

Do Principals Help Teach Principles? Administrator Compensation and School Performance in Colorado Public Schools

David K Carlson, Daniel KN Johnson*
Colorado College, Colorado, USA

Research Article

Received date: 14/12/2015
Accepted date: 28/12/2015
Published date: 31/12/2015

*For Correspondence

Daniel KN Johnson, Schlessman Professor, Department of Economics and Business, Colorado, United States of America, Tel: +719-389-6654

E-mail: djohnson@ColoradoCollege.edu

Keywords: Principal, Compensation, Salary, Education, Leadership, CSAP, Administration, School Performance.

ABSTRACT

This study examines the relationship between principal salaries and student performance on Colorado Student Assessment Program (CSAP) tests of math, reading and writing, using multivariate quartile regressions on data from the 2002-2007 school years. Using instrumental variables to allay the concerns of potential endogeneity in administrator and teacher salaries, and random effects panel data estimation to control for school-specific factors beyond those which we explicitly include, we confirm the intuitive and historically documented effects of funding and conduct violations on average scholastic achievement. Most interestingly, we find that high-performance schools receive an unambiguously positive impact from higher administrator salaries at the expense of teacher salaries, while low-performance schools see the reverse effect.

INTRODUCTION

“Given the importance of principals, and the role of compensation in determining the quality of people who opt to pursue this career path, it is shocking that we know so little about principal compensation”^[1] This paper explores the relationship between public school administrator compensation and the scholastic performance of the students in their care. More specifically, this study seeks to add to a nascent but growing literature on the tradeoff between administrator and teacher salaries, using panel data for all 1872 Colorado public schools (elementary, middle and secondary institutions) that reported state-wide standardized test scores in the three legally mandated disciplines of math, reading and writing over the 2002-07 academic year. Our analysis is sensitive to the potential endogeneity of principal and teacher salaries (using instrumental variables techniques to ameliorate the problem), as well as to the immeasurable differences between student populations and/or institutions (using a range of measured control variables, random effect panel data techniques, and quartile regression analysis to recognize inherent differences).

Colorado was chosen as the site for this research for three reasons. First, it has some of the longest-lived, most standardized and most transparent education databases in the country. Second, the Colorado Department of Education was kind enough to share those databases with us in their entirety. Third, Colorado faces a challenging financial situation, more constrained than most states by its own choice. The state has operated under a balanced budget amendment since its inception (Article X, 2), and under the Taxpayer Bill of Rights (TABOR) since 1992, requiring that any increase in expenditures that exceeds the combined rates of growth in population and inflation must be proposed and must be popularly approved via a ballot initiative. After accommodating mandated spending not subject to those bounds (e.g. on Social Security, Medicare/Medicaid), into a total budget limited in growth, the state found TABOR placed great tension on the K-12 public education budget. In order to reverse that trajectory, in 2000 Amendment 23 carved out funding specifically for K-12 public education, and there is intense public pressure to ensure that every dollar is spent wisely. Yet administrative hiring and firing decisions, and even salary negotiations, are all decentralized, left in the hands of school district officials (who have to propose and sponsor their own public ballot initiatives for any bond issue or spending beyond the TABOR cap).

The policy and academic implications of this paper's findings could be extensive, both in Colorado and in the national debate on education reform. Superintendants, school boards and lawmakers could use these findings to allocate resources more effectively, hold school leadership more accountable and make appropriate policy adjustments based upon differences in principal effectiveness across subjects and school performance levels. At the very least, we hope that these findings add to the nascent but rapidly growing study of compensation in educational leadership. This paper proceeds with a review of the related literature on school performance and education-related compensation in Section 2. The subsequent section describes our data, their advantages and limitations, along with the regression methodology we apply to them. Section 4 presents our econometric results, while Section 5 concludes with the implications for policy.

LITERATURE REVIEW

Despite the controversy concerning the type of assessment, standardized tests have become an obvious focal point in assessing school performance Papay et al. ^[2] for a useful recent review), so a wide array of research has focused on the factors associated with successful test-taking Bettinger ^[3] for a recent review).

Even for those who argue in favor of different metrics of scholastic success, the focus centers on attributes of organizational structure and administration. For example Lee and Burkham ^[4] evaluated drop-out rates and found that school organization and structure have a significant impact on high school students' attrition. The strength of student-faculty relationships, school size and academic offerings were all important determinants in students' decisions to drop out of high school, suggesting that smaller, academic-focused schools with compassionate teachers are best at keeping students through graduation.

