How Much Does Leadership Matter? Evidence from Public Schools[1]

by Gregory F. Branch[a], Eric A. Hanushek[b,c], Steven G. Rivkin[c,d], Jeffrey C. Schiman[e],

and Xinyu Chen[d2]

November 2026

ABSTRACT

Although much has been written about the importance of leadership in the determination of organizational success, difficulties separating the impact of leaders from other organizational components complicate the estimation of their productivity and effectiveness. Panel data on students and educators plus regular measurement of school output offer a window into the importance of leadership quality. High dimensional two-way fixed effects methods have been used widely to separate worker and firm effects, but the small number of schools with multiple principals who have worked in multiple schools limits their use in the such methods. We use school fixed effect and matching methods to estimate directly the variation in the effectiveness of public-school principals in the state of Texas and the city of Chicago. Effective principals have a large and significant impact on student achievement that is remarkably similar across the different samples. Moreover, this effect is larger in schools with concentrations of disadvantaged students. We show that part of this impact comes through the effects of principal quality on the composition of exiting teachers. Finally, we combine the quantitative data with principal, teacher and student surveys to investigate the contributions of instructional leadership and data management skills to principal effectiveness. In this descriptive analysis, we find achievement growth to be strongly related to the principal's self-reported focus on instructional leadership, to teacher impressions of instructional leadership, and to student interest in school and perceptions of safety.

---

[1] The Chicago component of this project is part of a larger collaboration with the University of Chicago Consortium on School Research on the study of school leadership in the Chicago Public Schools. We have received extensive support around data preparation and district engagement and invaluable feedback on the analysis and interpretation of the findings. The Texas component of this project was made possible by the data and facilities at the Texas Schools Project of the University of Texas at Dallas and the UT Dallas Education Research Center. The study has been supported by the Packard Humanities Institute, the Smith Richardson Foundation, the Spencer Foundation, the Laura and John Arnold Foundations, CALDER, and the Hewlett Foundation. The conclusions of this research do not necessarily reflect the opinions or official position of the Texas Education Agency, the Texas Higher Education Coordinating Board, the State of Texas, or the Chicago Public Schools. We thank Susan Kick and Linda Shay for their guidance on the survey component of the study.

[2] a: University of Texas at Dallas; b: Stanford University; c: National Bureau of Economic Research; d: University of Illinois, Chicago; e: Georgia Southern University

I. Introduction

Leadership quality is frequently cited as key to organizational success across such diverse places as boardrooms, sports arenas, national legislatures, and schools. Yet it is often difficult to separate the impact of the leader from a variety of external factors affecting organization performance.[3] School principals potentially provide an exception, where extensive administrative data plus regular measurement of school output offer a window into the importance of leadership quality.

Differences in principal effectiveness are especially relevant for school policy, given the strong consensus about the importance of teachers for achievement.[4] If, say, principals play an important role in raising the quality of instruction – through either teacher selection or teacher development, successful school policy might start with the selection of a particularly skillful principal. On the other hand, if principals are relatively unimportant for school operations, personnel policies centered entirely on teachers may be preferred.

Although the importance of addressing issues of principal effectiveness has been recognized, aspects of the problem have precluded simple estimation approaches even as rich panel data on student achievement and school performance have become increasingly available.[5] Somewhat naturally, early evaluations of variations in principal effects mimicked the well-developed estimation of teacher value-added through estimation of principal value added, sometimes in specifications that included school fixed effects or in analyses that demeaned value-added estimates by the school average .[6]

---

[3] Considerable analysis has investigated the importance of leaders through descriptive analyses of private firms, concentrating on how firm outcomes – typically revenues or profits – are associated with CEOs and other top managers who have varying backgrounds or experiences. There is an older literature in this genre employing event studies on corporate returns after changes in management (e.g., Johnson, Magee, Nagarajan, and Newman (1985)). Other work has looked at the impact of management in family owned firms (see Bertrand (2009)) and at changes from firm founders (Bennedsen, Meisner Nielsen, Pérez-González, and Wolfenzon (2007)). Much of the research on CEO performance has been motivated by investigations of CEO salaries (see, for example, Englmaier, Illing, and Sadka (2009) and the related symposium).

[4] Hanushek and Rivkin (2010b, 2012), Bacher-Hicks and Koedel (2023).

[5] Although the role of the principal has been highlighted for a long time, the earliest discussion focused on anecdotal accounts (See e.g. Edmonds (1979), Purkey and Smith (1983) and Carter (2000)).

[6] Studies that include school and principal fixed effects include Dhuey and Smith (2014, 2018), Grissom, Kalogrides, and Loeb (2015), Chiang, Lipscomb, and Gill (2016), Bartanen (2020), and Grissom, Egalite,

A natural extension involves the estimation of models with both principal (worker) and school (firm) fixed effects that use school switchers to separate principal from school effects and measure the productivity of principals relative to others in a connected network. However, as Bartanen et al (2024) points out, the small number of schools with principals who have led multiple schools severely limits the use of this approach.[7] Importantly, the very limited number of schools with multiple principals who have worked in other schools also precludes adoption of the Kline et al (2020) approach to dealing with bias introduced by the small numbers of movers.

Moreover, in contrast to the estimation of teacher value added that permits the inclusion of school by year fixed effects, the presence of only a single principal in a school at any point in time complicates efforts to separate principal from time-varying school, neighborhood, and peer effects. Direct estimation of the variance in principal effectiveness based on within school differences in achievement gains under different principals provides an attractive alternative approach, and we extend the analysis of variance approaches used in Rivkin, Hanushek, and Kain (2005), Coelli and Green (2012), Laing et al (2016), Branch et al (2020) and Bartanen et al (2024) that compares achievement gain fluctuations or correlations across principal transitions to those within principal spells to identify the variance in principal effectiveness.

This approach requires a valid counterfactual estimate of what the variance or correlation would have been in the absence of the principal change, and we adopt two alternative approaches for the construction of such a counterfactual. The first follows Laing et al (2016) and Branch et al (2020) and uses long-serving principals in each school that experiences a transition under the assumption that year-to-year fluctuations in school and neighborhood factors remain similar over time. The second follows Bartanen et al (2024) and compares schools that experience principal transitions with those that do not. Importantly, we extend this approach by using matching methods to select a

---

and Lindsay (2021). The small samples of principal movers is, however, likely to lead to estimation biases (Bonhomme et al. (2023)).

[7] This is far lower than the numbers referred to as "low" in the work by Bonhomme et al. (2023) and Kline, Saggio, and Sølvsten (2020) that provide methods that address biases introduced by small numbers of movers in many firms (or schools for our purposes).

comparison for each school with a transition. Specifically, we match each transition school with a school with a long-serving principal that exhibits similarly sized year-to-year achievement gain differences within principal spells. Importantly, the large set of potential donor schools permits us to compare principals the same number of years apart in at the same points in the tenure profile and to restrict matches to comparison schools with observations in the same years as each transition school.

Our empirical analysis relies on administrative longitudinal data for the period 1993 to 2018 for the state of Texas and for the Chicago Public Schools. We focus on the much larger Texas sample that permits more disaggregated analyses at the same years of tenure, though a comparison of estimates from the same basic specifications provides evidence on the generalizability of the results. Importantly, the two samples permit alternative explorations into what drives differences in principal effectiveness. The extensive Texas data allow an investigation of the role of the principal in adjusting the quality of the school teaching force, while the Chicago data incorporate teacher and student survey information on principal behaviors and perceptions of school climate and engagement that can be related to statistical estimates of productivity.

We find significant variation in principal value added, as the within school and matching estimates are significant and very similar in magnitude. Specifications that restrict treatment and control comparisons to be three years apart show that a principal one standard deviation above the mean in effectiveness will on average boost student achievement by 0.05 standard deviations in both Texas and Chicago. Restricting both the outgoing and incoming principals to be in their second year of tenure, the first observation for matched comparison principals to be in the second year of tenure and squared differences calculated alternately over observations three or four years apart increases the magnitude to slightly more than 0.06 standard deviations

Although these estimates are roughly half the magnitude at most of typical estimates of teacher effects (Hanushek and Rivkin (2010b)), they actually indicate a substantial impact of leadership because they apply to an entire school and not just a single classroom. When we separate Texas schools by the poverty rate, we find that the

variance in principal effectiveness increases monotonically with poverty. We cannot, however, determine whether leadership is simply more important in schools with concentrations of economically disadvantaged students or whether there are larger differences in skills of those leading high poverty schools.

We next explore potential channels through which more effective principals raise achievement. First, we use principals who lead multiple schools to investigate whether there is more negative selection of exiting teachers in schools led by more effective principals. By using value added measured in one school to investigate the link between principal effectiveness and the character of teacher exits in a second school we circumvent problems introduced by time-varying school factors and mechanical relationships between the value added of exiting teachers and of the principal. We find stronger negative selection of teacher exits in schools led by more effective principals, supporting the belief that management of the stock of teachers is an important channel for principals to affect the quality of instruction.

Second, we combine CPS surveys of teachers, students and principals with measures of principal productivity to extend existing work on school leadership practices that compares achievement with measures of practices and skills (Bloom, Lemos, Sadun, and Van Reenen (2015)).[8] The panel data enable us to account for unobserved heterogeneity with prior achievement and for unobserved school differences with fixed effects, thereby controlling for primary determinants of achievement that may be related to the survey measures.

