influence the development of teachers in the school as well as the retention of good teachers. As another example, some schools may have a core of teachers who work to help other teachers be effective, and these core teachers may have already been at the school before the principal arrived. Other schools may benefit from an unusually supportive and generous community leader, someone who helps the school even without the principal's efforts. In all of these cases, if the goal is to identify principal effectiveness, it will be important to net out the effects of factors that affect school effectiveness but are outside of the principal's control.<sup>4 5</sup>

How one thinks about these two theoretical issues—the timing of the principal effect and the extent of a principal's influence over schools—has direct implications for how we estimate the value that a principal adds to student performance. Three possible approaches for estimating value-added make different assumptions about these issues.

**Principal value-added as *school effectiveness*.** First, consider the simplest case, in which principals immediately affect schools and have control over all aspects of the school that affect learning except those associated with student characteristics. That is, school effectiveness is completely attributable to the principal. If this assumption holds, an appropriate approach to measuring the contribution of that principal would be to measure school effectiveness while the principal is working there, or how well students perform relative to students with similar background characteristics and peers. This approach is essentially the same as the one used for teachers; we assume that teachers have immediate effects on students during the year they have them, so we take students' growth during that year—controlling for various factors—as a measure of that teacher's impact. For principals, any growth in student learning that is different than that predicted for a similar student in a similar context is attributed to the principal.

This approach has some validity for teachers. Because teachers have direct and individual influences on their students, it makes sense to take the adjusted average learning gains of students during a year as a measure of that teacher's effect. The face validity of this kind of approach for principals, however, is not as strong. While the effectiveness of a school may be due in part to its principal, it may also result in part from factors that were in place before the principal took over. Many teachers, for example, may have been hired previously; the parent association may be especially helpful or especially distracting. Particularly in the short run, it would not make sense to attribute all of the contributions of those teachers to that principal. An excellent new principal who inherits a school filled with poor teachers—or a poor principal hired into a school with excellent teachers—might incorrectly be blamed or credited with results he had little to do with.

**Principal value-added as *relative school effectiveness*.** The misattribution of school effects outside of a principal's control can create bias in the estimates of principal effectiveness. One alternative is to compare the effectiveness of a school during one principal's tenure to the effectiveness of the school at other times. The principal would then be judged by how much students learn (as measured by test

scores) while that principal is in charge, compared to how much students learned in that same school when someone else was in charge. Conceptually, this approach is appealing if we believe that the effectiveness of the school that a principal inherits affects the effectiveness of that school during the principal's tenure. And it most likely does.

One drawback of this “within-school over-time” comparison is that schools change as neighborhoods change and teachers turn over. That is, there are possible confounding variables for which adjustments might be needed. While this need is no different than that for the first approach described above, the within-school over-time approach has some further drawbacks. In particular, given the small number of principals that schools often have over the period of available data, the comparison sets can be tiny and, as a result, idiosyncratic. If, in available data, only one principal serves in a school, there is no other principal to whom to compare her. If there are only one or two other principals, the comparison set is very small, leading to imprecision in the estimates. The within-school over-time approach holds more appeal when data cover a period long enough for a school to have had several principals. However, if there is little principal turnover, if the data stream is short, or if there are substantial changes in schools that are unrelated to the school leadership, this approach may not be feasible or advisable.

**Principal value-added as *school improvement*.** So far we have considered models built on the assumption that principal performance is reflected immediately in student outcomes and that this reflection is constant over time. Perhaps more realistic is an expectation that new principals take time to make their marks, and that their impact builds the longer they lead the school. School improvement that comes from building a more productive work environment (from skillful hiring, for instance, or better professional development or creating stronger relationships) may take a principal years to achieve. If it does, we may wish to employ a model that accounts explicitly for this dimension of time.

One such measure would capture the *improvement* in school effectiveness during the principal's tenure. The school may have been relatively ineffective in the year before the principal started, or even during the principal's first year, but if the school improved during the principal's overall tenure, that would suggest the principal was effective. If the school's performance declined, it would point to the reverse.

