



REVIEW OF TWO ARTICLES ON PRINCIPAL EFFECTS

Reviewed By

Margaret Terry Orr

Bank Street College of Education

March 2013

Summary of Review

Two articles by the same authors estimate the effects of principals on student achievement. Both report school-level, value-added Texas state test results for school principals, controlling for gender, race, free and reduced-price lunch, and earlier academic performance. They find large to small effect sizes depending on the model examined. The study is severely limited by methodological issues, including the age of the data (12-18 years old) and failure to account for several variables: experience as a principal, the effectiveness of comparison principals, district assignment policies, nonlinear effects of principals on schools, other influences on high-poverty schools, and non-principal influenced teacher mobility. Regarding principal mobility, the authors failed to account for district policies and career mobility patterns. Finally, the authors could not estimate the effects of first-year principals. These flaws raise serious questions about the actual effect sizes of principals on student test scores and thus the validity of the analyses. The report is most useful for methodological discussions about value-added estimates for principals, the validity of different models, and principal effects on teacher and principal turnover. The most important policy-relevant conclusion that can be derived from this report is that estimating principal effectiveness is simply not possible given current methodology and sample size restrictions. Thus, such estimates should not be used to evaluate principals.

Kevin Welner

Project Director

William Mathis

Managing Director

Erik Gunn

Managing Editor

National Education Policy Center

School of Education, University of Colorado

Boulder, CO 80309-0249

Telephone: (802) 383-0058

Email: NEPC@colorado.edu

<http://nepc.colorado.edu>

Publishing Director: Alex Molnar



This is one of a series of Think Twice think tank reviews made possible in part by funding from the Great Lakes Center for Education Research and Practice. It is also available at <http://greatlakescenter.org>.

This material is provided free of cost to NEPC's readers, who may make non-commercial use of the material as long as NEPC and its author(s) are credited as the source. For inquiries about commercial use, please contact NEPC at nepc@colorado.edu.

REVIEW OF *ESTIMATING THE EFFECT OF LEADERS ON PUBLIC SECTOR PRODUCTIVITY: THE CASE OF SCHOOL PRINCIPALS AND “SCHOOL LEADERS MATTER,”* IN *EDUCATION NEXT*

Margaret Terry Orr, Bank Street College of Education

I. Introduction

Estimating the effect of leaders on public sector productivity: The case of school principals, by Gregory F. Branch, Eric A. Hanushek and Steven G. Rivkin, and published by the National Bureau of Economic Research in 2012, and a related, less technical 2013 article, “School Leaders Matter,” by the same authors and published in *Education Next* (Branch et al., 2013), estimate how much high and low effective principals raise student achievement in a given year.¹ Both the report and the article estimate the effects that principals have on student test scores. They then attempt to show variations in principal effectiveness, controlling for students’ race and ethnicity, gender, free and reduced-price lunch eligibility, special education participation, and English as a second language status and the school’s previous test scores. Both also examine how teacher and principal turnover correlate to principal effectiveness classification and student achievement gains, particularly for high-poverty schools.

The report’s and article’s primary purpose was to measure the impact of school leaders by “separating their contributions from the many other factors that drive student achievement” (2013, p. 2). Both conclude that there is substantial variation in principal effectiveness: that, in comparison with other principals, highly effective principals produce higher annual student gains and have higher teacher turnover in grades where teachers are least effective. The reports find that least and most effective principals have more career transitions, raising principal labor market questions. The policy conclusion is to focus on the selection and retention of high quality principals as instrumental for improved student outcomes.

II. Findings and Conclusions of the Reports

The NBER report begins by situating its work within broader research on measuring public sector leadership quality. It reviews the importance of school leadership, the current state

of research on principal effectiveness, and considerations for separating the student achievement contributions of principal effectiveness from other factors. After describing its dataset, the report presents findings on principal transition patterns by job tenure, school poverty and achievement. Next it calculates estimates of principal effectiveness and looks at whether more effective principals were concentrated in one type of school (based on poverty) or, instead, equally distributed.