One of the most highly politicized and debated aspects of education policy is teacher compensation, presumably because it involves the livelihood of almost 3.2 million public employees nationwide, with a cost in excess of \$562 billion ^[5]. In their exhaustive study of Texas public schools, Hanushek and Rivkin ^[6] determine that salaries have a significant impact on both teacher retention and long-term student performance, leading them to advocate for compensation tied more closely to student performance, rather than traditional metrics like education level and experience. Interestingly, teachers that demonstrate the most trust in their principal are most likely to favor pay-for-performance programs, as was found in a recent survey of Washington state public school teachers ^[7]. These findings suggest that teachers are more willing to have their effectiveness (via student achievement in this case) evaluated if they believe in the person leading them a strong case for the effect of principal leadership on faculty outcomes.

There is a small but growing literature on K-12 school leaders, ranging from mobility and qualifications to performance and effectiveness. Ehrenberg ^[8] revealed that New York superintendants moved between districts for salary increases more than any other factor, in an industry where the pay structure was surprisingly tied to tax rates more than anything else. Notably, student achievement, school performance, standardized test score results or any other measure of successful outcomes had little to no bearing on superintendant mobility. Akiba and Reichardt ^[9] confirmed those findings for Colorado principals who, in their study, were motivated to change schools by pay and advancement possibilities more than their ability to improve student performance, though "student achievement" was a minor motivating factor in mobility.

Using the national Schools and Staffing Survey of 2003-04, Goldhaber^[11] determined that school profiles (such as size and demographics) and principal attributes (such as degree attained and experience) have some significance on determining compensation, and Billger ^[10] used the same data to find, interestingly, that principals receive lower salaries in schools that are required to meet state, local and district accountability goals. One possible explanation for this is that low-achievement schools are often in poorly funded districts, leading to a vicious cycle of weak attraction for qualified administrative candidates, weak leadership and rapid turnover. Since all Colorado schools are required to meet state accountability standards, Billger's findings are potentially highly relevant for the current study.

A trio of working papers from authors at the Center for the Analysis of Longitudinal Data in Education Research confirm the important correlation between principal compensation and student academic achievement. Clotfelter et al ^[11] provides strong evidence that the credentials, and therefore the expected salaries, of North Carolina teachers and principals is consistently lower in poorer school districts. Branch et al. ^[12] confirms that the variation in principal compensation and implicit effectiveness is largest for relatively poor Texas school districts. Clark et al. ^[13] find that there are positive returns to New York principal experience: more seasoned administrators, who on average have higher salaries than their less-experienced peers, are correlated with higher student achievement scores. Our work here aims to build upon this strong foundation, to compare Colorado's experience with these results, aiming for as close to best-practice methodology as the data will permit.

MODEL and DATA

The Colorado Department of Education has administered an annual state-mandated standardized test in math, reading and writing to every student in Colorado public schools grades 4-10, every February since 1997. Those raw Colorado Student Assessment Program (CSAP) scores are not published, but by state law the Department of Education publishes the percentage of eligible students (i.e. of the correct age, enrolled in public school, and physically present for the examinations) who score a

level of ‘proficient or above’ for every relevant public educational institution, aggregated by scholastic level (elementary, middle or secondary school). Thus for our purposes, we treat this percentage, in each testing section (math, reading or writing) as the dependent variable in the analysis which follows.

Following precedent in the literature ^[12,13], we posit that teacher and administrator salaries in part reflect the effectiveness of those public employees in generating scholastic results, so propose them as the primary explanatory variables. We assume that labor markets are fairly efficient, meaning that (in)effective employees will be discovered quickly, to be offered (lower) higher salaries, relying on the work of Ehrenberg et al. ^[8] and Akiba and Reichart ^[9] as evidence of that efficiency. However, as Branch et al. ^[12] and Clotfelter et al. ^[11] clearly demonstrate, salaries are endogenous in this explanation, a function not only of the individual but of the institution. Thus we emulate Goldhaber ^[1] and Billger ^[10] to instrument for teacher and administrator salaries in the primary model. After presenting the model, we describe and summarize the key variables.