We find a strong relationship between a principal's self-reported focus on instructional leadership and the school's value-added to achievement. Additionally, the estimates show significant, positive relationships between achievement growth on the one hand and teacher impressions of instructional leadership and student interest in school

[8] Bloom and Van Reenen (2007, 2011) and Bloom, Lemos, Sadun, and Van Reenen (2015) relate organization outcomes in different contexts including schools with self-reported manager skills and practices. The analysis of school management in Bloom, Lemos, Sadun, and Van Reenen (2015) finds a positive relationship between pupil achievement and the management index, controlling for enrollment, the pupil-teacher ratio, school admissions policies, and demographic characteristics. The limited controls for student and school heterogeneity, however, raise the possibility that other factors drive the associations between achievement and management skills.

and perceptions of school safety on the other. This evidence is consistent with the belief that instructional leadership and school safety are additional channels through which principals affect the quality of instruction.[9]

II. Identifying the Impact of School Principals

Estimation of two-dimensional worker and firm fixed effects models would appear to be a natural approach to the identification of principal value added, but the structure of the market for school leaders severely limits its usefulness in this context. First, less than 10 percent of Texas schools have principals who have worked in multiple schools, and the connected network of schools that have multiple principals who have worked in more than one school as required to implement the Kline et al (2020) approach to mitigating bias introduced by the low number of movers is very small. Second, the presence of only a single principal in a school at any point in time, frequent curricular, financial, and regulatory changes in the public-school sector, and demographic shifts amplify the limitations of this approach. Although the presence of only a single principal avoids the potential biases introduced by nonrandom assignment of students into classrooms, it removes a straightforward approach of accounting for time-varying school and neighborhood influences with school by year fixed effects.[10]

We therefore extend the analysis of variance approach used in Rivkin, Hanushek, and Kain (2005) for teachers and Coelli and Green (2012), Laing et al (2016), Branch et al (2020) and Bartanen et al (2024) for principals to identify the variance in principal value added. The following conceptual framework drawn from Rivkin and Shakeel (2026) highlights the key estimation issues including unobserved student heterogeneity,

[9]See also Grissom, Kalogrides, and Loeb (2015), and Chiang, Lipscomb, and Gill (2016).

[10] A widely-discussed problem in the estimation of teacher value-added is the purposeful classroom allocation of students, but that is not a major concern here because of the focus on aggregate school outcomes instead of performance outcomes across classrooms in the same school. Similarly, the much larger number of test-takers in schools as opposed to classrooms mitigates effects of test measurement errors that have been discussed in the teacher value-added literature. See, for example, the investigations in Kane, McCaffrey, Miller, and Staiger (2013), Chetty, Friedman, and Rockoff (2014), and Rothstein (2010) on the presence and magnitude of biases introduced by nonrandom assignment to classrooms. For larger discussions of issues specific to the estimation of teacher value-added, see Hanushek and Rivkin (2012) and Koedel, Mihaly, and Rockoff (2015).

fixed and time-varying school effects, and the dynamics of principal influences on the quality of instruction and achievement.

A. Conceptual framework

A basic value-added model for principals with achievement gain on the left hand side illustrates the key issues for estimation. In Eq. 1, achievement gain (ΔA) for student i attending school s with principal j in year Y is a function of a principal effect specific to school s (θ), demographic and program characteristics (X), and a random error:

$$\Delta A_{isjY} = \beta X_{iY} + \theta_{js} + \varepsilon_{isjY} \quad (1)$$

Although estimation in this form has been common,[11] it is susceptible to bias introduced by unobserved school factors. If higher value-added principals tend to work in higher value-added schools, this approach would on average lead to overestimates of the value-added of more effective principals and underestimates of that for less effective principals, thus leading to overestimation of the variation in principal effectiveness. Importantly, even if principal and school quality are uncorrelated, the estimated variance of principal effectiveness calculated from the variance of $\hat{\theta}$ will be biased because it will include a component of the variation in school quality.

More generally, this fixed effect formulation ignores important dynamics of principal and school effects. To fix ideas, it is convenient to consider a more general formulation that incorporates important elements hypothesized in previous analyses of principal effects. Eq. 2 permits principal effectiveness to evolve over time along with time-varying school factors, allows for potential match effects between principal and school, and recognizes that prior principal actions can affect current school quality through various personnel decisions, teacher development, and establishment of school climate and practices:

[11] See, for example, Branch, Hanushek, and Rivkin (2012), Grissom and Loeb (2011), Bartanen, Husain, Liebowitz, and Rogers (2024).

$$\Delta A_{isjY} = \beta X_{iY} + \theta_{jsY} + \sum_{y=Y-n}^{Y-1} \lambda^{Y-y} \theta_{lsy} + \pi_{sY} + \varepsilon_{isjY} \quad (2)$$

where λ is the annual rate of persistence of principal effects in the n prior years and π_{sY} captures both fixed and time-varying school and neighborhood factors outside of the principal's control. The subscript l indexes the principal in a prior year y: l denotes the same principal as j in the absence of a transition between year y and year Y and a different principal if there is a transition.

The persistent influences of principal effects in prior years, represented by the sum of principal l effects each of the y years earlier multiplied by λ^{Y-y}, captures the impacts a principal exerts on school quality in subsequent years through personnel decisions and other channels.[12] Although many principal actions affect only the current year and staff turnover lessens persistence, some principal decisions will affect the quality of instruction in subsequent years. As Gilraine and McCarthy (forthcoming) highlights in the case of teachers, less than full persistence likely elevates value added estimates for teachers whose students had more effective teachers in the prior year relative to those with less effective teachers if the fadeout rate is higher for teacher produced knowledge than for knowledge gain from other sources. The mechanism linking prior and current principal value-added estimates differs from that for teachers; in the case of principals it is the persistence of principal actions including teacher hiring that connects prior year and current year value added. Consequently, higher prior principal value added will tend to inflate estimates of current principal value added because prior principals affect the current quality of instruction and not just the magnitude of lagged achievement. As we discuss below, this both frames the interpretation of variance

[12] This possibility raises complications not faced in the estimation of teacher value-added where controlling for lagged-test scores likely captures most of prior teacher's effects. Persistence of principal effects on future school quality is distinct from depreciation of effects on student knowledge. The latter refers to teacher or school effects on knowledge that may depreciate over time. The persistence of principal effects on future school quality refers to the effect of current decisions or actions on skill acquisition at the school in the future.

estimates and supports comparisons of achievement gains for outgoing and incoming principals the same number of years apart and at the same point in the tenure profile.

B. Variance of principal value added

Direct estimation of the variance offers an attractive alternative to a two-step approach that requires estimation of principal value added. We begin with a simplified version of Eq. 2 that assumes principal effects are fixed throughout their tenure at school s and that they fully depreciate each year ($i.e., \theta_{j,s,y} = \theta_{js}$ and $\lambda = 0$). Eq. 3 focuses on the squared difference in the school average achievement gain ($\overline{\Delta A}$) between years Y and $Y-k$:

$$\left(\overline{\Delta A}_{s,Y} - \overline{\Delta A}_{s,Y-k}\right)^2 = \beta\left(\bar{X}_{s,Y} - \bar{X}_{s,Y-k}\right)^2 + \left(\theta_{j,s,Y} - \theta_{l,s,Y-k}\right)^2 + \left(\pi_{s,Y} - \pi_{s,Y-k}\right)^2 + V_{s,Y,Y-k} \quad (3)$$

All covariances across the components and the variance of the error are included in V.

Expansion of the squared difference in principal and school effect terms illustrates the key idea underlying the use of principal transitions to identify the variance in principal effectiveness.

$$\left(\overline{\Delta A}_{s,Y} - \overline{\Delta A}_{s,Y-k}\right)^2 = \beta\left(\bar{X}_{s,Y} - \bar{X}_{s,Y-k}\right)^2 + \left[\theta_j^2 + \theta_l^2 - 2\theta_j\,\theta_l\right] + \left[\pi_{s,Y}^2 + \pi_{s,Y-k}^2 - 2\pi_{s,Y}\,\pi_{s,Y-k}\right] + V_{s,Y,Y-k} \quad (4)$$

Under the assumption of constant principal effectiveness, $J = 1$ in the absence of a principal transition, meaning that $2\theta_j\,\theta_l = \theta_j^2 + \theta_l^2$ and the principal term in brackets drops out. If principals in each school are independently drawn from a common distribution and the school has a transition, $2\theta_j\,\theta_m = 0$, and the term in brackets becomes $2\theta_j^2$—that is, twice the variance in principal effectiveness.

If the squared difference in school effects are the same on average for schools with and those without transitions, the squared difference in achievement gain within principal spells would provide a valid counterfactual for what the squared difference in outcomes across principal transitions would have been had the principal not changed. This key assumption cannot be tested directly, but we can restrict comparisons to those

schools more likely to satisfy the conditions for a valid counterfactual, compare estimates for different control groups that are susceptible to the influences of different confounding factors, and examine the sensitivity to the inclusion of time-varying student characteristics.

Relaxing the assumption that prior principal effects on school quality depreciate fully each year adds the terms $\lambda^{2(Y-(Y-y))}\left(\theta_{j,s,Y-y} - \theta_{l,s,Y-k-y}\right)^2$ for each of the y years prior to year Y.[13] With less than full persistence, the prior principal will contribute more to school quality in the current principal's second year than in their third year. In addition, the cumulative effects of principals will evolve with tenure as the influences of the prior principal wane.

The waning influence of the prior principal as current-principal tenure increases supports an additional estimation restriction suggested by Bartanen et al (2024): principals should be compared at the same point in their tenure profile. Because of the small number of schools with observations y years apart for two different principals at the same point in the tenure profile AND with observations y years apart for a long-serving principal, we rely on the matching approach for this component of the analysis.

Miller (2013) also illustrates an additional dynamic component: a substantial decline in average achievement in the year prior to a transition. This leads us to account directly for the turbulence around transitions in many of the specifications described below.

C. Empirical Specifications

The empirical analysis considers a series of regressions of the squared difference in outcomes calculated across a fixed number of years on an indicator for a principal transition and in some specifications the squared difference in demographic variables.[14] The counterfactuals for the baseline models come from either long-serving principals in

[13] Note that the magnitude of the squared difference in persistent principal effects in prior years decreases substantially with y, as $\lambda^{2(Y-(Y-y))}$ quickly becomes quite small. For example, with a 50 percent rate of persistence of principal actions, the value of $\lambda^{2(Y-(Y-y))}$ equals 0.0625 for y=2 and 0.0156 for y=3.

[14] Altonji, Elder, and Taber (2005) describe an informal approach to measuring the bias from unobserved factors on the basis of the sensitivity of the results to the inclusion of observed characteristics. See also Oster (2019).

the same school (school fixed effects models) or a long-serving principal from a matched comparison school.