The researchers use three value-added models, all of which take into account school demographic characteristics. The first model restricts the sample to principals in their first three years at their current school and reports the standard deviation of principal effectiveness and the estimated variance within each principal effectiveness percentile and by share of students living in poverty. The report raises test measurement issues (such as the structure of the tests, test measurement error, construction of the test [the TAAS is focused on lower-level skills]), the distribution of skill proficiency among students and how these proficiency distributions may be related to school poverty levels. The researchers tested two alternative methods of addressing these issues in how prior achievement is specified and principal performance is weighted. These attempts to shrink the variance had little effect on the variance estimates for principal effectiveness. The authors also thought that reweighting the tests' score distribution—because the “greatest concentration of test scores at the lower end of the achievement distribution causes the larger variance in principal quality observed for high poverty schools”(p. 14)—was exacerbating unobserved influences for students in small cells.

The second value-added method added school fixed-effect estimates, restricting the comparison to other principals who lead the same school. It uses the full sample of principals and does not control for years at the school. The report shows the estimated standard deviation of principal value added by school poverty quartile with and without school fixed effects, yielding 0.11 and 0.22 s.d., respectively.

A third approach examines principal effects by evaluating “how the variance in achievement growth changes across student cohorts as new principals assume leadership in a school” (p. 16). It uses year-to-year squared differences in school average student gains for different principals and for principals in their first year in adjacent and nonadjacent years, with student controls. Further, the authors consider several assumptions about the relationship and the bias that would result when these assumptions are violated. Ultimately, the three models resulted in very different effect sizes ranging from 0.05 to 0.22 s.d.

The report then addresses the relationship between principal quality and teacher turnover generally and among teachers in lower-performing grades in particular. The final findings of the paper document principal moves from school to school, as well as from employment in different non-principal positions within schools and in the central office, and relates these to both school poverty levels and principal effectiveness levels. In short, this section provided a comparison of career pathways for principals leaving their position from various schools and performance experiences. The report concludes by restating the importance of a “systematic analysis of principals and their impact on student outcomes”

(p. 27) and acknowledging the analytic challenges such as controlling for unobserved characteristics of schools through the use of peer effects, addressing test measurement issues, and including beginning principals in models using school fixed effects. It restates the report's key findings and returns to the larger context of understanding the variation of quality in public sector managers generally.

The non-technical article follows a similar organization, but excludes the larger public sector leader quality context, the statistical modeling behind each estimation method, and some of the bias considerations. While the technical report includes 10 tables and one figure on the various principal estimates, the non-technical article only includes a text table on methods and two figures illustrating the relationship between principal turnover, on the one hand, and student achievement, poverty, and principal quality, on the other. The central finding as summarized in the *Education Next* article is as follows: "Our results indicate that highly effective principals raise the achievement of a typical student in their schools by between two and seven months of learning in a single year; ineffective principals lower achievement by the same amount" (p. 1).

III. The Report's and Article's Rationale for Its Findings and Conclusions

The research report and non-technical article define their purposes based on somewhat different rationales. The report focuses on the need to distinguish leadership quality as key to organizational success. The article, in contrast, cites federal policy and foundation investment in school leadership and its preparation as key to school success. Both the report and article stress how little rigorous research exists on principal contributions to growth in student achievement. While the article does not explain the conceptual basis for the study's research, the report explains that it is using Bertrand and Schoar's (2003) semi-parametric approach to estimate the variance in principal effectiveness.² Semi-parametric analyses are often used when the relationships between variables are not linear. For example, the influence of a principal on student test scores may be small in the first year and grow systematically in a non-linear fashion through years five or six, then remain relatively constant over time. The report also draws on other recent research that similarly investigates how to model these estimates, and on some of the authors' prior work on teacher estimates.³

IV. The Report's Use of Research Literature

While both the report and the article cite how little rigorous research exists on this topic, only the research report discusses related and contextual research. This includes 27 peer-reviewed research articles, three non-peer reviewed articles, two book chapters and seven governmental, foundation and organizational reports and unpublished research papers

(including two from the National Bureau of Economic Research). Only a small portion of the cited research is used to frame the study or guide the methodology, and most literature is referenced in footnotes.