The appeal of such an approach is its clear face validity. However, it has disadvantages. In particular, the data requirements are substantial. There is error in any measure of student learning gains, and calculating the difference in these imperfectly measured gains to create a principal effectiveness measure increases the error.<sup>6</sup> Indeed, this measure of principal effectiveness may be so imprecise as to provide little evidence of actual effectiveness.<sup>7</sup> In addition, as with the second approach, if the school were already improving because of work done by former administrators, we may overestimate the performance of principals who simply maintain this improvement.

We have outlined three general approaches to measuring principal value-added. The ***school effectiveness*** approach attributes all of the learning benefits of attending a given school while the principal is leading it to that principal. The ***relative school effectiveness*** approach attributes the learning benefits of attending a school while the principal is leading it relative to the benefits of the same school under other principals. The ***school improvement*** approach attributes the changes in school effectiveness during a principal's tenure to that principal. These three approaches are each based on a conceptually different model of principals' effects, and each will lead to different concerns about validity (or bias) and precision (or reliability).

## WHAT IS THE CURRENT STATE OF KNOWLEDGE ON THIS ISSUE?

Value-added measures of teacher effectiveness and school effectiveness are the subject of a large and growing research literature summarized in part by this series.<sup>8</sup> In contrast, the research on value-added measures of principal effectiveness—as distinct from school effectiveness—is much less extensive. Moreover, most measures of teachers' value-added and schools' value-added are based on a shared conception of the effect that teachers and schools have on their students. By contrast, value-added measures of principals can vary both by their statistical approach and their underlying logic.

One set of findings from Miami-Dade County Public Schools compares value-added models based on the three conceptions of principal effects described above: school effectiveness, relative school effectiveness, and school improvement. A number of results emerge from these analyses. First, the model matters a lot. In particular, models based on school improvement (essentially *changes* in student test score gains across years) tend not to be correlated at all with models based on school effectiveness or relative school effectiveness (which are measures of student test score gains over a single year).<sup>9</sup> That is, a principal who ranks high in models of school improvement is no more or less likely to be ranked high in models of school effectiveness than are other principals. Models based on school effectiveness and those based on relative school effectiveness are more highly correlated, but still some principals will have quite different ratings on one than on the other. Even within conceptual approaches, model choices can make significant differences.

Model choice affects not only whether one principal appears more or less effective than another but also how important principals appear to be for student outcomes. The variation in principal value-added is greater in models based on school effectiveness than in models based on improvement, at least in part because the models based on improvement have substantial imprecision in estimates.<sup>10 11</sup> Between models of school effectiveness and models of relative school effectiveness (comparing principals to other principals who have taught in the same school), the models of school effectiveness show greater variation across principals.<sup>12</sup> For example, in one study of North Carolina schools, the estimated variation in principal effectiveness was more than four times greater in the model that attributes school effects to the principal than in the model that compares principals within schools.<sup>13</sup> This finding is not surprising given that the models of relative school effectiveness have taken out much of the variation that exists across schools, looking only within schools over time or with a group of schools that share principals.

The Miami-Dade research also provides insights into some practical problems with the measures introduced above. First, consider the model that compares principals to other principals who serve in the same school. This approach requires each school to have had multiple principals. Yet in the Miami-Dade study, even with an average annual school-level principal turnover rate of 22 percent over the course of eight school years, 38 percent of schools had only one principal.<sup>14 15</sup> Even when schools have had multiple principals over time, the number in the comparison group is almost always small. The within-school relative effectiveness approach, in essence, compares principals to the few other principals who have led the schools in which they have worked, then assumes that each group of principals (each set of principals who are compared against each other) is, on average, equal. In reality, they may be quite different. In the Miami-Dade study, the average principal was compared with fewer than two other principals in value-added models based on within-school relative effectiveness. The other two approaches (school effectiveness and school improvement) used far larger comparison groups.