We propose a simple linear reduced form explanation for CSAP scores:

$$\text{score}_i = \beta_0 + \beta_1 \text{adminsals}_i + \beta_2 \text{teachsals}_i + \beta_3 \text{lunch}_i + \beta_4 \text{stratio}_i + \beta_5 \text{conduct}_i + \beta_6 \text{perpupilfund}_i + \beta_7 \text{trend}_i + u_i + \varepsilon \quad (1)$$

where score_i is the percentage of students in a given school who score at the proficient level or higher in the (math/reading/writing) section of the standardized test;

adminsals is the cost-of-living adjusted average administrator salary in school i ;

teachsals is the cost-of-living-adjusted average teacher salary in school i

lunch is the percentage of students in school i who qualify for federal free or reduced-fee lunch programs;

stratio is the average student-to-teacher ratio in school i ;

conduct is the number of reported conduct code violations per student in school i ;

perpupilfund is the total school district revenue divided by the total number of students enrolled in the district within which school i resides;

trend is a time variable taking the value of 1 in the 2002-03 academic year; and

u_i is the school-specific effect to be estimated using panel data methodology.

To instrument for adminsals and teachsals , we propose a suite of potential proxy variables, as informed by the literature and as available in the circumstances. Specifically,

$$\begin{aligned} \text{adminsals}_i = & \delta_0 + \delta_1 \text{lunch}_i + \delta_2 \text{stratio}_i + \delta_3 \text{conduct}_i \\ & + \delta_4 \text{perpupilfund}_i + \delta_5 \text{trend}_i + \delta_6 \text{averageperf}_i \\ & + \delta_7 \text{localshare}_i + \delta_8 \text{adminft}_i + \delta_9 \text{adminpt}_i \\ & + \delta_{10} \text{teachft}_i + \delta_{11} \text{teachpt}_i + \varepsilon \end{aligned} \quad (2)$$

$$\begin{aligned} \text{teachsals}_i = & \rho_0 + \rho_1 \text{lunch}_i + \rho_2 \text{stratio}_i + \rho_3 \text{conduct}_i \\ & + \rho_4 \text{perpupilfund}_i + \rho_5 \text{trend}_i + \rho_6 \text{averageperf}_i \\ & + \rho_7 \text{localshare}_i + \rho_8 \text{adminft}_i + \rho_9 \text{adminpt}_i \\ & + \rho_{10} \text{teachft}_i + \rho_{11} \text{teachpt}_i + \varepsilon \end{aligned} \quad (3)$$

where averageperf_i is the unweighted average score for school i across math, reading and writing scores in the same year;

localshare_i is the percentage of total school district revenue contributed by local property taxes;

adminft_i is the number of full-time administrators at school i ;

adminpt_i is the number of part-time administrators at school i ;

teachft_i is the number of full-time teachers at school i ; and

teachpt_i is the number of part-time teachers at school i .

The first two instruments, averageperf and localshare , are designed to reflect the Akiba and Reichert ^[9] finding that employees are drawn to positions not only by the quality of the scholastics but by the pay (given Colorado’s constrained state budget, higher

salaries have to supported by localities willing to pass bond issues, i.e., pay a larger share of total expenses using local funding). In addition, average administrator salaries may be lower where the position is shared with other administrators, but higher where there are more teachers to supervise. For completeness, we include all original variables in these instrument equations as well, to proxy for poorer districts (higher value of lunch_i), less intimate work environments (higher value of stratio_i), more dangerous working conditions (higher value of conduct_i), the size of the resource pool (higher value of perpupilfund_i), and a time trend.

Each administrator (and teacher) average salary is adjusted using the annual cost of living adjustment as determined by the Colorado Department of Education in order to allocate state funds to each of its 178 school districts. For example, the correction factors in 2003-04 range from 1.64 in Aspen (a high-income mountain community) to 1.07 in Stratton (a low-income ranching community).

Turning now to a description of the variables themselves, we present summary statistics in **Table 1** for our unbalanced panel. The data represent the 2002-07 academic years (omitting 2005-06 due to data availability issues), spanning 1872 academic institutions at least once in the period, for 6484 observations in all.

Table 1. Summary statistics of key variables.