The report's analysis of the research literature on principal effectiveness ignores a large body of research, claiming instead that "[m]ost prior analysis of principal effectiveness has been qualitative, though a small number of papers examine the determinants of principal effectiveness and any links to the labor market." (p. 5). Yet there has been significant large scale research that has combined survey and student achievement data in order to tease out the effects of specific principal practices on school and student outcomes.⁴

The analysis in this report only cites three earlier economic studies on the relationship of principal practices and student outcomes: Brewer (1993), Eberts and Stone (1988) and Cullen and Mazzeo (2008). The first two articles examined the relationships between principal characteristics and behaviors with student achievement while the third article was an unpublished manuscript that examined the relationships between principal salary

The report's analysis of the research literature on principal effectiveness ignores a large body of research.

growth, employment transitions, and student achievement.⁵ While conceptually and methodologically limited in comparison to other more recent, large-scale research, these three studies found that selected principal behaviors and characteristics are positively related to improvements in student achievement.

The principal transition and related career research is only discussed in a footnote and limited to two studies.⁶ The report misses other important research, one using the same data source to investigate principal career patterns⁷ and others addressing conceptual and methodological issues to be considered in estimating principal effects on student test scores.⁸ At the time this report was released but before the article was published, several other value-added estimates of principal effectiveness studies were released or published.⁹ These more recent reports included this report in their background and comparison of methods and findings, and yielded somewhat more modest estimates, particularly when trying to isolate principal effects from other factors.

V. Review of the Report's Methods

The report is framed as a productivity study and applies a conventional management effectiveness approach to determine the productivity of school leaders. It uses student standardized test scores, rather than management outcomes of revenue or return on investment, as its performance measure. It modifies this model by applying a second approach to account for year-to-year fluctuations in achievement used by some of the

report authors to assess teacher effectiveness, to generate lower-bound estimates of the variance based on year-to-year variation in school achievement and career transitions.¹⁰

The report acknowledges, in analyzing prior studies, that creating measures of principal effectiveness may not be meaningful if too few relevant student, school and district controls are considered. Factors included in the report's analyses are eligibility for free and reduced-price lunch, gender, race and ethnicity, special education participation, English language learners, and the proportion of students new to the school. It does not take into account differences by school level. Most other studies of principal effects are limited to either grades 4-8 or high school and thus do not conflate school levels .

The study's methodology is ambitious in its scope and size, covering a six-year period and including all principals in one state for whom data are available. The authors use a Texas public administration database for the years 1995 through 2001 for their analyses. The database combines different data sources on students, teachers and principals, including individual and school-based demographic information and student achievement data. The data set includes "7,420 unique principals that yield 28,147 annual observations [for each principal's year] of principals" (p. 7). The subset of these principals and observations used in the analyses below are unspecified, however, making their representation difficult to determine.

The study's methodology is limited, however, because it lacks information on the quality and completeness of the data set, how school populations change over the analysis periods, the effects of the age of the data set, and its choice of performance measure. The authors explain little about the quality or completeness of the data set, except to note in a footnote that an unspecified number of "special education and limited English proficient students are exempted from the tests" (p. 7) and that about "15 percent of students do not take the tests, either because of an exemption or because of repeated absences on testing days" (p. 7). Nowhere do the authors address how these absences and exemptions are independently distributed among schools, how they affect the analyses or how their year-to-year variations may independently influence the variance estimates. The authors did not account for the likelihood that these absences and exemption rates were greater in high-poverty schools, thus possibly contributing to the greater variance in principal effects.

In addition, schools' student cohort population characteristics are not consistent over time and across schools. The authors do not discuss what impact student transfers and attrition might have on estimates of school and principal effectiveness. The proportion and performance level of student transfer and attribution is also likely to be uneven across schools and districts, with the proportion leaving likely to increase with school poverty rates. The authors could have included the proportion of students in a school with valid test scores for two consecutive years in an attempt to control for these issues.