The matching approach matches each school with a principal transition with the closest comparison school in terms of the residual squared difference in gains two years apart within a principal spell, controlling for year. We match on squared within-principal achievement-gain differences because unobserved, time varying school influences constitute the greatest threat to identification. The nearest match to each school that has a principal transition provides the counterfactual estimate of what the squared difference in gains would have been in the absence of the principal change. Note that the donor pool for each specification includes principals from schools without transitions who serve long enough to match the transition school in terms of distance between observations plus at least one more year.

Transition and comparison schools in the matching models are matched on gains between observations two years apart under the same principal. To avoid introducing mechanical correlations, the observations used to estimate the variance exclude the observations used to select the matched comparisons.

The main outcome is the squared difference in adjusted average achievement gain under the assumption of 30 percent knowledge depreciation.[15] Achievement gain in year t therefore equals test score in year t minus 0.7 times test score in year $t-1$. Notice that this refers to the standard depreciation of teacher and school effects and not the previously discussed persistence of the effects of principals' decisions or actions on future school quality. We also present estimates from squared difference in gains models that assume no knowledge depreciation.

[15] For a discussion of the potential biases from analyzing the dynamics with no knowledge depreciation, see Meghir and Rivkin (2011). In our sensitivity analysis, we assume an overall depreciation rate of 0.3, which corresponds to commonly estimated values (Hanushek and Rivkin (2012). Jacob, Lefgren, and Sims (2010) provide evidence of a larger depreciation rate for teacher effects.

III. Texas and Chicago Data

Estimation of the variance in principal quality requires long panel data sets of matched school personnel and students to extract the systematic input of principals from other influences on achievement gains. With more than twenty years of data for both Texas and Chicago we can observe multiple principals in the vast majority of schools. In Texas, the availability of linkages between classroom teachers and their students facilitates causal analysis of how principal quality affects the average quality of teachers who leave their schools. In Chicago, teacher and student survey data permit the construction of principal and school quality indexes along multiple dimensions while principal survey data yield self-reported measures of practices and skills. Both provide information about the channels through which principals affect achievement.

Because of the very large sample size and number of principal transitions, the Texas sample provides the primary data set for the analysis of overall principal effectiveness. Variance estimates for the Chicago Public Schools sample can be compared to those for the state of Texas, but the availability of teacher and student survey data that enable comparisons of principal fixed effects and survey-based principal quality measures provides the primary purpose for the Chicago Public Schools sample.

A. Administrative Data from Texas

The Texas database was assembled at the Texas Schools Project of the University of Texas at Dallas from school administrative data developed for the Texas Education Agency (TEA). The analytic panel data sets include linked information on all Texas public school teachers, administrators, staff, and students in each year since 1993.

The Public Education Information Management System (PEIMS), Texas's statewide educational database, reports key demographic and program data including race, ethnicity, subsidized lunch eligibility, and special education classification. The student and teacher data can be merged by school-grade-year throughout the sample period and by classroom teacher of record beginning in 2012.

The sample period is divided among three state standardized test regimes, and the test data can be merged with the PEIMS data by student and year. Beginning in 1993, the

Texas Assessment of Academic Skills (TAAS) was administered each spring to eligible students enrolled in grades 3 through 8.[16] In 2003 the state substituted the *Texas Assessment of Knowledge and Skills* (TAKS) in place of the TAAS, and in 2012 the *State of Texas Assessment of Academic Readiness* (STAAR) replaced the TAKS. We transform all test results into standardized scores with a mean of zero and variance equal to one for each subject, grade, and year, meaning that our achievement measures describe students by their relative position in the overall state performance distribution.[17]

One of the strengths of this sample is the large number of principals and schools that are observed. Over the sample period from 1996 to 2018 used in the direct variance analysis, we observe 20,389 unique principals and 119,006 principal-year observations. The sample size facilitates analyses of heterogeneity by school-poverty rate and within school estimators that account for fixed differences across schools. The teacher turnover estimation focuses on the STAR period between 2013 and 2018 when student-teacher matches become available, though estimates of principal effectiveness used in this analysis come from the TAAS and TAKS period as well.

B. Administrative Data from Chicago Public Schools

Chicago is the third largest school district in the U.S. with enrollment fluctuating around 400,000 students over the past quarter century. [18] There were 585 schools in 2016 (although we focus on elementary schools that consistently report achievement data).

The analysis of CPS principals uses longitudinal administrative data for the period 1994 to 2018, which include test scores, program information, and demographics for all students and extensive information on staff. In the spring of 1994 through 2005, students

[16] Many special education and limited English proficient students are exempted from the tests. In each year roughly 15 percent of students do not take the tests, either because of an exemption or because of repeated absences on testing days.

[17] Given the ordinal nature of standardized scores and focus of the tests on less advanced skills, comparisons of schools with students concentrated at different points on the achievement distribution may miss differences in school quality and principal effectiveness in terms of gains in knowledge in an absolute sense. Preliminary work (not shown) indicated that a much more flexible parameterization of prior achievement had virtually no effect on the estimates in specifications focusing on within-school differences with controls for student and school-average demographics.

[18] Chicago enrollment was 408,830 in 1990, rises to 435, 261 in 2000, and falls to 378,199 in 2016 (U.S. Department of Education (2018)).

in grades 3 through 8 were tested using the *Iowa Test of Basic Skills* (ITBS). In 2006, the *Illinois Standards Achievement Test* (ISAT) test replaced the ITBS, and in 2015 tests from the *MAP Suite of NWEA* replaced the ISAT.[19] As in Texas, we transform CPS test scores to have a mean of zero and a standard deviation of one within each subject, grade, and year. Importantly, standard deviations on the National Assessment of Educational Progress (NAEP) for Chicago and Texas are quite similar, indicating that it is reasonable to compare estimates across settings directly (Ost, Gangopadhyaya, and Schiman (2017)).

Principals must be linked over time by name, as there are no unique ID numbers that span the period. We take great care to account for name changes following marriage or divorce, other name changes, and spelling or punctuation inconsistencies. We merge principals by school and year to the student data. The principal panel contains almost 1,700 principals in over 600 schools.

Survey responses to questions about principals (from teacher surveys), schools and instruction (from student surveys) and skills and practices (from principal surveys) form the final components of the data. The surveys were administered every other year between 1995 and 2011 and annually starting in 2012. Principal surveys contain information on own practices, though questions vary considerably from year to year. Based on the findings in Bloom, Lemos, Sadun, and Van Reenen (2015), we focus on the 2013 survey which asked a variety of questions about the use of assessment data to measure performance (e.g. how much do you use standardized test scores to gauge performance or to what extent do you use test data to evaluate progress of subgroups of students) and instructional leadership practices (e.g. when working with teachers, how often do you observe teacher instruction in the classroom). These various survey responses provide information on the practices and skills related to principal effectiveness.

[19] https://www.nwea.org/

IV. The variance of principal effectiveness

Direct estimation of the variance in principal effectiveness relies on changes in student outcomes that are associated with principal transitions, and Table 1 shows the distributions of transitions among Texas schools, by poverty quartile. Most schools experience multiple principal transitions over the sample period, with a mean number of transitions of slightly more than two. The distribution of principal transitions by poverty quartile of the student population shows substantially less turnover in the lowest poverty schools. Schools in the lowest quartile average 1.77 principals over the period 1996-2011, while those in the upper two quartiles average 2.24 principals with little difference among these.

A. Baseline estimates for Texas

Table 2 compares within-school and matching specifications, where the sources of any biases differ substantially across the estimators. The bottom panel presents two matching estimators, one that matches over all potential comparison schools and one that matches only over donor pools for each school where the transition and comparison school observations come from the same years. A comparison of results in these two panels thus illustrates the sensitivity of the estimates to the comparison group designation and the likelihood that unobserved, time-varying school or neighborhood factors confound the estimates. The school fixed effects specifications require at least one of the two principals to have at least four consecutive years in the school, so we drop all transition observations without at least one such principal. This reduces the number of transitions in the sample to 9,174.[20] Note that Appendix Table a2 shows that the estimates

[20] The principal with at least four years of data is used in the matching process unless both principals serve at least four years during the sample period, in which case we use the chronologically first principal. When we use the second principal in the matching process, we take the squared difference between achievement gain in the first and third years, while when we use the first principal, we take the squared difference in gain between the penultimate year at the school and two years earlier. Importantly, the observations used to identify a matched comparison school never contribute to the estimates of the variance in principal effectiveness.

are insensitive to the assumption of a 30 percent knowledge depreciation rate and fixing the difference between observations at three years.

The estimates for the school fixed effects (upper panel) and the unrestricted matching (upper row of lower panel) models are all significant at the 1 percent level and almost identical in all specifications regardless of whether specifications include indicators for first and last spell years or demographic characteristics. The second row of the lower panel reports coefficients from the restricted matching model, and the coefficients are quite similar to the corresponding estimates for the unrestricted matching and school fixed effects models and significant at the 1 percent level. The insensitivity of the matching estimates to the year restriction holds despite the fact that it substantially decreases the average quality of the match, increasing the mean distance between the squared difference in gains in the transition and matched school by a factor of 10 (Appendix Table a1).

The pattern of estimates strongly supports the belief that the transition coefficients identify the variance of principal effectiveness. Not only do comparison groups that are susceptible to different confounding factors produce quite similar estimates, the controls for turbulence around transitions and fluctuations in demographic characteristics increase slightly rather than decrease the magnitude of the estimate.

These estimated standard deviation of principal effectiveness of slightly less than 0.05 is less than half as large as the typical range of estimates of the variance in teacher effectiveness summarized in Hanushek and Rivkin (2010b). Yet unlike a teacher who primarily influences just the students in their classes, a principal affects all children in the school. Average enrollment in Texas elementary and middle schools is 563 and 735, respectively. Therefore, even a one standard deviation increase in the quality of school leadership of 0.04 to 0.05 standard deviations of achievement has an educationally significant effect on cognitive skills.

B. Fixed tenure comparisons

The specifications that produced the estimates in Table 2 hold constant the distance between observations in the transition and counterfactual schools but allow

principal tenure to vary. If the impacts of principals persist into subsequent leadership regimes, this potentially muddies the interpretation of the transition estimates because the influences of the previous principal are related to principal tenure which may differ for the pre- and post- transition principals. The following estimates compare principals in their second year, where the observations are either three or four years apart so that the second year is not the final year of the spell. This requires long-serving principals in the comparison schools to serve at least six or seven year spells to ensure that the second observation is not the final year at the school.