Measures of principal value-added based on school improvement also require multiple years of data. There is no improvement measure for a single year, and even two or three years of data are often insufficient for calculating a stable trend. Requiring principals to lead a school for three years in order to calculate value-added measures reduced the number of principals by two-thirds in the Miami-Dade study.<sup>16</sup> A second concern with using school improvement is imprecision. As described above, there is more error in measuring *changes* in student learning than in measuring *levels* of student learning. There simply may not be information left in the measures based on school improvement to be useful as a measure of value-added.

While there are clear drawbacks to using value-added measures based on school improvement, the approach also has substantial conceptual merit. In many cases, good principals do, in fact, improve schools. The means by which they do so can take time to reveal themselves.<sup>17</sup> Moreover, one study of high schools in British Columbia points to meaningful variation across principals in school improvement.<sup>18</sup>

To better understand the differences in value-added measures based on different approaches, the Miami-Dade study compared a set of value-added measures to: schools' accountability grades;<sup>19</sup> the district's ratings of principal effectiveness; students', parents' and staff's assessments of the school climate; and to principals' and assistant principals' assessments of the principal's effectiveness at certain tasks. These comparisons show that the first approach—attributing school effectiveness to the principal—is more predictive of all the non-test measures than are the other two approaches, although the second approach is positively related to many of the other measures as well. The third approach, measuring value-added by school improvement, is not positively correlated with any of these other measures. The absence of a relationship between measures of school improvement and measures of these other things could be the result of imprecision, or it could be because the improvement is based on a different underlying theory about how principals affect schools.

The implications of these results may not be as clear as they first seem. The non-test measures appear to validate the value-added measure that attributes all school effectiveness to the principal. Alternatively, the positive relationships may represent a shortcoming in the non-test measures. District officials, for example, likely take into account the effectiveness of the school itself when rating the performance of the principal. When asked to assess a principal's leadership skills, assistant principals and the principals themselves may base their ratings partly on how well the school is performing instead of solely on how the principal is performing. In other words, differentiating the effect of the principal from that of other school factors may be a difficulty encountered by both test-based and subjective estimates of principal performance.

In sum, there are important tradeoffs among the different modeling approaches. The simplest approach—attributing all school effectiveness to the principal—seems to give the principal too much credit or blame, but it produces estimates that correlate relatively highly across math and reading, across different schools in which the principal works, and with other measures of non-test outcomes that we care about. On the other hand, the relative school effectiveness approach and the school improvement approach come closer to using a reasonable conception of the relationship between principal performance and student outcomes, but the data requirements are stringent and may be prohibitive. These models attempt to put numbers on phenomena when we may simply lack enough data to do so.



Other research on principal value-added goes beyond comparing measurement approaches to using specific measures to gain insights into principal effectiveness. One such study, which used a measure of principal value-added that was based on school effectiveness, found greater variation among principal effectiveness in high-poverty schools than in other schools. This study provides some evidence that principals are particularly important for student learning in these schools, and it highlights the point about the effects of model choice on the findings.<sup>20</sup> A number of studies have used value-added measures to quantify the importance of principals for student learning. The results are somewhat inconsistent, with some finding substantially larger effects than others. One study of high school principals in British Columbia that used the within-schools approach finds a standard deviation of principal value-added that is even greater than that which is typical for teachers. Most studies, however, find much smaller differences, especially when estimates are based on within-school models.<sup>21</sup>

### WHAT MORE NEEDS TO BE KNOWN ON THIS ISSUE?

Using student test scores to measure principal performance faces many of the same difficulties as using them to measure teacher performance. As an example, the test metric itself is likely to matter.<sup>22</sup> Understanding the extent to which principals who score well on measures based on one outcome (e.g., math performance) also perform well on measures based on another outcome (e.g., student engagement) would help us understand whether principals who look good on one measure also look good on other measures. If value-added based on different measures is inconsistent, it will be particularly important to choose outcome measures that are valued.