Limiting the study to data for the years 1995 through 2001 is puzzling for methodological and policy reasons. First, more recent data were available. Other researchers have drawn on the same data source to examine principal career patterns, and were able to include more recent years, up through 2007¹¹ and 2008 to examine the relationship among

principal preparation, principals, and teacher selection.¹² Second, other researchers have used newer state or district data sets, spanning 2004-2008. Third, the years covered by the data in this analysis precede federal, state and local assessment-based accountability policies, which may have an independent influence on the quality and completeness of the assessment data, principal assignment policies, principal career patterns, and teacher selection and retention practices. Specifically, the state and districts may have been far more lax about testing conditions, alignment of instruction to assessment outcomes, and

The report's terminology in differentiating principal effectiveness is sloppy, adding ambiguity to the analyses and discussion.

principal and teacher selection, assignment and retention practices to improve student outcomes. Only in recent years as accountability pressures have increased with No Child Left Behind achievement goals and the human resource management aims of the federal Race to the Top and other grants, have districts begun to encourage schools and principals to align curriculum, instruction, school organization and human resource management practices in ways that would positively influence student achievement outcomes.

The report uses mathematics achievement as the performance measure, although reading achievement scores were also available. The report offers no explanation for selecting only mathematics and why these achievement data were not separately analyzed to validate the models. Other studies on principal effects used both reading and mathematics achievement scores, in part to test the relationships on separate performance outcomes.

In the value-added model without school fixed-effects, the authors focus only on newly hired principals in their first three years at a school. However, the authors do not appear to control for the prior experience as a principal in previous schools or for experience as an assistant principal in the same school or in other schools. This is critically important, given that more experienced principals tend to become employed in higher-performing schools; thus there is a non-random sorting of experienced, but newly hired, principals. The omission of such variables could certainly influence the results found in the study.

In the school fixed-effects model, the authors do not control for principal tenure for either the principal in question or the prior principal in the school. This is a serious flaw, since this model compares the effects of the current principal to the prior principal in the same school. Thus, for a particular school, the authors could be comparing the effects for a principal with no prior experience to a principal with ten years of experience. Other research from Texas suggests that newly hired principals at that time would systematically have had less experience than prior principals, particularly in high poverty districts.¹³

Finally, the report's terminology in differentiating principal effectiveness is sloppy, adding ambiguity to the analyses and discussion. The terms more or less "effective" principals and more or less "successful" principals, "highly rated" and "less highly rated" principals, "better" principals and others, and "principal quality" are used interchangeably throughout the report.

VI. Review of the Validity of the Findings and Conclusions

The report and article are among the first in a new wave of research that estimates the effects of principals on student achievement, building on and extending prior research on teacher effects and public sector leader effects. As such, the report's findings have already been cited in even more recent research and its methodology and findings are used by other authors for validation and comparison.

The report begins by estimating principal fixed effects in a student achievement model, while acknowledging that the quality of these estimates is dependent in part on how well the model can account for “potential confounding factors including test measurement error and issues surrounding test construction” (p. 9) and that “other unmeasured factors will be correlated with the principal in each school and with the principal's effectiveness” (p. 9). The authors rightly assert that principals' influence may increase over time, although not in a linear manner, as they learn more about school operations and environment, and that in the process principals affect teacher hiring and retention. Their principal fixed-effects model estimates the overall standard deviation of principal's effectiveness as 0.207, which they describe as “unbelievably large” (p. 10-11). The authors acknowledge that these estimates may be due to other unexplained factors, including the relationship between principal experience and school poverty rates and test measurement issues, but do not explain why their estimates were so large.