Variable	Mean	Standard Deviation	Minimum	Maximum	Frequency of zeroes
Math	50.5	23.2	0	100	52
Read	66.4	17.8	0	100	11
Writ	52.3	18.9	0	100	2
Adminsals	58815	10348	3869	181421	0
Teachsals	34756	5498	1216	58794	0
Lunch	0.31	0.24	0	1	136
Stratio	16.5	12.6	4	421	0
Conduct	0.14	0.40	0	24	569
Perpupilfund	7079	1601	1159	78783	0
Averageperf	57.4	18.7	0	98.7	2
Localshare	0.36	0.14	0	0.87	1
Adminft	1.53	1.13	0	8	601
Adminpt	0.23	0.55	0	7	5245
Teachft	26.9	19.1	0	186	19
Teachpt	4.10	4.20	0	50	719

Notice first that there is significant variance in each of the three CSAP test sections, each one ranging across the entire spectrum from 0 to 100, with averages in the range of 50 to 66. Almost one percent of the observations show a 0 for the math score (more than show a math score of 95 or higher). In fact, further down in the table, the variable averageperf shows that while one school scored 98.7 as the average across all three scores, 2 schools scored a 'perfect' zero in all three scores. It is due to this enormous variation that we will report quartile regressions in the analytical stage of this study.

Salaries are no less interesting in range. While administrator salaries average \$58,815 in cost-of-living-corrected currency, there is enormous variability there as well, ranging from a paltry \$3869 per year to an impressive \$181421 annually. Naturally, that range masks the differences between full-time and part-time positions, as well as small and large schools. Teacher salaries average just below sixty percent of administrator salaries, topping out just below the administrator average and reaching a minimum of \$1216 for four part-time positions.

We treat "lunch", the percentage of students within each school on free or reduced lunch, as determined by the National School Lunch Program ^[14], as a proxy for the relative income of the students that attend each school, since this federally standardized data is widely regarded as an accurate reflection of average family income. While 136 observations in our sample see no students eligible for the program, the average school sees just under one-third of all students eligible. Three schools see one hundred percent eligibility.

Average class size, or more appropriately, student-to-teacher ratio, is 16.5 in our sample, but ranges as dramatically as every other variable, from 4 to over 400 students per teacher. Conduct violations show similar variation, averaging 0.14 violations per student, with nine percent of all schools reporting no violations at all, but peaking at an alarming 24 violations per student per academic year.

Funding averages a little over \$7000 per student per year, but some schools run on less than 1/5 of that while others use eleven times the average. On average, that funding relies on local communities to support roughly one-third of the expenses (localshare), but one school received no local support at all and one received 87 percent from locally approved taxes. Obviously, the rate at which local communities are taxed is based upon the community's willingness to be taxed (expressed through local mill levy elections) which is independent of its ability to be taxed (the value of the property in the district).

The average school functions with 1.5 full-time administrators, another 0.25 part-time administrators, 27 full-time and 4 part-time teachers. There are institutions which do not have one category of employee or another, but every school registers at least 1 administrator and 1 teacher of some variety.

RESULTS

For clarity and brevity, we exclude the first-stage IV regression results of equations (2) and (3) here, but they are available from the authors upon request. They show uniformly very strong Wald chi-squared statistics, and intuitive coefficients: administrator and teacher salaries are higher in schools that are well-funded (lower incidence of eligibility for free and reduced-fee lunches, higher funding per pupil, more local funding support), and have risen marginally with the time trend. Both administrators and teachers tend to make higher salaries in schools where there are fewer administrators and more teachers, a result that we find intuitive for administrator salaries but surprising for teacher salaries. All other variables show significance in certain regressions, but no overall pattern across specifications and subsamples. **Table 2** presents the primary results of this paper, presenting quartiles of the population in the columns and different sections of the standardized test in the rows. In order to keep all observations of an institution together in one quartile (to take advantage of random school-specific effects, despite the ability of a school to improve or decline across the time periods observed), quartiles are defined based on each school's average performance across all time periods and all three standardized tests. Thus, the top quartile represents the same institutions in each time period, in each test, subject to data availability. There are 1686 (467), 1657 (467), 1672 (465) and 1456 (467) observations (schools) in each quartile respectively. We chose to break out the top 10% of all institutions (654 observations, 186 schools) for special consideration, to determine whether the elite institutions were in any way different than their peers in the top quartile. Sensitivity analysis separating out elementary, middle and secondary schools by quartile show which in every meaningful way are identical to the aggregate results, so are not presented here but are available from the authors. Similar tests using lagged instrumented salaries also show similar results. In all cases, F-statistics confirm the statistical significance of the proposed model at all reasonable confidence levels.

Table 2. Primary regression results.