Nonetheless, there are challenges to using test scores to measure principal effectiveness that differ from those associated with using such measures for teachers. These, too, could benefit from additional research. In particular, a better understanding of how principals affect schools would be helpful. For example, to what extent do principals affect students through their influence on veteran teachers, providing supports for improvement as well as ongoing management? Do they affect students primarily through the composition of their staffs, or can they affect students, regardless of the staff, with new curricular programs or better assignment of teachers? To what extent do principals affect students through cultural changes? How long does it take for these changes to have an impact? Clearer answers to these questions could point to the most appropriate ways of creating value-added measures.

No matter how much we learn about the many ways in which principals affect students, value-added measures for these educators are going to be imperfect; they probably will be both biased and imprecise. Given these imperfections, can value-added measures be used productively? If so, under what circumstances? As do many managers, principals perform much of their work away from the direct observation of their employers. As a result, their employers need measures of performance other than observation. Research can clarify where the use of value-added improves outcomes, and whether other measures, in combination with or instead of value-added, lead to better results. There is now little empirical evidence to warrant the use of value-added data to evaluate principals, just as there is little clear evidence against it.

### WHAT CAN'T BE RESOLVED BY EMPIRICAL EVIDENCE ON THIS ISSUE?

The problems with outcome-based measures of performance are not unique to schooling. Managers are often evaluated and compensated based on performance measures that they can only partially



control.<sup>23</sup> Imperfect measures can have benefits if they result in organizational improvement. For example, using student test scores to measure productivity may encourage principals to improve those scores even if the value-added measures are flawed. However, whether such measures actually do lead to improvement will depend on the organizational context and the individuals in question.<sup>24</sup>

This brief has highlighted many of the potential flaws of principal value-added measures, pointing to the potential benefit of additional or alternative measures. One set of measures could capture other student outcomes, such as attendance or engagement. As with test scores, highlighting these factors creates incentives for a principal to improve them, even though these measures likely would share with test-based value-added the same uncertainty about what to attribute to the principal. Another set of measures might more directly gauge principals' actions and the results of those actions, even if such measures are likely more costly than test-score measures to devise. These measures might come from feedback from teachers, parents, students, or from a combination of observations and discussions between district leaders and principals.

Research can say very little about how to balance these different types of measures. Would the principals (and their schools) benefit from the incentives created by evaluations based on student outcomes? Does the district office have the capacity to implement more nuanced evaluation systems? Would the dollars spent on such a system be worth the tradeoff with other potentially valuable expenditures? These are management decisions that research is unlikely to directly inform.

## CONCLUSION

The inconsistencies and drawbacks of principal value-added measures lead to questions about whether they should be used at all. These questions are not specific to principal value-added. They apply, at least in part, to value-added measures for teachers and to other measures of principal effectiveness that do not rely on student test performance. There are no perfect measures, yet district leaders need information on which to make personnel decisions. Theoretically, if student test performance is an outcome that a school system values, the system should use test scores in some way to assess schools and hold personnel accountable. Unfortunately, we have no good evidence about how to do this well.

The warning that comes from the research so far is to think carefully about what value-added measures reveal about the contribution of the principal and to use the measures for what they are. What they are *not* is a clear indicator of a principal's contributions to student test-score growth; rather, they are an indicator of student learning in that principal's school compared with learning that might be expected in a similar context. At least part of this learning is likely to be due to the principal, and additional measures can provide further information about the principal's role. To the extent that districts define what principals are supposed to be doing—whether that is improving teachers' instructional practice, student attendance, or the retention of effective teachers—measures that directly capture these outcomes can help form an array of useful but imperfect ways to evaluate principals' work.