The authors analyze differences in the estimated variance in principal effectiveness against differences in their schools' poverty rates and show that variance increases with poverty rates. They acknowledge that this relationship is difficult to explain, suggesting that less effective principals are more likely to be in high-poverty schools, or that their influence may be more adverse in these schools. The authors tried several adjustments—a statistical shrinkage estimator and reweighting scores to account for or eliminate differences in the distribution of initial achievement among principals' schools— but found that the adjustments yielded larger variance estimates and thus bias. The authors chose not to use reweighted data, but without further insight into the influence of unexplained factors.

To reduce inflated estimates of the variance in principal quality, the authors tested an alternative approach that “focuses on how the variance in achievement growth changes across student cohorts as new principals assume leadership in a school” (p. 16). They evaluated the validity of three hypotheses while controlling for student demographics: the effect of a principal does not change over time; principals are assigned randomly to schools; and principal turnover is unrelated to other school changes that affect achievement gains. Narrowing the focus to the variance associated with a change in principal and including school fixed effects reduces the magnitude of the principal value-added estimates to 0.11 s.d. Based on their model approaches, the authors concluded that “there is significant variation in principal effectiveness” (p. 20), this variation increases with school poverty rates, and that “the magnitude of the estimates falls substantially following the inclusion of school fixed effects” (p. 20) and when focusing on the change in principal. The report authors found differences in principal effects by school poverty rate

that appear not to be biased by other changes or inflated by observed demographic factors and other unobserved factors. The different principal effectiveness coefficients increase by the schools' poverty quartile from 0.08 to 0.14 s.d. (lowest to highest poverty rate).

Like other studies, this report looks at within-school comparisons of principal performance in order to determine variability in principal effects. It does not consider how the effectiveness of prior principals biases the comparison for more recent principals. It also does not clarify how restricted its sample is: the study is limited to schools with two or

The report does not take into account how districts actually assign principals to schools, the impact of school growth on the mobility of principals, and the influence of working conditions on principal transfer patterns.

more principals during the study period and to principals with at least three years of school achievement data. This represents a notable sample bias when considering the applicability of findings to all principals. The authors do note that the reduction in sample size was so severe that analyses using only newly hired principals in their first three years of leadership was not a fruitful path to pursue, but do not report the numbers. Chiang, Lipscomb and Gill (2012) argue that principal value-added estimates can only be created for principals with a transition during the analysis period and that principals new to the role should be excluded, rendering such analyses inappropriate for principal evaluation purposes because estimates could only be calculated for a small subset of principals, as noted above. Grissom, Kalgorides, and Loeb (2012) argue that under such conditions, the ability of the prior principal greatly influences the estimate of the effectiveness of the principal under question. More specifically, a principal with a true effectiveness of X could be considered effective if following an ineffective principal, or could be considered ineffective if following an effective principal. Such reliance on the idiosyncratic nature of placement after a previous principal renders the entire analysis inaccurate and not useful for evaluative purposes.

While the report's third model helps with variance estimate problems, the report does not adequately address timeframe issues related to effects. Other studies point out that principal effects and school outcomes lag and do not occur simultaneously, since principal performance actions in one year are not experienced as school effects until subsequent years.¹⁴ The report does point out, however, in a related analysis, that one direct influence principals have is over the quality of teachers hired or replaced. It will take several years for changes in teaching staff to yield positive (or negative) student achievement effects. As well, a principal's effect cannot be assumed to be immediate or constant over time, a problem that the report does not address. Moreover, other studies have found that principal effects decline when they change schools, suggesting the contribution of other unaccounted-for factors.¹⁵

Another study is far more cautious about the use of estimating principal effects (including the approach used in this report), concluding that “the inconsistencies and drawbacks of the measures lead to consideration of whether they should be used at all...The warning that comes from these analyses is that it is important to think carefully about what the measures are revealing about the specific contribution of the principal and to use the measures for what they are, not for as a clear indicator of principals’ specific contributions”.¹⁶ Unlike other studies, this report, while providing evidence of meaningful variation between ways its authors measure principal effects, does not look at the relationship among measures or with non-test based measures, as other researchers have done.¹⁷