Variable	Top 10%		Top 25%		Second 25%		Third 25%		Bottom 25%	
	coeff.	t-stat.	coeff.	t-stat.	coeff.	t-stat.	coeff.	t-stat.	coeff.	t-stat.
Math scores										
Adminal	3.11×10^{-4}	(1.56)	6.16×10^{-4}	(3.35)***	5.21×10^{-4}	(2.68)***	-1.72×10^{-3}	(4.12)***	-1.05×10^{-4}	(0.83)
Teachsal	6.20×10^{-4}	(1.36)	-2.52×10^{-3}	(3.21)***	-3.65×10^{-3}	(2.11)**	4.31×10^{-3}	(4.97)***	1.74×10^{-3}	(5.96)***
Lunch	0.16	(0.03)	5.71	(1.21)	8.54	(1.79)*	19.18	(5.49)***	14.07	(6.77)***
Stratio	0.02	(1.67)*	-6.52×10^{-3}	(0.45)	3.54×10^{-2}	(0.95)	2.63×10^{-2}	(0.82)	-2.69×10^{-2}	(0.99)
Conduct	-25.85	(4.65)***	-6.96	(2.48)**	-4.52	(2.10)**	-8.65	(3.52)***	-1.47	(3.41)***
Perpupilfund	-2.43×10^{-6}	(0.01)	1.14×10^{-3}	(1.99)**	1.14×10^{-3}	(2.28)**	9.01×10^{-4}	(1.95)*	1.69×10^{-4}	(1.06)
Trend	0.68	(2.53)***	2.53	(4.77)***	3.85	(3.01)***	0.98	(2.88)***	0.66	(2.52)***
Constant	40.69	(4.69)***	111.68	(5.17)***	132.14	(2.38)**	-9.83	(0.95)	-39.86	(5.59)**
F-stat		25.49***		40.96***		41.47***		29.14***		45.87***
Reading scores										
Adminal	2.50×10^{-4}	(2.16)**	4.04×10^{-4}	(3.88)***	6.20×10^{-4}	(2.98)***	-3.40×10^{-3}	(4.90)***	-8.36×10^{-4}	(5.42)***
Teachsal	6.02×10^{-5}	(0.25)	-1.17×10^{-3}	(2.66)***	-8.92×10^{-4}	(2.41)**	6.67×10^{-3}	(4.84)***	3.13×10^{-3}	(9.39)***
Lunch	-2.18	(0.69)	-8.37	(3.08)***	-11.13	(7.10)***	-20.25	(4.13)***	-21.40	(8.76)***
Stratio	8.47×10^{-4}	(0.12)	-1.12×10^{-2}	(1.35)	-2.34×10^{-2}	(0.88)	4.60×10^{-3}	(0.09)	-1.05×10^{-2}	(0.30)
Conduct	-4.14	(1.34)	-3.49	(2.17)**	-0.43	(0.34)	-3.96	(1.03)	-0.98	(1.77)*
Perpupilfund	4.15×10^{-4}	(0.87)	3.20×10^{-4}	(0.97)	7.56×10^{-4}	(3.29)***	-7.92×10^{-4}	(1.19)	5.31×10^{-4}	(2.73)***
Trend	-0.49	(3.32)***	0.34	(1.13)	0.15	(0.92)	0.74	(1.41)	-0.54	(1.63)
Constant	71.12	(15.61)***	100.08	(8.14)***	66.00	(16.01)***	38.90	(2.80)***	-6.01	(0.75)
F-stat		3.56***		4.64***		10.16***		5.30***		29.31***
Writing scores										
Adminal	4.06×10^{-4}	(2.64)***	5.10×10^{-4}	(3.90)***	7.06×10^{-4}	(3.04)***	-3.57×10^{-3}	(4.67)***	-7.06×10^{-4}	(4.75)***
Teachsal	5.71×10^{-5}	(0.14)	-1.26×10^{-3}	(2.26)**	-5.86×10^{-4}	(1.40)	7.28×10^{-3}	(4.88)***	3.18×10^{-3}	(9.66)***
Lunch	3.58	(0.72)	-1.59	(0.46)	-8.78	(4.71)***	-16.20	(3.20)***	-11.43	(4.79)***
Stratio	6.82×10^{-3}	(0.67)	-2.04×10^{-3}	(0.19)	-4.99×10^{-2}	(1.67)*	-1.91×10^{-2}	(0.36)	-1.34×10^{-2}	(0.41)
Conduct	-6.04	(1.34)	-4.39	(2.15)**	1.95	(1.34)	-0.75	(0.18)	-0.50	(0.95)
Perpupilfund	7.25×10^{-4}	(0.99)	9.21×10^{-4}	(2.19)**	9.99×10^{-4}	(3.76)***	-1.08×10^{-3}	(1.56)	4.51×10^{-4}	(2.43)**
Trend	-0.68	(2.97)***	8.20×10^{-2}	(0.21)	-0.24	(1.30)	0.53	(0.94)	-1.14	(3.68)***
Constant	50.54	(6.55)***	81.09	(5.22)***	34.43	(6.89)***	12.41	(0.88)	-33.01	(4.14)***
F-stat		3.15***		3.19***		(8.85)***		4.25***		18.96***