## ENDNOTES

<sup>1</sup> The Introduction is drawn from: Grissom, Jason A., Demetra Kalogrides, and Susanna Loeb, "Using student test scores to measure principal performance," (National Bureau of Economic Research, No. w18568, 2012).

<http://www.nber.org/papers/w18568>

<sup>2</sup> The research base documenting the large array of roles principals play and how execution of those roles impacts the school is extensive. For a review, see Hallinger, Philip, and Ronald H. Heck, "Exploring the principal's contribution to school effectiveness: 1980-1995," *School Effectiveness and School Improvement* 9 (2) (1998): 157-191. <http://www.tandfonline.com/doi/abs/10.1080/0924345980090203#preview>

On instructional leadership and student outcomes, see Robinson, Viviane MJ, Claire A. Lloyd, and Kenneth J. Rowe, "The impact of leadership on student outcomes: An analysis of the differential effects of leadership types," *Educational Administration Quarterly* 44 (5) (2008): 635-674. For a more recent study linking principals' management skills to student learning growth, see Grissom, Jason A., and Susanna Loeb, "Triangulating principal effectiveness: How perspectives of parents, teachers, and assistant principals identify the central importance of managerial skills," *American Educational Research Journal* 48(5) (2011): 1091-1123.

<http://cepa.stanford.edu/content/triangulating-principal-effectiveness-how-perspectives-parents-teachers-and-assistant>

<sup>3</sup> For a more detailed discussion of school effects see Raudenbush, Stephen W., and J. Douglas Willms, "The estimation of school effects," *Journal of Educational and Behavioral Statistics* 20 (4) (1995): 307-335. The authors point out the difficulty of separating peer and community characteristics from effective school practices.

<sup>4</sup> This same issue arises with teachers and has received considerable attention from researchers. For an example, see Rothstein, Jesse, "Student sorting and bias in value-added estimation: Selection on observables and unobservables," *Education Finance and Policy* 4(4) (2009): 537-571. <http://www.nber.org/papers/w14666>

<sup>5</sup> Teacher value-added measures are not immune to similar concerns. For example, a teacher might have a class of students who are unusually supportive of each other. However, the issues for estimating value-added are somewhat less problematic for teachers because the students they work with vary from class to class and year to year whereas most of the teachers and students that principals work with remain in the same school from year to year.

<sup>6</sup> Kane, Thomas J., and Douglas O. Staiger. "The promise and pitfalls of using imprecise school accountability measures," *The Journal of Economic Perspectives* 16 (4) (2002): 91-114.

<http://www.dartmouth.edu/~dstaiger/pub.html>; Boyd, Donald, Hamilton Lankford, Susanna Loeb, and James Wyckoff, "Measuring test measurement error: A general approach," (National Bureau of Economic Research, No. w18010, 2012). <http://www.nber.org/papers/w18010>

<sup>7</sup> Kane, et al, 2002, *ibid*; Boyd, et al, 2012, *ibid*.

<sup>8</sup> See for example: Aitkin, Murray, and Nicholas Longford, "Statistical modeling issues in school effectiveness studies," *Journal of the Royal Statistical Society, Series A (General)* (1986): 1-43.

<sup>9</sup> Grissom, et al, 2012, *ibid*.

<sup>10</sup> *ibid*.

<sup>11</sup> Estimates of improvement have greater measurement error because there is imprecision in both the starting effectiveness of the school and the ending point effectiveness. The improvement is the difference between these two imprecise measures and, as a result, it has two sources of error.

<sup>12</sup> Grissom, et al, 2012, *ibid*.

<sup>13</sup> Dhuey, Elizabeth and Justin Smith, "How school principals influence student learning?" (Working Paper, 2013).

<http://homes.chass.utoronto.ca/~edhuey/index.php/home/page/research>

<sup>14</sup> Authors' calculations.

<sup>15</sup> Beteille, Tara., Demetra Kalogrides, and Susanna Loeb, "Stepping stones: Principal career paths and school outcomes," *Social Science Research* 41(4) (2012): 904-919.