The report acknowledges, as found in other large-scale principal effectiveness research, that principals have only an indirect impact on student test scores and that such analyses do not provide information about the principal behaviors associated with test score increases, although the authors do investigate one potentially direct principal effect: changing the composition of the teaching force within a school. The results showed mixed relationships among school poverty rates, principal effectiveness and teacher career patterns, with some relationship for high poverty schools and schools with the least effective principals. The report does not take into account whether principals had hiring, retention and removal authority, whether this authority was a district responsibility, and whether principal authority over these decisions and practices was consistent throughout the study period. Such authority is likely to vary by district size and wealth and thus not be evenly distributed. In addition, the report authors were not able to identify more and less effective teachers, only whether teachers from grades with lower value-added scores were more likely to leave than were other teachers in the same schools.

The report also presents results to understand “the extent and character of principal turnover” (p. 7), and makes comparisons by school poverty and mathematics achievement. The authors find that 30% of principals leave their school each year, a rate 50% higher than the national average of 20% (although it should be noted that the authors compared rates at two different time periods—1995-2001, in Texas, when employment opportunities were robust, and for 2009, nationally, when economic conditions were tighter and there were fewer non-principal employment opportunities). Similarly, the report does not take into account school-district policies and relationships when comparing principal transitions based on principal effectiveness and school poverty rates. The report’s interpretation that the high rate at which low-performing principals in high-poverty schools change schools is one-sided, by suggesting that districts just move these principals around, rather than screen them out. The report also does not take into account how districts actually assign principals to schools, the impact of school growth on the mobility of principals, and the influence of working conditions on principal transfer patterns.

Other research has found an unequal distribution of principals, with first-time principals often assigned to the most challenging schools. Principals generally prefer to work in schools with easier-to-serve students and more favorable working conditions. Consequently, there will be higher turnover in low-performing schools, based on school assignment preferences and the difficulty of being successful.¹⁸ Other research, using

another more recent state data set, identified several conceptual and methodological issues in investigating principal turnover and retention, related to choice (whether they can choose or are required to move based on district policy or fit), opportunity (which is not distributed equally within or among systems), role and place, that this report did not take into account.¹⁹

VII. Usefulness of the Report for Guidance of Policy and Practice

Taken as a whole, the report has limited utility to guide policy and practice. It serves as one of several recent efforts to develop value-added measures of principal effectiveness not conflated with school effectiveness measures and controlled sufficiently for unmeasured factors. The results confirm, as do other studies, that principals have a positive, independent influence achievement and that the size of this influence varies by school poverty rates. Methodological flaws, however, raise serious questions about the actual effect sizes principals might register on student test scores and thus the validity of such analyses.

The most important policy-relevant conclusion that can be derived from this report is that estimating principal effectiveness using student achievement data for performance evaluation decisions is simply not possible given current methodology and sample size restrictions. This conclusion is further confirmed by the other recently released studies of value-added indicators of principal effectiveness. All found that isolating principal effects requires controlling for school effects, which can only be done with principals who are new to the position who can be compared with prior principals of the same school (or vice versa), and having available test grade data. From other studies, this limitation makes value-added analyses applicable to less than half of all principals in the state. This has a profound effect on the multiple states that have adopted principal evaluation systems that include estimates of principal effectiveness relative to changes in student test scores.

Notes and References

1 Branch, G. F., Hanushek, E. A., & Rivkin, S. G. (2012). *Estimating the effect of leaders on public sector productivity: The case of school principals*. Cambridge, MA: National Bureau of Economic Research. Retrieved February 18, 2013, from <http://www.nber.org/papers/w17803>;

Branch, G. F., Hanushek, E. A., & Rivkin, S. G. (2013). School leaders matter. *Education Next*, 13(1), 63-69, Retrieved February 18, 2013, from <http://educationnext.org/school-leaders-matter/> .

2 Bertrand, M. & Schoar, A. (2003). Managing with style: The effect of managers on firm policies. *The Quarterly Journal of Economics*, 118(4), 1169-1208.