The upper panel of Table 3 reports matching estimates based on squared difference in gains three years apart for two principals in the second year in the school; we focus here on the matching estimators because of the small sample of schools with these transitions and a long-serving principal. Estimates from both the unrestricted and restricted matching models are highly significant regardless of whether the specification includes first and last year controls or demographic characteristics. The magnitudes are slightly larger than those that combine principals with different tenure levels, translating to a one standard deviation difference in principal effectives equal to roughly 0.06 standard deviations. Despite the fact that the year restriction increases the average distance between transition and matched comparison schools by roughly a factor of 10, it actually leads to slightly smaller standard errors.

The lower panel of Table 3 reports estimates for observations in the second year of tenure that are four rather than three years apart. These estimates are somewhat larger than those in the top panel and also highly significant. The implied standard deviation of principal effectiveness is roughly 15 to 20 percent larger with the sample of observations four years apart. Again, the year restriction has little effect on the estimates.

C. Heterogeneity by school poverty

The large Texas sample of principals and the diversity of school composition permit an investigation of heterogeneity across districts by the student poverty rate. Table 4 reports school fixed effect estimates of the variance in principal effectiveness separately by quartile of the school poverty rate based on specifications that combine tenure levels

and have gains differenced across three years, a full set of indicators for first and final years of spells, and knowledge depreciation of 30 percent.[21] The estimated variance in principal quality increases monotonically with the poverty rate and is significant for all but the lowest poverty quartile. By far the largest jump occurs for the highest poverty schools, where a one standard deviation increase in principal effectiveness is associated with an average student achievement increase of 0.09 standard deviations, almost double the size found in the other three quartiles.

We cannot determine whether the quality of leadership is more important in high poverty schools or whether the variance of underlying principal skills of those willing to lead one of those schools is larger. The sample of principals who lead schools in different poverty categories is too small to measure with any degree of precision how the effectiveness of principals might change if they move across poverty quartiles. In any case, these findings highlight the much greater importance of hiring an effective school leader in a high poverty school than in a school with fewer impoverished children.

D. Principal Effectiveness in Chicago

The primary value from incorporating the Chicago Public Schools (CPS) comes from the unique survey data that allow investigation of underlying management skills presented below, but it is also useful to compare results for a large midwest urban district to those for the state of Texas. The Texas analysis considers the range of effectiveness across some 1,200 districts and more than 9,000 schools in a state that does not allow collective bargaining for teachers or principals. In contrast, CPS has both unionized teachers and principals along with a single human resource and hiring authority and an extensive regulatory structure.

Because of the much smaller number of schools and relatively long panel, the CPS analysis uses the school fixed effects models and combines tenure levels. Similar to the baseline estimates for Texas, there are significant differences in effectiveness among

[21] Results for school fixed effect models with no knowledge depreciation and matching models produce the same pattern at quite similar magnitudes.

CPS principals that are quite similar across specifications (Table 5). Thus, even in the much-restricted CPS labor market for principals, statistically and educationally significant variation in principal productivity emerge that are only modestly smaller than that found for Texas principals.

V. The Role of Teacher Selection

Given the compelling evidence that teacher effectiveness is a primary determinant of the quality of instruction, it is natural to investigate the linkages of principal and teacher effectiveness. Principals can affect instruction and learning through a combination of personnel decisions and instructional leadership. Estimating the influence of the principal on the effectiveness of incumbent teachers is a particularly difficult analytical task (Although we provide some descriptive evidence in the next section). It is, however, possible to provide direct causal evidence about the interaction of principal effectiveness and the quality composition of teachers. Specifically, we investigate the effect of principal quality on teacher selection by looking at the quality of exiting teachers relative to their peers who remain in the school.[22] Although teacher selection may enter more broadly if some principals hire more effectively than others, the absence of quality measures for new teachers and the fact that many principals have little control over new hires lead us to focus on the composition of leavers.[23]

A challenge in analyzing teacher selection is that estimates of principal effectiveness in part aggregate value-added of teachers to the school level. This introduces a mechanical correlation between principal quality and teacher quality that complicates identification of principal quality on the effectiveness distribution of teacher exits. Even random departures would introduce a spurious negative relationship between

[22] Teacher turnover has been a significant concern in past research (see Hanushek, Kain, and Rivkin (2004), Boyd, Lankford, Loeb, and Wyckoff (2005). Ronfeldt, Loeb, and Wyckoff (2013), Hanushek, Rivkin, and Schiman (2016), and Adnot, Dee, Katz, and Wyckoff (2017)). Studies on the quality dimension of teacher turnover include Hanushek and Rivkin (2010a), Goldhaber, Gross, and Player (2011), and Hanushek, Rivkin, and Schiman (2016). None, however, focus on the role of the principal in this turnover.

[23] Grissom and Bartanen (2019) pursue a similar approach except that it relies on subjective evaluations of principal effectiveness and does not account fully for time-varying school factors.

principal quality and teacher exits if the years following the teacher transitions are included in the sample used to estimate principal effectiveness. Consider, for example, a school where a group of more effective teachers happened to exit a school for reasons unrelated to the principal. Average teacher quality and therefore estimated principal effectiveness would be lower in the subsequent period. The opposite would hold if less-effective teachers were to disproportionately exit from a school; in both cases these dynamics would introduce a spurious negative relationship between estimates of principal quality and the average effectiveness of exiting teachers.

To purge the estimates of this mechanical connection, we focus on principals who lead multiple schools, allowing us to estimate principal effectiveness in their first school (s) and the relationship between principal effectiveness and the character of teacher transitions in their subsequent school (s').[24] We employ a two-stage estimation procedure over the subsample of Texas principals who lead multiple schools by first estimating value-added in their initial school and then using this measure to estimate the effect of principal quality on the relative effectiveness of departing teachers in their second school.[25]

We estimate principal value added during the second year of tenure both to make the estimates comparable across principals and to address concerns related to the dynamics around principal transitions discussed previously. The value-added specification regresses math achievement of student i in school s with principal j in year t on prior math achievement entered as a cubic polynomial, student demographic controls

[24] An alternative would be to use the initial year or two that a principal is at a school to measure principal effectiveness and then subsequent years to analyze teacher departures. We do not use this approach because of the turbulence around principal transitions that makes it undesirable to include achievement in the first year of a principal's tenure in the estimation of principal effectiveness. Using only the second year at a school would preclude first year in charge, which is likely to be the period in which the principal exerts the most influence on teacher composition.

[25] Note that this analysis of teacher value-added requires student-teacher matches that are available only beginning in 2012 under the STAR testing regime. Principal effectiveness can be estimated over the longer period that includes the TAKS tests.

(X), school characteristics (Z) including peer measures comprised of campus-year-grade averages of X, a year fixed effect τ, and a principal fixed effect ρ.[26]

$$A_{ist} = f(A_{i,t-1}) + \beta X_{ist} + \varphi Z_{st} + \rho_{js} + \tau_t + \epsilon_{ist} \quad (5)$$

As previously discussed, the estimate of principal quality, $\widehat{\rho_{js}}$, is equivalent to a school-by-year fixed effect, meaning that other factors outside of the control of the principal will also influence the estimate. In addition, changes over time in principal skills or the match quality with the school environment will introduce additional noise, attenuating the estimated effect of principal quality on the character of teacher transitions. Finally, principals are not randomly placed in schools, and endogenous sorting might introduce bias. However, the focus on the interaction with teacher effectiveness and the inclusion of school-by-year fixed effects in the teacher transition specification mitigate concerns about selection bias. Note that teacher effectiveness is measured by average math achievement growth for all students with valid current and prior grade test scores taught by teacher k.[27]

For the second stage, Equation 6 models the probability teacher k in school s' in year t' exits the school as a function of teacher effectiveness ($T_{ks't'}$), the interaction between teacher and principal effectiveness, ($T_{ks't'} * \widehat{\rho_{js,s't}}$), average characteristics of students taught by teacher $k(Z)$, a dummy variable equal to one if principal effectiveness is estimated under a test regime other than STAR ($TAKS_{ks't'}$), separate indicators for teachers with 0, 1 and 2 years of prior experience, an indicator for female teachers, a school-year fixed effect, (μ), and an error term (υ),

$$Exit_{ks't'} = \alpha T_{ks't'} + \phi\left(T_{ks't'} * \widehat{\rho_{js,s't'}}\right) + \varphi Z_{ks't'} + \zeta TAKS_{js,s't'} + \vartheta novice_{kt'} + \alpha female_{kt'} + \mu_{s't'} + \upsilon_{ks't'} \quad (6)$$

[26] Also included are mobility indicators to identify students who are in their first year at the school and students in the lowest grade at a school to account for the effects of structural moves between levels of schooling.

[27] Although classroom factors contribute to average math achievement growth, a leave out estimator that mitigates their influence is not feasible in this setting. Not only is this the only year of data available for first year teachers who do not return for a second year, but it will also be the only year in which the new principal has not influenced the composition of the teaching staff.

Robust standard errors are clustered by principal.

If higher quality principals are more likely to retain effective teachers and to move out ineffective teachers, we would expect $\phi < 0$. Importantly, the inclusion of school-by-year fixed effects fully accounts for all school-level influences on teacher transitions in that year that may be associated with principal quality. Only within-school variation in teacher effectiveness at a point in time is used to identify the relationship between principal quality and the pattern of teacher exits.[28]

Table 6 reports estimates of impact of the interaction between principal effectiveness and teacher- average achievement growth on the probability a teacher exits, by whether the regression is weighted by the number of test-takers. Regardless of whether the regression is weighted, ϕ is negative and significant, consistent with the hypothesis that skill at managing the stock of teachers is one determinant of overall principal effectiveness. The much larger interaction coefficient in the weighted regressions is consistent with principals' placing greater weight on average student growth when the teachers have more students. In terms of magnitude, the point estimate of -0.83 suggests that a one standard deviation movement up the principal effectiveness distribution increases the probability that a low achievement-growth teacher (one standard deviation below the mean) exits by around 5 percentage points.[29]

It is informative to compare these results to the association between principal effectiveness and the overall level of teacher turnover ignoring teacher effectiveness. In contrast to the findings in Table 7, Figure 1 reveals little relationship between the teacher transition rate and principal effectiveness. This is consistent with the belief that some turnover is productive.