<sup>16</sup> Grissom, Jason A., Demetra Kalogrides, and Susanna Loeb, "Using student test scores to measure principal performance," (National Bureau of Economic Research, No. w18568, 2012). <http://www.nber.org/papers/w18568>

Note: The principals for whom school improvement models are feasible – those who remain in the same school for a longer period of time – are those for whom the relative effectiveness measures are likely to be less feasible, because fewer principals have likely led their school.

<sup>17</sup> See for example, Anthony S. Bryk, Penny Bender Sebring, Elaine Allensworth, Stuart Luppescu, and John Q. Easton, *Organizing Schools for Improvement: Lessons from Chicago*, (Chicago: University of Chicago Press, 2010).

<sup>18</sup> Coelli, Michael, and David A. Green, "Leadership effects: School principals and student outcomes," *Economics of Education Review* 31(1) (2012): 92-109.

<sup>19</sup> Florida grades each school on a 5-point scale (A, B, C, D, F) that is meant to succinctly capture performance. Grades are based on a scoring system that assigns points to schools for their percentages of students achieving the highest levels in reading, math, science, and writing on Florida's standardized tests in grades 3 through 10, or who make achievement gains. Grades also factor in the percentage of eligible students who are tested and the test gains of the lowest-performing students.

<sup>20</sup> Branch, Gregory F., Eric A. Hanushek, and Steven G. Rivkin, "Estimating the effect of leaders on public sector productivity: The case of school principals," (National Bureau of Economic Research, No. w17803, 2012).

<http://www.nber.org/papers/w17803>

<sup>21</sup> Dhuey, Elizabeth, and Justin Smith, "How important are school principals in the production of student achievement?" (Working Paper, 2013). <http://homes.chass.utoronto.ca/~edhuey/index.php/home/page/research>

<sup>22</sup> Kane, Thomas, and Steven Cantrell. "Learning about teaching: Initial findings from the measures of effective teaching project," (Bill & Melinda Gates Foundation, MET Project Research Paper, 2010).

<http://www.metproject.org/reports.php>; Lockwood, J. R., Daniel F. McCaffrey, Laura S. Hamilton, Brian Stecher, Vi-Nhuan Le, and José Felipe Martinez, "The sensitivity of value-added teacher effect estimates to different mathematics achievement measures," *Journal of Educational Measurement* 44(1) (2007): 47-67; Kane, Thomas J., Daniel F. McCaffrey, Trey Miller, and Douglas Staiger, "Have we identified effective teachers? Validating measures of effective teaching using random assignment," (Bill & Melinda Gates Foundation, MET Project Research Paper, 2013). <http://www.metproject.org/reports.php>

<sup>23</sup> Graham, John R., Si Li, and Jiaping Qiu, "Managerial attributes and executive compensation," (National Bureau of Economic Research, No. w17368, 2011). <http://www.nber.org/papers/w17368>

<sup>24</sup> Heinrich, Carolyn J., and Gerald Marschke, "Incentives and their dynamics in public sector performance management systems," *Journal of Policy Analysis and Management* 29(1) (2010): 183-208.

## AUTHORS



**Susanna Loeb** is the Barnett Family Professor of Education at Stanford University, Faculty Director of the Center for Education Policy Analysis, and a Co-Director of Policy Analysis for California Education (PACE). She specializes in the economics of education and the relationship between schools and federal, state and local policies. Her research addresses teacher policy, looking specifically at how teachers' preferences affect the distribution of teaching quality across schools, how pre-service coursework requirements affect the quality of teacher candidates, and how reforms affect teachers' career decisions. She also studies school

leadership and school finance, for example looking at how the structure of state finance systems affects the level and distribution of resources across schools. Susanna is a senior fellow at the Stanford Institute for Economic Policy Research, a faculty research fellow at the National Bureau of Economic Research, a member of the Policy Council of the Association for Policy Analysis and Management, and Co-Editor of *Educational Evaluation and Policy Analysis*.