3Brewer, D. J. (1993). Principals and student outcomes: Evidence from U. S. high schools. *Economics of Education Review*, 12(4), 281-292;

Eberts, R. W., & Stone, J. A. (1988). Student achievement in public schools: Do principals make a difference? *Economics of Education Review*, 7(3), 291-299;

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools and academic achievement. *Econometrica*, 73(2), 417-458.

4Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago, Ill: University of Chicago Press;

Hallinger, P., & Heck, R. (1998). Exploring the principal's contributions to school effectiveness. *School Effectiveness and School Improvement*, 9(2), 157-191;

Leithwood, K., & Jantzi, D. (2008). Linking leadership to student learning: The contributions of leader efficacy. *Educational administration quarterly*, 44(4), 496-528;

Louis, K. S., Leithwood, K., Wahlstrom, K. L., & Anderson, S. E. (2010). *Investigating the links to improved student learning. Final report of research findings*. Minneapolis, MN: University of Minnesota;

Robinson, V. M. J., Lloyd, C. A., & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational administration quarterly*, 44(5), 635-674.

5 Brewer, D. J. (1993). Principals and student outcomes: Evidence from U. S. high schools. *Economics of Education Review*, 12(4), 281-292;

Cullen, J. B., & Mazzeo, M. J. (2008). *Implicit performance awards: an empirical analysis of the labor market for public school administrators*. San Diego, CA: University of California;

Eberts, R. W., & Stone, J. A. (1988). Student achievement in public schools: Do principals make a difference? *Economics of Education Review*, 7(3), 291-299.

6 Battle, D. (2010). *Principal attribution and mobility: Results from the 2008-09 principal follow-up survey*. Washington, DC: National Center for Education Statistics;

Gates, S. M., Ringel, J. S., Santibanez, L., Guarino, C., Ghosh-Dastidar, B., & Brown, A. (2006). Mobility and turnover among school principals. *Journal of policy analysis and management*, 25(3), 289-302.

7 Fuller, E. J., & Young, M. D. (2008). The revolving door: Principal turnover in Texas. *Texas Study of Secondary Education*, 17(2), 14-18.

8 Cullen, J. B., & Mazzeo, M. J. (2008). Implicit performance awards: an empirical analysis of the labor market for public school administrators. San Diego, CA: University of California;

Farley-Ripple, E. N., Solano, P. L., & McDuffie, M. J. (2012). Conceptual and methodological issues in research on school administrator career behavior. *Educational Researcher*, 41(6), 220-229;

Loeb, S., Kalogrides, D., & Horng, E. L. (2010). Principal preferences and the uneven distribution of principals across schools. *Educational evaluation and policy analysis*, 32(2), 205-229.

9 Chiang, H., Lipscomb, S., & Gill, B. (2012). *Is school value-added indicative of principal quality?* Cambridge, MA.: Mathematica Policy Research;

Dhuey, E., & Smith, J. (2012). *How school principals influence student learning*. Toronto, CA: University of Toronto.

This study estimated school and principal value added on independent samples (different grades, years and students) for both math and reading student test scores, using all principals who were involved in a leadership transition between some pair of successive years during the analysis period (50% of all principals) and for whom there were comparison principals in their school network (where a large enough one exists and where there were at least 20 students with assessment test scores) (6% of all principals who had served in Pennsylvania schools during the analysis period and 12% of the principals who had a leadership transition). The authors found within-network principal effects ranging from 0.8 to 0.14 student-level standard deviations, but that school value-added is not an accurate predictor of principal value-added estimates (or vice versa). They concluded that “school-effectiveness measures cannot disentangle principals’ contributions from unobserved, school-specific factors” (p. 25), including how principals are assigned to schools and the extent to which principals can “leave a lasting or widespread impact on their schools’ contributions to student learning in the first few years of their tenure” (p. 25) and whether excluding principals who are new to their role would improve the relationships.

Grissom, J. A., Kalogrides, D., & Loeb, S. (2012). Using student test scores to measure principal performance. Nashville, TN: Vanderbilt University.