[28] Note that the direct impact of principal quality on the probability a teacher exits is not identified because principal quality does not vary within a school and year.

[29] Note, however, that this is not a clear estimate of the general equilibrium effect, since the exiting teacher may simply move to another school.

VI. Management Practices and Achievement Growth

Using survey responses of school managers in eight countries, Bloom, Lemos, Sadun, and Van Reenen (2015) describe the relationships between student outcomes and management practices in public and private schools. Their rigorous survey methodology provides information about differences in school management within and across countries, and the empirical analysis reveals a positive relationship between achievement on the one hand and manager skills and use of practices considered to be effective on the other.[30] However, the lack of controls for prior achievement and limited information on students and schools raise questions about the influences of student and school heterogeneity on the estimated effects of management skills and practices. By combining the principal, teacher and student survey information with longitudinal achievement data on Chicago students, we can extend this line of research with much richer controls for underlying differences in students and schools.

CPS principal surveys contain information on each principal's practices, though questions vary considerably from year to year.[31] We focus on the 2013 survey because principals were asked a variety of questions about the use of assessment data to measure performance and instructional leadership practices. Appendix Tables A3 (use of assessment data for evaluation) and A4 (use of instructional leadership practices) report the questions used to assess principal management along with descriptive information on responses. These surveys show, for example, that virtually all principals report some use of assessment data for evaluative and support purposes but that there is considerable variation in the actual usage patterns.

Teacher and student surveys contain questions related to leadership along with questions on various aspects of school climate. We focus on survey responses that appear particularly relevant to principal effectiveness, re-appear in virtually all years of the data, and are reasonably objective in nature to lessen the contribution of the principal's

[30] These include practices elucidated in Fryer (2014) that were found to have a causal effect on achievement.

[31] Measuring principal management practices through surveys of the principals themselves was an innovation of Bloom, Lemos, Sadun, and Van Reenen (2015) and their earlier management studies in other areas.

popularity to the variation in responses.[32] Teachers are asked the extent to which their principal sets clear expectations, communicates a vision, understands student learning, sets high learning standards, implements professional development, and tracks student achievement. Students are asked the extent to which they feel safe in the classroom and the hallways, the extent to which they find the curriculum interesting and challenging, and whether they work hard to do their best. The distributions of student and teacher survey responses are reported in Appendix Table A5.

Although estimates of principal value added likely conflate principal and school effects, they come much closer than average achievement to indexing principal effectiveness. Table 8 shows that average principal value-added increases monotonically as student and teacher responses become more positive about principal leadership and school climate. The patterns are remarkably similar for student responses regarding safety and their level of engagement with the material and for teacher responses regarding instructional leadership and focus on achievement. The differences in average principal value added between the most positive and most negative responses are around 0.02 standard deviations for the student responses and somewhat larger for the teacher responses, particularly for questions related to the focus on achievement. By comparison, the pattern in Table 9 for the principal self-survey responses is far less pronounced, particularly for the responses under use of assessment in performance measurement.

The overall value-added comparisons are consistent with a positive association between leadership skills and certain management practices, but unobserved school differences, other principal practices, and shocks associated with the final or first years in a school may have substantial influence on the patterns in Tables 7 and 8. To account for these differences and to dampen the noise associated with responses to each question, we produce a series of indexes based on the responses and estimate a variant of Equation 5 in which one or more of the indexes replaces the principal-school fixed effects:

$$A_{ijst} = f(A_{it-1}) + \kappa S_{jst} + \beta X_{ist} + \varphi Z_{st} + \mu_s^* + \tau_t + \mu_{ist} \qquad (6)$$

where $f(A_{it-1})$ is a cubic in prior achievement, S is a vector of indices of skills and

[32] We thank Principals Susan Kick and Linda Shay for their guidance on item selection.

practices from the surveys for principal j in school s, student controls (X_{ist}), school controls (Z_{st}), μ_s^* is a school fixed effect that is included in specifications based on teacher and student responses, and τ_t is a year fixed effect. Student controls include indicators for new student, special education, race, sex, grade repeater, and an index of socioeconomic status, while school controls also include school-grade-year averages of the student characteristics. Some specifications also include indicators for the first or last years of spells in a school. The indexes are created using polychoric factor analysis which accounts for the categorical nature of the survey responses and are scaled such that higher index values correspond to more frequent use of a practice or more favorable perception of the principal or school climate. The availability of only one year of principal surveys precludes the inclusion of school fixed effects in specifications that include indexes constructed from principal responses. Further, to provide a direct comparison with Bloom, Lemos, Sadun, and Van Reenen (2015), we estimate commensurate models that exclude the control for prior achievement.

Table 9 reveals strong, highly significant associations between achievement and all three student and teacher indices in specifications that include prior achievement and school fixed effects regardless of whether the specifications also include indicators for first or last year of a spell (Columns 2 and 3). Removal of the school fixed effects leads to little change in the coefficient on the teacher leadership index but substantial changes to the coefficients on the student interest and safety indexes. The decrease in the magnitude and significance of the coefficient on the interest index and increase in magnitude of the coefficient on school safety index moving from Column 2 to Column 1 suggest the presence of unobserved school differences that affect perceptions about academic engagement and school safety and are related to achievement growth. When neither prior achievement nor school fixed effects are included, the estimated coefficients are much larger, and there is a surprisingly negative estimate in the case of the student interest index.

On balance, the findings in the preferred specifications that account for prior achievement and school fixed effects are consistent with instructional leadership and the

establishment of a safe and engaging climate being channels through which principals raise achievement. Though not causal, these estimates strongly suggest that more effective principals tend to be perceived as better instructional leaders and tend to produce safer and more engaging school environments. The sensitivity of the estimates to the inclusion of prior achievement and school fixed effects highlights the value of the longitudinal data that contain measures of prior achievement and focus on principal comparisons within schools through school fixed effects.

The self-reported survey data from principals provide another window into differences in principal effectiveness. Table 10 shows the relationship of the single index of principal-reported skills to achievement, by controls. Columns 1-3 show a positive and highly significant relationship between achievement and the overall index after controlling for prior student achievement. Notice that the inclusion of the teacher and student indexes from Table 9 increase the magnitude of the estimate. In contrast, the estimates in Columns 4-6 based on specifications that do not control for prior achievement are uniformly smaller and not significant. The fact that the control for prior achievement strengthens rather than weakens the relationship provides added support for a conclusion that better leadership practices and the use of data to measure performance are associated with higher school quality and achievement.

Overall, these results support the conclusions in Bloom, Lemos, Sadun, and Van Reenen (2015) that self-reported principal practices capture skill and practice differences related to effectiveness and the quality of instruction. The analysis here, however, expands on the previous evidence with additional information on instructional leadership practices and school climate and much more extensive controls for student and school differences. Rather than reducing the magnitudes and significance of the coefficients, the controls for student and school differences tend to bolster the findings and the interpretation that they are channels through which principals affect the quality of instruction and achievement.

VII. Conclusions

Although unobserved student, school, and neighborhood influences complicate the estimation of the variance of principal value added to achievement, the sum total of results in this study support the belief that school leadership quality makes a substantial contribution to achievement, especially in high-poverty schools. First is the finding of significant variation in principal effectiveness where the magnitude is similar whether or not tenure is held fixed, whether counterfactuals come from long-serving principals in the same school or matched comparison schools, whether the matching donor pools are restricted to comparisons during the same years as the transition schools, and whether there are controls for demographic shifts and the turbulence around transitions. Second is the finding of more adverse teacher selection out of schools led by more effective principals. And third is the strong association between estimates of principal value added on the one hand, and teacher perceptions of the principal as an instructional leader and student perceptions of the safety and engagement of the school.

Even with the constraints on public sector operations and the diffuse performance incentives, principals significantly affect achievement. It is unlikely that this finding is limited to the education sector, and from a policy perspective improvement in the selection, development, incentives and retention of high-quality public sector leaders would appear to have a high payoff.

References

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High wage workers and high wage firms." *Econometrica* 67, no. 2 (March): 251-333.

Adnot, Melinda, Thomas Dee, Veronica Katz, and James Wyckoff. 2017. "Teacher Turnover, Teacher Quality, and Student Achievement in DCPS." *Educational Evaluation and Policy Analysis* 39, no. 1: 54-76.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools." *Journal of Political Economy* 113, no. 1: 151-184.

Bacher-Hicks, Andrew, and Cory Koedel. 2023. "Estimation and interpretation of teacher value added in research applications." In *Handbook of the Economics of Education, Vol 6*, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann: Elsevier: 93-134.

Bartanen, Brendan, A. 2020. Principal Quality and Student Attendance. *Educational Researcher* 49(2): 101–113.

Bartanen, Brendan, Aliza N. Husain, and David D. Liebowitz. 2024. "Rethinking principal effects on student outcomes." *Journal of Public Economics* 234(June): 105115.

Bartanen, Brendan, Aliza N. Husain, David D. Liebowitz, and Laura K. Rogers. 2024. "The Returns to Experience for School Principals." EdWorkingPaper No. 24-978. Brown University: Annenberg (June).

Bennedsen, Morten, Kasper Meisner Nielsen, Francisco Pérez-González, and Daniel Wolfenzon. 2007. "Inside the Family Firm: the Role of Families in Succession Decisions and Performance." *Quarterly Journal of Economics* 122, no. 2: 647-691.

Bertrand, Marianne. 2009. "CEOs." *Annual Review of Economics* 1: 121-149.

Bloom, Nicholas, Renata Lemos, Raffaella Sadun, and John Van Reenen. 2015. "Does management matter in schools?" *Economic Journal* 125, no. 584 (May): 647-674.

Bloom, Nicholas, Renata Lemos, Raffaella Sadun, and John Van Reenen. 2020. "Healthy Business? Managerial Education and Management in Health Care." *The Review of Economics and Statistics* 102, no. 3: 506-517.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices across Firms and Countries." *The Quarterly Journal of Economics* 122, no. 4 (November): 1351-1408.