Random school-specific effects have been estimated but are not reported here. *** represents confidence at the 99% level, ** represents the 95% level, * represents the 90% level.

First, notice that administrator and teacher salaries are significant in 24 of the 30 coefficient estimates presented in **Table 1**. Even more interestingly, administrator salaries show a positive coefficient for high quartiles and a negative coefficient for low quartiles, while teacher salaries show the reverse pattern. In other words, attracting great administrators seems more important for scholastic success at strong schools, even at the expense of weaker teachers (because the students and institution are already highly capable). Conversely, funds spent on productive administrators are wasted at weaker schools, where the resources

should be spent on better teachers to help students excel. It appears that strong institutions need great leadership, while weaker institutions need stronger workers.

The size of the coefficients is also compelling. In almost every case, the positive salary coefficient (whether for the administrator or teacher) outweighs the negative salary coefficient, implying that across-the-board percentage increases in the average salary at the each institution are associated with better scholastic outcomes. However, notice that it is more expensive to raise the average salaries of teachers than to raise the average salaries of administrators due to the relative size of their populations (there are over 174,000 full-time teachers represented in the data, compared to a little under 10,000 administrators). Furthermore, across-the-board percentage increases are easiest to administer, but are ill-advised as a strategy for improving scholastic performance, as it is clear that raises in some salaries are associated with much more dramatic outcomes. For example, an increase of \$10,000 in average administrator salaries in second quartile schools is associated with better than 5-, 6- and 7-point increases in math, reading and writing scores respectively. The same increase in average teacher salaries in the third quartile is associated with better than 4-, 6- and 7-point increases respectively. On the other hand, increases in administrative salaries in the third quartile and increases in teacher salaries in the second quartile are associated with marked drops in scholastic results.

Turning now to the control variables, there are some interesting effects to consider. "Lunch" had a strong negative correlation with student performance in writing and reading, but curiously has a positive correlation with math scores in the lower half of all schools. Given the range of the variable, between 0 and 1, and the difference between average math scores in each quartile, the effects are small but significant. For example, the poorest school populations in the lowest quartile (where every student qualifies for the lunch program) have math scores average 11.3, so even with the boost from the "lunch" coefficient, their predicted scores just barely reach the average in among own peers in the bottom quartile.

The student-to-teacher ratio appears insignificant in virtually every regression, fluctuating between positive and negative estimated coefficients. On the other hand, the number of conduct violations appear negatively related to performance in all but one regression, and are strongly significant for math scores in particular. Funding per pupil is not always relevant, but where it is statistically significant it is small and positive. For example, an extra \$1000 per student in second-quartile schools is associated with an increase in each average test score by roughly 1 point.

The time trend is largely insignificant, except for in math scores, where each quartile appears to be improving gradually over time once other factors have been controlled. Apparently, the gap between best and worst school performances has been broadening over time, as the two top quartiles have been progressing much faster than have the lower quartiles. It is also disturbing that the lowest quartile of schools has actually been deteriorating in writing scores, *ceteris paribus*.

CONCLUSIONS

Administrator and teacher salaries have been found to have a complicated correlation with student performance, with test scores rising with administrator salaries at high-quartile schools and falling with administrator salaries at low-quartile schools. Teacher salaries have shown the reverse pattern. This finding concurs with the literature which suggests that schools that are already willing and able to attract better leadership talent with the lure of higher salaries are able to achieve better standardized test scores^[9], leading to a virtuous cycle. Strong schools attract good teachers a priori, so they can save on teacher salaries to spend even more in attracting or retaining great leadership for the next year. For weak schools, the results suggest that they settle for lower-paid (and presumably lower-quality administrators) as they pay more to attract and retain high-quality teachers (which thankfully boosts their achievement scores). However, this raises the obvious question about where the crossover from a teacher-centric to a leadership-centric strategic approach occurs.