**Jason A. Grissom** is an Assistant Professor of Public Policy and Education, at the Peabody College of Education and Human Development at Vanderbilt University. He studies the policy, politics, and governance dimensions of K-12 education. He is particularly interested in identifying the impacts of school and district leaders on teacher and student outcomes. He has published widely in education, public administration, and political science, with articles appearing in such outlets as *American Educational Research Journal*, *American Journal of Education*, *Educational Evaluation and Policy Analysis*, *Journal of Policy Analysis and Management*, *Journal of Politics*, *Journal of*

*Public Administration Research and Theory*, *Public Administration Review*, and *Teachers College Record*. Currently, he is principal investigator on an Institute of Education Sciences-funded, longitudinal study of school leader effectiveness in four large, urban districts. Other current research explores the causes and consequences of administrator turnover, school board decision-making, and public opinion in education. Professor Grissom holds a Masters degree in Education and a PhD in Political Economics from Stanford University.

## ABOUT THE CARNEGIE KNOWLEDGE NETWORK

The Carnegie Foundation for the Advancement of Teaching has launched the Carnegie Knowledge Network, a resource that will provide impartial, authoritative, relevant, digestible, and current syntheses of the technical literature on value-added for K-12 teacher evaluation system designers. The Carnegie Knowledge Network integrates both technical knowledge and operational knowledge of teacher evaluation systems. The Foundation has brought together a distinguished group of researchers to form the *Carnegie Panel on Assessing Teaching to Improve Learning* to identify what is and is not known on the critical technical issues involved in measuring teaching effectiveness. Daniel Goldhaber, Douglas Harris, Susanna Loeb, Daniel McCaffrey, and Stephen Raudenbush have been selected to join the Carnegie Panel based on their demonstrated technical expertise in this area, their thoughtful stance toward the use of value-added methodologies, and their impartiality toward particular modeling strategies. The Carnegie Panel engaged a User Panel composed of K-12 field leaders directly involved in developing and implementing teacher evaluation systems, to assure relevance to their needs and accessibility for their use. This is the first set of knowledge briefs in a series of Carnegie Knowledge Network releases. Learn more at [carnegieknowledgenetwork.org](http://carnegieknowledgenetwork.org).

## CITATION

Loeb, Susanna, and Jason A. Grissom. Carnegie Knowledge Network, "What Do We Know About the Use of Value-Added Measures for Principal Evaluation?" Last modified July 2013. URL = <http://carnegieknowledgenetwork.org/briefs/value-added/principal-evaluation/>



## **Carnegie Foundation** for the Advancement of Teaching

Carnegie Foundation for the Advancement of Teaching  
51 Vista Lane  
Stanford, California 94305  
650-566-5100

Carnegie Foundation for the Advancement of Teaching seeks to vitalize more productive research and development in education. We bring scholars, practitioners, innovators, designers, and developers together to solve practical problems of schooling that diminish our nation's ability to educate all students well. We are committed to developing networks of ideas, expertise, and action aimed at improving teaching and learning and strengthening the institutions in which this occurs. Our core belief is that much more can be accomplished together than even the best of us can accomplish alone.

**[www.carnegiefoundation.org](http://www.carnegiefoundation.org)**

We invite you to explore our website, where you will find resources relevant to our programs and publications as well as current information about our Board of Directors, funders, and staff.



This work is licensed under the Creative Commons Attribution-Noncommercial 3.0 Unported license.

To view the details of this license, go to:  
<http://creativecommons.org/licenses/by-nc/3.0/>.

**Knowledge Brief 9**  
**JULY 2013**

**[carnegieknowledgenetwork.org](http://carnegieknowledgenetwork.org)**

*Funded through a cooperative agreement with the Institute for Education Science.  
The opinions expressed are those of the authors and do not represent views of the  
Institute or the U.S. Department of Education.*