Bloom, Nicholas, and John Van Reenen. 2011. "Human Resource Management and Productivity." In *Handbook of Labor Economics*, edited by Ashenfelter Orley and Card David: Elsevier: 1697-1767.

Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler. 2023. "How Much Should We Trust Estimates of Firm Effects and Worker Sorting?" *Journal of Labor Economics* 41, no. 2 (April): 291-322.

Boyd, Don, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2005. "The draw of home: How teachers' preferences for proximity disadvantage urban schools." *Journal of Policy Analysis and Management* 24, no. 1 (Winter): 113-132.

Branch, Gregory F., Eric A. Hanushek, and Steven G. Rivkin. 2012. "Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals." NBER Working Paper W17803. Cambridge, MA: National Bureau of Economic Research (January).

Carter, Samuel Casey. 2000. *No Excuses: Lessons from 21 High-Performing, High-Poverty Schools*. Washington, D.C.: The heritage Foundation.

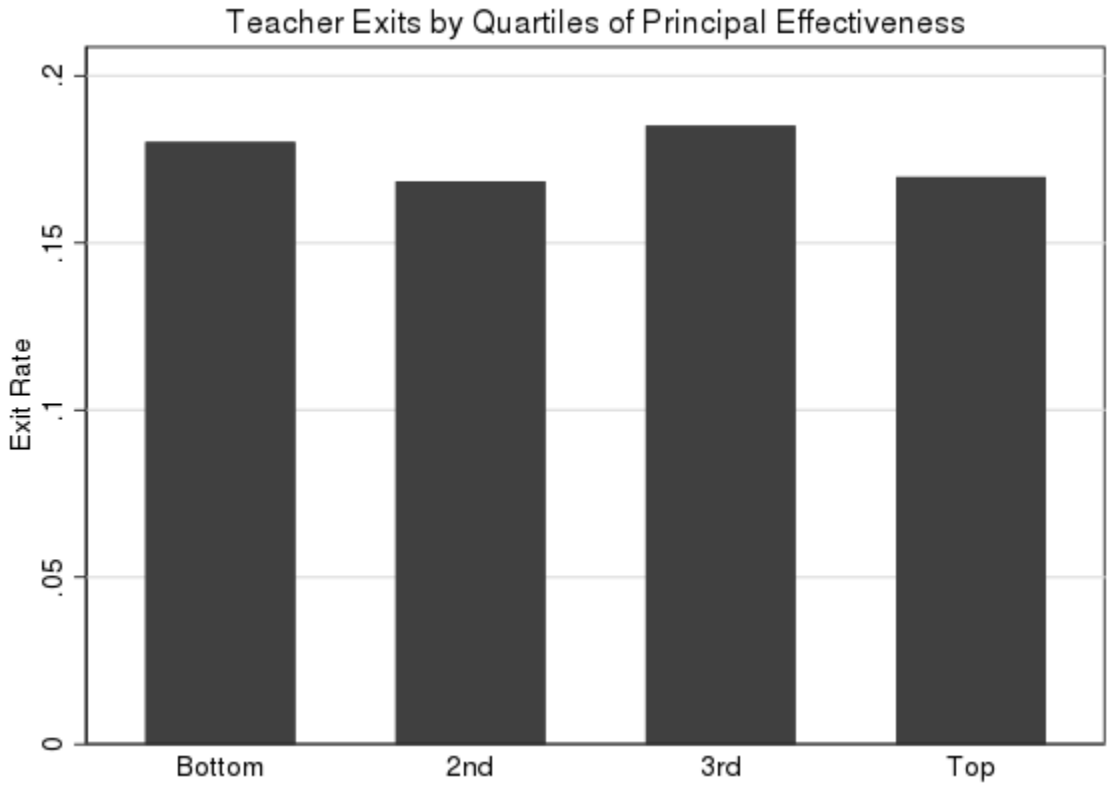
Chetty, Raj, John N. Friedman, and Jonah Rockoff. 2014. "Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates." *American Economic Review* 104, no. 9 (September): 2593-2632.

Chiang, Hanley, Stephen Lipscomb, and Brian Gill. 2016. "Is School Value Added Indicative of Principal Quality?" *Education Finance and Policy* 11, no. 3 (Summer): 283–309.

Coelli, Michael, and David A. Green. 2012. "Leadership effects: school principals and student outcomes." *Economics of Education Review* 31, no. 1 (February): 92-109.
Dhuey, Elizabeth, and Justin Smith. 2014. "How important are school principals in the production of student achievement?" *Canadian Journal of Economics/Revue canadienne d'économique* 47, no. 2 (May): 634-663.
Dhuey, Elizabeth, and Justin Smith. 2018. "How school principals influence student learning." *Empirical Economics* 54, no. 2: 851-882.
Edmonds, Ronald R. 1979. "Effective schools for the urban poor." *Educational Leadership* 37, no. 1 (October): 15-24.
Englmaier, Florian, Gerhard Illing, and Efraim Sadka. 2009. "Introduction to the Symposium on Executive Pay." *CESifo Economic Studies* 55, no. 3/4 (September/December): 399-404.
Fryer, Roland G., Jr. 2014. "Injecting Charter School Best Practices into Traditional Public Schools: Evidence from Field Experiments." *Quarterly Journal of Economics* 129, no. 3 (August): 1355–1407.
Goldhaber, Dan, Betheny Gross, and Daniel Player. 2011. "Teacher career paths, teacher quality, and persistence in the classroom: Are public schools keeping their best?" *Journal of Policy Analysis and Management* 30, no. 1 (Winter): 57-87.
Grissom, Jason A., and Brendan Bartanen. 2019. "Strategic Retention: Principal Effectiveness and Teacher Turnover in Multiple-Measure Teacher Evaluation Systems." *American Educational Research Journal* 56, no. 2: 514-555.
Grissom, Jason A., Anna J. Egalite, and Constance A. Lindsay. 2021. *How Principals Affect Students and Schools: A Systematic Synthesis of Two Decades of Research*: The Wallace Foundation (February).
Grissom, Jason A., Demetra Kalogrides, and Susanna Loeb. 2015. "Using Student Test Scores to Measure Principal Performance." *Educational Evaluation and Policy Analysis* 37, no. 1 (March): 3-28.
Grissom, Jason A., and Susanna Loeb. 2011. "Triangulating Principal Effectiveness: How Perspectives of Parents, Teachers, and Assistant Principals Identify the Central Importance of Managerial Skills." *American Educational Research Journal* 48, no. 5 (October): 1091-1123.
Hanushek, Eric A., John F. Kain, and Steve G. Rivkin. 2004. "Why public schools lose teachers." *Journal of Human Resources* 39, no. 2 (Spring): 326-354.
Hanushek, Eric A., and Steven G. Rivkin. 2010a. "Constrained Job Matching: Does Teacher Job Search Harm Disadvantaged Urban Schools?" NBER w15816. Cambridge, MA: National Bureau of Economic Research (March).
Hanushek, Eric A., and Steven G. Rivkin. 2010b. "Generalizations about using value-added measures of teacher quality." *American Economic Review* 100, no. 2 (May): 267-271.
Hanushek, Eric A., and Steven G. Rivkin. 2012. "The distribution of teacher quality and implications for policy." *Annual Review of Economics* 4: 131-157.
Hanushek, Eric A., Steven G. Rivkin, and Jeffrey C. Schiman. 2016. "Dynamic effects of teacher turnover on the quality of instruction." *Economics of Education Review* 55(December): 132-148.

Jacob, Brian A., Lars Lefgren, and David Sims. 2010. "The persistence of teacher-induced learning." *Journal of Human Resources* 45, no. 4 (Fall): 915-943.
Johnson, Bruce, Robert P. Magee, Nandu J. Nagarajan, and Harry A. Newman. 1985. "An analysis of the stock price reaction to sudden executive death: Implications for the management labor market." *Journal of Accounting and Economics* 7, no. 1-3: 151-174.
Kane, Thomas J., Daniel F. McCaffrey, Trey Miller, and Douglas O. Staiger. 2013. *Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment*. MET Project: Bill and Melinda Gates Foundation (January).
Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten. 2020. "Leave-Out Estimation of Variance Components." *Econometrica* 88, no. 5 (September): 1859-1898.
Koedel, Cory, Kata Mihaly, and Jonah E. Rockoff. 2015. "Value-added modeling: A review." *Economics of Education Review* 47: 180-195.
Laing, Derek, Steven G. Rivkin, Jeffrey C. Schiman, and Jason Ward. 2016. "Decentralized Governance and the Quality of School Leadership." NBER Working Paper No. 22061. Cambridge, MA: National Bureau of Economic Research (March).
Meghir, Costas, and Steven Rivkin. 2011. "Econometric Methods for Research in Education." In *Handbook of the Economics of Education*, edited by Stephen Machin Eric A. Hanushek and Ludger Woessmann: Elsevier: 1-87.
Miller, Ashley. 2013. "Principal turnover and student achievement." *Economics of Education Review* 36(October): 60-72.
Ost, Ben, Anuj Gangopadhyaya, and Jeffrey C. Schiman. 2017. "Comparing standard deviation effects across contexts." *Education Economics* 25, no. 3 (May): 251-265.
Oster, Emily. 2019. "Unobservable selection and coefficient stability: Theory and evidence." *Journal of Business & Economic Statistics* 37, no. 2 (April): 187-204.
Purkey, Stewart C., and Marshall S. Smith. 1983. "Effective Schools: A Review." *The Elementary School Journal* 83, no. 4 (March): 427-452.
Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, schools, and academic achievement." *Econometrica* 73, no. 2 (March): 417-458.
Ronfeldt, Mathew, Susanna Loeb, and James Wyckoff. 2013. "How Teacher Turnover Harms Student Achievement." *American Educational Research Journal* 50, no. 1 (February): 4–36.
Rothstein, Jesse. 2010. "Teacher quality in educational production: Tracking, decay, and student achievement." *Quarterly Journal of Economics* 125, no. 1 (February): 175-214.
Schiman, Jeffrey C., Derek Laing, Steven G. Rivkin, and Jason Ward. 2018. "Decentralized Governance and the Quality of School Leadership." (mimeo) Chicago: University of Illinois Chicago.
U.S. Department of Education. 2018. *Digest of Education Statistics, 2017*. Advanced Release ed. Washington, DC: National Center for Education Statistics.