We have attempted in this paper to control for obvious endogeneity issues, using an instrumental variables approach combined with random school-specific effects to account for unobservable factors that may bias our results. Nevertheless, we are sensitive to the fact that our model specification is only one of many that could be proposed. Our concerns are allayed somewhat by the fact that our results accord with the existing literature and that sensitivity tests on subsets of schools, by age of student or by quality of school, show complementary conclusions.

We had anticipated that math scores might be most sensitive to school attributes, hypothesizing that reading and writing skills are more continually supplemented outside the classroom, both through the everyday life of children (like reading the menu at a restaurant) and by the conscious effort of parents (through bedtime stories, ample home libraries and forced letters or emails to relatives). However, math especially above simple addition, subtraction, multiplication and division is a subject generally less practiced outside the classroom in the lives of 6 to 18 year olds, even in homes that insist on practicing other skills. Therefore, we surmised, students are on a more even educational playing field with math testing, making the influence of school factors more important in determining standardized scores in the subject. However, we found little evidence to support this hypothesis.

Admittedly, this study has much room for improvement and future research in this area should be informed by our shortcomings. For example, the authors would have preferred to include data on teacher quality, teacher time-on-task, principal and teacher tenure/duration in the position, and the availability of co-curricular activities at each school.

Most interestingly, teacher salaries had the largest impact on student test scores in the lowest performing schools, a novel

observation that deserves both consideration and further study. This result could be attributable to the fact that the mobility rate for teachers is highest in the worst schools, meaning that increases in teachers' pay could allow schools in this quintile to attract the best teachers available and willing to teach in these worst performing schools^[6]. Policymakers should take particular note of this finding.

The primary finding of this paper that principal and teacher salaries do indeed correlate with student performance in quantifiable ways should inform state policymakers, School Boards, and superintendents alike. If policymakers intend to make schools more accountable for the performance of the children within their classrooms, they must appreciate the impact of district-based resource allocation decisions, even within the salary pool of a single school.

REFERENCES

1. Goldhaber D. Principal Compensation: More Research Needed on a Promising Reform. The Center for American Progress. 2007.
2. Papay J, RJ Murnane and JB Willett. How Performance Information Affects Human-Capital Investment Decisions: The Impact of Test-Score Labels on Educational Outcomes. NBER Working Paper. 2011.
3. Bettinger E. Paying to Learn: The Effect of Financial Incentives on Elementary School Test Scores. NBER Working Paper. 2010; 16333.
4. Lee V and DT Burkham. Dropping out of High School: The Role of School Organization and Structure. American Educational Research Journal. 2003;40:353-393.
5. National Center for Education Statistics (NCES), data from nces.ed.gov.
6. Hanushek E and SG Rivkin. Pay Working Conditions and Teacher Quality. The Future of Children Excellence in the Classroom. 2007; 69-86.
7. Goldhaber D, M DeArmond and S DeBurgomaster. Teacher Attitudes about Compensation Reform: Implications for Reform implementation. School Finance Redesign Project University of Washington. 2007.
8. Ehrenberg R, RP Chaykowski and RA Ehrenberg. Determinants of the Compensation and Mobility of School Superintendents. Industrial and Labor Relations Review. 1988;41:386-401.
9. Akiba M and R Reichert. What Predicts the Mobility of Elementary School Leaders? An Analysis of Longitudinal Data in Colorado. Education Policy Analysis Archives. 2007.
10. Billger SM. Principals as Agents? Investigating Accountability in the Compensation and Performance of School Principals. Industrial and Labor Relations Review. 2007;1:90-107.
11. Clotfelter CT, HF Ladd, J Vigdor and J Wheeler. High poverty schools and the distribution of teachers and principals. The Urban Institute CALDER Program Working Paper. 2006.
12. Branch G, E Hanushek and S Rivkin. Estimating principal effectiveness. The Urban Institute CALDER Program Working. 2009;32.
13. Clark, D, P Martorell and J Rockoff. School Principals and School Performance, Working Paper 38 from National Center for Analysis of Longitudinal Data in Education Research. 2009.
14. United States Department of Agriculture. National School Lunch Program data at www.fns.usda.gov/cnd/Lunch/.