This study similarly tried to model principal effects on student learning, as distinct from other school effects, testing out a school effectiveness model, a relative within-school effectiveness model, and a school improvement model. Using state math and reading assessment scores for students in grades 3-8 from a very large urban school district, the authors created principal value-added measures for 523 principals. They also collected district evaluation scores for the principals, school climate results and attendance data as comparative non-test-based performance measures. They criticize the school effectiveness model, which yielded standard deviations of 0.09-0.15 (with shrinkage), for only adjusting for student characteristics and not for factors outside the principals’ control. Their relative within-school effectiveness model restricted analysis to only those principals who can be compared with other principals of the school during the study period, and yielded estimates ranging from 0.03-0.08 s.d. The authors acknowledge that this method is problematic because it uses “comparison sets that can be tiny and, as a result, idiosyncratic” (p. 11) which is even more problematic when there are multiple principal turnovers in a study period, and because principal effectiveness is determined in part by the comparative quality of the prior principal’s effectiveness. The school improvement model must be measured over multiple years of a principal’s tenure (reducing the principal sample by one-third), which can aggregate measurement error and may

still capture contributions made by prior principals. Their analysis used four years of data and yielded standard deviations of 0.02-0.04, which are poorly correlated with non-test-based measures.

10 Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools and academic achievement. *Econometrica*, 73(2), 417-458.

11 Fuller, E. J., & Young, M. D. (2008). The revolving door: Principal turnover in Texas. *Texas Study of Secondary Education*, 17(2), 14-18.

12 Fuller, E., Young, M. D., & Baker, B. D. (2011). Do principal preparation programs influence student achievement through the building of teacher-team qualifications by the principal? An exploratory analysis. *Educational Administration Quarterly*, 47(1), 173-216.

13 Fuller, E. & Young, M. D. (2009). *Tenure and retention of newly hired principals in Texas*. Manuscript. Austin, TX: University of Texas at Austin.

14 Grissom, J. A., Kalogrides, D., & Loeb, S. (2012). *Using student test scores to measure principal performance*. Nashville, TN: Vanderbilt University.

15 Cannon, S., Figlio, D., & Sass, T. (2012). *Principal quality and the persistence of school policies*. Evanston, IL: Northwestern University.

16 Grissom, J. A., Kalogrides, D., & Loeb, S. (2012). *Using student test scores to measure principal performance*, 32. Nashville, TN: Vanderbilt University.

17 Grissom, J. A., Kalogrides, D., & Loeb, S. (2012). *Using student test scores to measure principal performance*. Nashville, TN: Vanderbilt University.

18 Loeb, S., Kalogrides, D., & Hornig, E. L. (2010). Principal preferences and the uneven distribution of principals across schools. *Educational evaluation and policy analysis*, 32(2), 205-229.

19 Farley-Ripple, E. N., Solano, P. L., & McDuffie, M. J. (2012). Conceptual and methodological issues in research on school administrator career behavior. *Educational Researcher*, 41(6), 220-229.

DOCUMENTS REVIEWED:

Estimating the effect of leaders on public sector productivity: The case of school principals, and

“School Leaders Matter” in *Education Next*

AUTHORS:

Gregory F. Branch, Eric A. Hanushek and Steven G. Rivkin

PUBLISHER/THINK TANKS:

National Bureau of Economic Research
Education Next

DOCUMENT RELEASE DATES:

2012 and 2013

REVIEW DATE:

March 5, 2013

REVIEWER:

Margaret Terry Orr, Bank Street College of Education

E-MAIL ADDRESS:

morr@bankstreet.edu

PHONE NUMBER:

212-875-4546

SUGGESTED CITATION:

Orr, M. T. (2013). *Review of “Estimating the effect of leaders on public sector productivity: The case of school principals,” and “School Leaders Matter.”* Boulder, CO: National Education Policy Center. Retrieved [date] from <http://nepc.colorado.edu/thinktank/review-estimating-effect-principals>.