Figure 1. Teacher Exit Rates by Quartile of Principal Value-Added



Notes: Exit rates are defined by the ratio of the number of teachers exiting to the total number of teachers, averaged across the period of study (academic years 2014-2017) and principal effectiveness quartile. Exit rates shown are for the schools and years included in the teacher exit regression sample.

Table 1. Distribution of Principal Transitions across Chicago Public Schools and the State of Texas, Overall and By Poverty Quartile in Texas

	Distribution by Number of Transitions						Expected transitions	Number of schools
	0	1	2	3	4	5+		
Chicago Public Schools	0.024	0.197	0.232	0.230	0.150	0.167	2.76	706
State of Texas								
Overall	0.154	0.230	0.238	0.184	0.119	0.072	2.1	6084
1st (Lowest) Poverty Quartile	0.195	0.287	0.240	0.153	0.079	0.044	1.77	1472
2nd Poverty Quartile	0.158	0.218	0.226	0.196	0.128	0.073	2.14	1538
3rd Poverty Quartile	0.140	0.212	0.234	0.186	0.138	0.090	2.24	1541
4th (Highest) Poverty Quartile	0.126	0.206	0.255	0.203	0.128	0.082	2.24	1533

Notes: Distributions are based on the entire sample, spanning 1994 to 2018 in Chicago and 1993 to 2018 in Texas. Within each school, we count the number of principal transitions that occurred during the sample period. The distribution of transitions is reported for the overall sample of schools in the CPS and both overall and by quartile of poverty in Texas.

Table 2. School fixed effect and matching model estimates of transition effects on the squared difference in achievement gains for observations 3 years apart

	(1)	(2)	(3)
Panel A. School fixed effects	0.0048***	0.0049***	0.0037***
	(0.0014)	(0.0014)	(0.0008)
Implied standard deviation	0.049	0.049	0.043
Number of school-year observations	45,327	45,327	45,327
Panel B. Matching			
1. Unrestricted matching model	0.0046***	0.0047***	0.0044***
	(0.0011)	(0.0011)	(0.0012)
Implied standard deviation	0.048	0.048	0.047
2. Restricted matching model	0.0042***	0.0044***	0.0043***
	(0.0011)	(0.0011)	(0.0011)
Implied standard deviation	0.046	0.047	0.046
Number of school-year observations	18,348	18,348	18,348
Indicators for first and last years of spells	Yes	Yes	No
Demographic controls	Yes	No	No

Notes: The coefficients come from regressions of the squared difference in depreciated test score gain between years Y and Y-3 on an indicator for a principal transition. The transition indicator equals one if there is a new principal in year Y relative to Y-3. The school fixed effects specification includes all transition events from the matching specifications and all continuous leadership counterfactuals. Matching is based on squared differences in achievement gains, measured either between the first and third years for incoming principals, or between the penultimate year and two years prior for outgoing principals, requiring a minimum of four years of data within a principal's tenure and defaulting to the outgoing principal when both are eligible. Restricted specifications match exactly on the years over which we compute the squared differences in gains. Unrestricted specifications match on residualized squared differences that first partial out calendar year effects. All regressions include year fixed effects. The student demographic controls are the squared difference in proportions black, Hispanic, female, special education and new entrants to the school, school enrollment, and share eligible for a subsidized lunch in multiple years. Standard errors clustered by school are in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table 3. Matching model estimates of transition effects on the squared difference in achievement gains for transition observations with two years of tenure, by number of years between observations

	(1)	(2)	(3)
Panel A. 2nd year of tenure, 3 years apart			
1. Unrestricted matching model	0.0060**	0.0066**	0.0050**
	(0.0030)	(0.0030)	(0.0025)
Implied standard deviation	0.055	0.057	0.050
2. Restricted matching model	0.0070***	0.0071***	0.0056**
	(0.0027)	(0.0027)	(0.0023)
Implied standard deviation	0.059	0.060	0.053
Number of school-year observations	2,316	2,316	2,316
Panel B. 2nd year of tenure, 4 years apart			
1. Unrestricted matching model	0.0099***	0.0097***	0.0094***
	(0.0038)	(0.0037)	(0.0030)
Implied standard deviation	0.070	0.070	0.069
2. Restricted matching model	0.0094**	0.0099***	0.0086***
	(0.0039)	(0.0038)	(0.0030)
Implied standard deviation	0.069	0.070	0.066
Number of school-year observations	1,684	1,684	1,684
Indicators for first and last years of spells	Yes	Yes	No
Demographic controls	Yes	No	No

Table 4. School fixed effect estimates of transition effects on the squared difference in achievement gains for observations 3 years apart, by Poverty Quartile

	Poverty Quartile			
	1st (Lowest)	2nd	3rd	4th (Highest)
Coefficient on new principal	0.0005	0.0032**	0.0056**	0.0155***
indicator	(0.001)	(0.0016)	(0.0023)	(0.0057)
Implied standard deviation	0.016	0.040	0.053	0.088
Number of school-year observations	10,703	11,164	11,068	12,203

Note: See notes to Table 2. Separate estimates are computed by poverty quartile.

Table 5. School fixed effect estimates of transition effects on the squared difference in achievement gains for the Chicago Public Schools

	(1)	(2)	(3)	(4)
Coefficient on new principal indicator	0.0033**	0.0039***	0.0039***	0.0038***
	(0.0013)	(0.0014)	(0.0014)	(0.0012)
Implied standard deviation	0.041	0.044	0.044	0.044
Number of school-year observations	9,836	8,491	8,491	8,491
Number of years between observations	1	3	3	3
Indicators for first and last years of spells	No	Yes	Yes	No
Demographic controls	Yes	Yes	No	No

Notes: The coefficients come from regressions of the squared difference in test score gain between years *t* and *t-k* on an indicator for a principal transition, assuming 30% depreciation. n is the number of years between observations, and the principal transition indicator equals one if there is a new principal in year *t* relative to *t-n.* All regressions include school fixed effects and year fixed effects. The student demographic controls are the squared difference in proportions black, Hispanic, female, classified as special education and new entrants to the school, in school enrollment, and in average parental SES. Standard errors clustered by school are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 6. Estimated effect of principal quality on the selectivity of teacher exits

	(1)	(2)
teacher effectiveness	-0.20	-0.10
	(0.183)	(0.235)
principal effectiveness*teacher effectiveness	-0.43*	-0.83**
	(0.230)	(0.276)
Teacher-years (N)	1,373	1,373
Campus-years (groups)	322	322
Regression weighted by number of test-takers	No	Yes

Note: The coefficients come from linear probability models of the probability a teacher exits on classroom average depreciated achievement gain and that gain interacted with an estimate of principal value-added in their previous school. All specifications include campus-by-year fixed effects, an indicator that the principal value-added estimate came from TAKS data, separate indicators for teachers with 0, 1 , and 2 years of prior experience, and an indicator for a female teacher. Standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table 7. Average Principal Value-Added by Student and Teacher Survey Items

	(1)	(2)	(3)	(4)
	Student Surveys			
How safe do you feel…	Not Safe	Somewhat Safe	Mostly Safe	Very Safe
In your class	-0.0103	-0.0063	-0.0020	0.0072
In the hallways/bathrooms	-0.0105	-0.0057	-0.0005	0.0109
How much do you agree	Strongly Disagree	Disagree	Agree	Strongly Agree
The topics we are studying are interesting and challenging	0.0027	0.0028	0.0047	0.0106
I work hard to do my best in this class	-0.0074	-0.0046	0.0020	0.0102
	Teacher Surveys			
The principal at this school	Strongly Disagree	Disagree	Agree	Strongly Agree
Makes teaching expectations clear	-0.0157	-0.0107	0.0007	0.0149
Communicates clear vision for school	-0.0157	-0.0111	0.0021	0.0152
Understands how students learn	-0.0112	-0.0006	0.0033	0.0131
Sets high standards for student learning	-0.0261	-0.0196	-0.0006	0.0168
Presses teachers to implement professional development	-0.0192	-0.0020	0.0015	0.0146
Tracks student academic progress	-0.0260	-0.0112	0.0012	0.0180

Notes: Student and teacher survey data from Chicago Public Schools covering years in which surveys were administered – 1997, 1999, 2001, 2003, 2005, 2007, 2009, 2011, 2012, 2013, 2014, and 2015.

Table 8. Average Principal Value-added by Principal Survey Items

Use of Assessment Data in Performance Measurement	not at all	a little	somewhat	very much
1. How much do you use each of the following to gauge how well your school is performing?				
a. results from the ISAT test	0.0279	0.0014	0.0101	0.0150
b. results from the NWEA test	n/a	0.1290	0.0346	0.0070
c. student grades	-0.0342	-0.0111	0.0288	-0.0040
d. benchmark assessments designed by your teachers	0.0490	-0.0249	0.0093	0.0227
e. classroom observations	-0.2001	0.0872	-0.0064	0.0169
2. To what extent do you use assessment data to				
a. evaluate the progress of individual students	n/a	-0.1378	0.0208	0.0124
b. evaluate the progress of subgroups of students	n/a	-0.1089	0.0109	0.0222
c. compare performance among classrooms in your school	-0.1105	-0.0238	-0.0142	0.0273
d. identify teachers who could serve as instructional leaders	-0.0749	0.0251	0.0101	0.0154
3. To what extent do you use assessment data to				
a. evaluate teacher performance	-0.1011	0.0128	-0.0019	0.0278
b. inform and communicate with parents	n/a	0.0034	0.0205	0.0087
c. assess schoolwide instructional programs	n/a	0.0286	0.0265	0.0069
d. monitor changes in your schools test scores over time	n/a	0.0002	0.0135	0.0136

Table 8 (cont.) Average Principal Value-added by Principal Survey Items

Instructional leadership practices	Never	a few times per year	a few times per month	1-2 days per week	>2 days per week
1. In general, when working directly with teachers this year, how often did you do each of the following:					
a. Demonstrate instructional practices in a classroom	0.0043	0.0177	-0.0003	0.0406	0.0047
b. Observe teacher instruction in a classroom	n/a	-0.0210	-0.0029	0.0051	0.0208
c. Advise teachers about classroom practices	n/a	-0.0212	0.0048	-0.0005	0.0294
d. Examine what students were working on during a teacher's lesson	0.0232	0.0021	0.0123	-0.0082	0.0311
e. Review student assessment data with your teachers	-0.2000	0.0165	-0.0057	0.0238	0.0381
f. Meet with teachers to give feedback on their instruction.	-0.2001	0.0266	0.0096	-0.0078	0.0436
	Never	Once a year	A few times a year	A few times a semester	Monthly
2. How often do you review assessment data (of any type)?					
a. I review assessment data independently	n/a	-0.2000	-0.0300	-0.0014	0.0059
b. I review assessment data with other administrators in my school	0.3986	-0.2000	-0.0003	-0.0088	0.0049
c. I review assessment data with other administrators in the district	0.0262	0.0032	0.0090	0.0243	-0.0049
d. I review assessment data with network staff	-0.0004	0.0137	-0.0138	0.0297	0.0244
e. I review assessment data with teachers	-0.2001	n/a	-0.0284	0.0226	-0.0042
f. I review assessment data with groups of teachers	-0.2000		-0.0523	0.0131	0.0124

Notes: Chicago Public Schools survey data are for 2013

Table 9. Relationship between achievement and teacher and student rating indexes and principal management indexes

	Controlling for prior student achievement			Not controlling for prior student achievement		
	(1)	(2)	(3)	(4)	(5)	(6)
Student safety index	0.036***	0.019***	0.019***	0.140***	0.056***	0.055***
	(0.004)	(0.003)	(0.003)	(0.012)	(0.006)	(0.006)
Student interest index	0.0022	0.013***	0.013***	-0.060***	0.0011	0.00079
	(0.0028)	(0.003)	(0.003)	(0.010)	(0.0042)	(0.0042)
Teacher leadership impression index	0.015***	0.014***	0.014***	0.038***	0.028***	0.027***
	(0.002)	(0.002)	(0.002)	(0.006)	(0.004)	(0.003)
School fixed effect	N	Y	Y	N	Y	Y
Indicators for first and last years of spells	N	N	Y	N	N	Y
Observations	1,086,081	1,086,081	1,086,081	1,488,497	1,488,497	1,488,497

Notes. Coefficients come from student-level regressions of achievement on indexes base on student and teacher responses to the surveys. The indexes are based on responses to the questions in Table 6 and are standardized to have a mean of zero and standard deviation of one. Each regression controls for student-level and school-grade-year averages of student race, sex, special education, enrollment, share new students, and parental SES as well as enrollment. Regressions including prior achievement control for a cubic function of lagged achievement. Standard errors clustered by principal spell are in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.001$

Table 10. Relationship between achievement and principal management, student, and teacher indexes

	Controlling for prior student achievement			Not controlling for prior student achievement		
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Combined Principal Index					
Combined Principal Management Index	0.0297**	0.0360***	0.0354***	0.0059	0.0200	0.0199
	(0.0131)	(0.0134)	(0.0131)	(0.0232)	(0.0216)	(0.0216)
Observations	51,279	49,684	49,684	67,717	64,919	64,919
Includes teacher and student indexes	N	Y	Y	N	Y	Y
Indicators for first and last years of spells	N	N	Y	N	N	Y

Notes: Coefficients come from student-level regressions of achievement on indexes based on principal responses about their practices. The indexes are based on responses to the questions in Table 7 and are standardized to have a mean of zero and standard deviation of one. The estimates are based on the 2013 survey that included principal practice measures. Each regression controls for student level and school-grade-year averages of student race, sex, special education, enrollment, share new students, and parental SES. Regressions for value-added control for a cubic function of lagged achievement. Robust standard errors are in parentheses.
* p<0.10, ** p<0.05, *** p<0.001

Appendix Table A1. Distribution of absolute distances between the squared difference in gains for the first principal in the transition school and in the comparison schools, by tenure and years between observations

	p10	p50	p90	mean	N
No Tenure Restriction					
Unrestricted matching	0.000000	0.000002	0.000021	0.000034	9,174
Restricted matching	0.000001	0.000014	0.000240	0.000405	
2nd year of tenure, 3 years apart					
Unrestricted matching	0.000001	0.000007	0.000107	0.000116	1,158
Restricted matching	0.000004	0.000061	0.000998	0.000930	
2nd year of tenure, 4 years apart					
Unrestricted matching	0.000001	0.000011	0.000210	0.000344	842
restricted matching	0.000004	0.000082	0.001790	0.002394	

Appendix Table A2. Estimated Variation in Texas Principal Effectiveness, by assumed rate of annual student knowledge depreciation

	(1)	(2)	(3)	(4)
Panel A. 30 % student achievement depreciation				
Coefficient on transition indicator	0.0046***	0.0053***	0.0053***	0.0041***
	(0.0010)	(0.0014)	(0.0014)	(0.0008)
Implied standard deviation	0.048	0.052	0.046	0.045
Panel B. No student achievement depreciation				
Coefficient on transition indicator	0.0050***	0.0056***	0.0056***	0.0032***
	(0.0012)	(0.0018)	(0.0017)	(0.001)
Implied standard deviation	0.050	0.053	0.041	0.040
Number of school-year observations	57,230	45,138	45,138	45,138
Number of years between observations (k)	1	3	3	3
Indicators for first and last years of spells	No	Yes	Yes	No
Demographic controls	Yes	Yes	No	No

Notes: The coefficients come from regressions of the squared difference in test score gain between years *t* and *t-k* on an indicator for a principal transition, where k is the number of years between observations. The new principal indicator equals one if there is a new principal in year *t* relative to *t-k.* All regressions include school fixed effects and year fixed effects. The student demographic controls are the squared difference in proportions black, Hispanic, female, classified as special education and new entrants to the school, in school enrollment, and in average parental SES. Standard errors clustered by school are in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Appendix Table A3. Distributions of Responses to CPS Principal Survey Questions Used To Construct the Data Usage Practice Index

Use of Assessment Data in Performance Measurement	not at all	a little	somewhat	very much	Obser-vations
1. How much do you use each of the following to gauge how well your school is performing?					
a. results from the ISAT test	0.029	0.135	0.356	0.480	281
b. results from the NWEA test	0.032	0.014	0.118	0.836	280
c. student grades	0.015	0.107	0.460	0.418	337
d. benchmark assessments designed by your teachers	0.074	0.165	0.425	0.336	339
e. classroom observations	0.003	0.030	0.233	0.734	335
2. To what extent do you use assessment data to					
a. evaluate the progress of individual students	0.006	0.033	0.274	0.688	336
b. evaluate the progress of subgroups of students	0.003	0.063	0.325	0.609	335
c. compare performance among classrooms in your school	0.015	0.083	0.292	0.610	336
d. identify teachers who could serve as instructional leaders	0.021	0.054	0.316	0.609	335
3. To what extent do you use assessment data to					
a. evaluate teacher performance	0.027	0.078	0.451	0.444	333
b. inform and communicate with parents	0.012	0.115	0.436	0.436	330
c. assess schoolwide instructional programs	0.003	0.033	0.283	0.681	332
d. monitor changes in your schools test scores over time	0.009	0.015	0.223	0.753	332
e. compare your school to similar schools	0.033	0.103	0.320	0.544	331

Appendix Table A4. Distributions of Responses to CPS Principal Survey Questions Used To Construct the Instructional Leadership Index

Instructional leadership practices	Never	a few times per year	a few times per month	1-2 days per week	>2 days per week	
1. In general, when working directly with teachers this year, how often did you do each of the following:						
a. Demonstrate instructional practices in a classroom	0.158	0.436	0.257	0.085	0.064	
b. Observe teacher instruction in a classroom	0.000	0.021	0.094	0.271	0.615	
c. Advise teachers about classroom practices	0.000	0.026	0.167	0.368	0.439	
d. Examine what students were working on during a teacher's lesson	0.003	0.021	0.138	0.355	0.484	
e. Review student assessment data with your teachers	0.006	0.135	0.394	0.279	0.185	
f. Meet with teachers to give feedback on their instruction.	0.003	0.053	0.232	0.382	0.329	
2. How often do you review assessment data (of any type)?						
	Never	Once a year	A few times a year	A few times a semester	Monthly	Weekly
a. I review assessment data independently	0.003	0.006	0.044	0.144	0.359	0.444
b. I review assessment data with other administrators in my school	0.009	0.009	0.053	0.190	0.377	0.362
c. I review assessment data with other administrators in the district	0.124	0.080	0.239	0.230	0.254	0.074
d. I review assessment data with network staff	0.044	0.062	0.280	0.283	0.271	0.059
e. I review assessment data with teachers	0.003	0.000	0.077	0.233	0.325	0.363
f. I review assessment data with groups of teachers	0.006	0.003	0.080	0.220	0.333	0.357

Appendix Table A5. Distributions of Responses to CPS Student and Teacher Questions Used To Construct the Student and Teacher Indices

	(1)	(2)	(3)	(4)
	Student Surveys			
How safe do you feel…	Not Safe	Somewhat Safe	Mostly Safe	Very Safe
In your class	0.0273	0.0751	0.2334	0.6642
In the hallways/bathrooms	0.0569	0.1563	0.3322	0.4546
How much do you agree	Strongly Disagree	Disagree	Agree	Strongly Agree
The topics we are studying are interesting and challenging	0.0832	0.2173	0.4589	0.2406
I work hard to do my best in this class	0.0199	0.0717	0.5100	0.3984
	Teacher Surveys			
The principal at this school	Strongly Disagree	Disagree	Agree	Strongly Agree
Makes teaching expectations clear	0.0266	0.0754	0.4104	0.4876
Communicates clear vision for school	0.0336	0.1001	0.3908	0.4755
Understands how students learn	0.0364	0.1018	0.4250	0.4369
Sets high standards for student learning	0.0227	0.0572	0.4097	0.5104
Presses teachers to implement professional development	0.0251	0.0959	0.4346	0.4445
Tracks student academic progress	0.0277	0.0994	0.4365	0.4363