How Should School Districts Shape Teacher Salary Schedules? Linking School Performance to Pay Structure in Traditional Compensation Schemes

Jason A. Grissom and Katharine O. Strunk

Educational Policy published online 25 August 2011
DOI: 10.1177/0895904811417583

Educational Policy
XX(X) 1–33

DOI: 10.1177/0895904811417583
http://epx.sagepub.com
SAGE

How Should School Districts Shape Teacher Salary Schedules? Linking School Performance to Pay Structure in Traditional Compensation Schemes

Jason A. Grissom[1] and Katharine O. Strunk[2]

Abstract

This study examines the relative distribution of salary schedule returns to experience for beginning and veteran teachers. We argue that districts are likely to benefit from structuring salary schedules with greater experience returns early in the teaching career. To test this hypothesis, we match salary data to school-level student performance data on math and reading achievement tests across states. We find that frontloaded compensation schemes—those that allocate greater salary returns to experience to novice teachers—are associated with better performance in multiple grades and throughout the achievement distribution. Our results contribute to national debates concerning teacher compensation policies.

[1]Vanderbilt University, Nashville, TN, USA
[2]University of Southern California, Los Angeles, CA, USA

Corresponding Author:
Jason A. Grissom, Peabody College, Vanderbilt University, PMB 0414 GPC, Nashville, TN 37203, USA
Email: jason.grissom@vanderbilt.edu

Keywords

teachers' salaries, student achievement, compensation policy

As districts and schools search for ways to strengthen student performance in the current era of high-stakes state and national accountability policies, researchers, and policy makers are examining the role of teacher compensation structures in promoting student achievement. Recent debate over how best to pay teachers has focused mainly on the use of merit pay to reward teachers for their students' test score gains. This compensation strategy differs fundamentally from the traditional salary schedule model—used by 92% of school districts, according to 2007-08 Schools and Staffing Survey calculations—which bases compensation for teachers on years of experience and educational attainment. While proponents see promise in pay-for-performance (see Figlio & Kenny, 2007; Lavy, 2007; Podgursky & Springer, 2007), for a number of reasons such compensation plans are not yet in widespread use. First, critics express concerns about the mismeasurement of teacher performance via student test score gains (e.g., Rothstein, 2010) and the difficulties of distinguishing individual teachers' contributions within the "team" production of schooling. Second, early merit pay programs, including the Denver ProComp plan, the Dallas Educators' Excellence Grants and District Awards for Teacher Excellence, and the Minnesota Q Comp plan, have proven quite expensive, often relying on large one-time grants from private or government sources to increase spending on teachers' salaries (Podgursky & Springer, 2007). Last, in many districts, implementation of compensation models based on merit pay necessitates support from teachers' unions, who often are unenthusiastic about substantial changes to the traditional salary schedule. These political and resource challenges have thus far proved insurmountable in most places, though incentives attached to Race to the Top and other federal policy programs are likely to make state and local experimentation more prevalent. Even so, precedents set by ProComp and other pay-for-performance systems suggest that merit pay is most likely to augment rather than supplant traditional compensation schemes, at least in the near future.

In other words, despite ongoing policy discussions related to pay-for-performance, the traditional uniform salary schedule dominates teacher compensation policy. The likelihood of its continued widespread usage, whether on its own or in conjunction with merit pay, makes the question of whether the uniform salary schedule should be replaced altogether perhaps less immediate than the question of *how best to structure* these schedules to recruit and retain effective teachers. Most salary schedules share a similar architecture:

as teachers gain years of experience, they advance down the rows of the schedule, receiving pay increases at each "step"; as they gain education, they advance across the schedule's columns, shifting pay upward to reward the attainment of a master's degree or some other accumulation of credits. Yet the *sizes* of the pay increases associated with moves along the rows and columns vary substantially from district to district. Local policy makers have begun to recognize that the distribution of those increases across the schedule may matter for outcomes, as witnessed by recent proposals in San Francisco and Denver to increase salaries for newer teachers relative to more veteran teachers. Similar proposals have been considered in New York City and Washington, D.C. (Sawchuk, 2009).

Unfortunately, research attention to salary schedule structure so far has lagged behind school district interest. This study contributes to the relatively small research base on the structure of traditional salary schedules by examining how districts formulate pay raises for early-career teachers relative to veteran teachers. Specifically, we ask whether some patterns of rewards to experience and education are more consistent with district goals of promoting student achievement than others. In particular, we weigh the arguments for *frontloading*—giving larger raises early in a teacher's career and smaller raises later—versus *backloading*—concentrating raises among veteran teachers—which leads us to the hypothesis that frontloading districts are likely to see more positive school performance. Then, using a unique data set that matches compensation data to school-level test score data in 28 states, we test this hypothesis, examining the relationship between salary schedule frontloading and school performance across grades and at multiple points in the achievement distribution (basic competence, proficient, and advanced proficiency). Our results contribute to the ongoing debate over how to choose teacher compensation policies that best serve district goals.

A Rationale for Frontloaded Salary Schedules

Research provides strong evidence that teachers are the most important school factor in predicting student performance (e.g., Rivkin, Hanushek, & Kain, 2005; Wright, Horn, & Sanders, 1997). School district success thus depends on the district's capacity to maintain a high-quality teacher workforce. Providing teachers with higher pay is one means for achieving this goal (Figlio, 1997; Imazeki , 2005; Loeb & Page, 2000;). The realities of district budget constraints, however, mean that any dollar directed toward a given teacher's salary is one that cannot be spent elsewhere. This tradeoff suggests an optimization problem: Can districts maximize their ability to

employ effective teachers under the constraints of local budgets by structuring compensation in a way that targets pay increases at some teachers instead of others? Advocates of pay-for-performance argue that the answer is yes, and that pay increases should be directed toward teachers who demonstrate greater effectiveness via higher student test scores, principal evaluations, and other means. However, within the context of traditional compensation schemes, can we expect that some salary structures might be associated with higher outcomes than others?

Unlike pay-for-performance, salary schedules are pegged to teacher characteristics. Ostensibly, districts gain by rewarding characteristics most closely associated with teacher effectiveness, since such rewards provide a means of attracting and retaining teachers with those characteristics. The difficulty is that very few observable teacher characteristics have been linked consistently to teacher performance. The exception is teacher experience, though substantial gains to effectiveness from experience do not appear to accumulate beyond the 4th or 5th years of teaching (Rivkin et al., 2005; Hanushek & Rivkin, 2007; Rockoff, 2004). Thus a salary schedule designed simply to reward teachers for quality gains associated with experience clearly would be very frontloaded, with the largest raises given in the first 4 or 5 years and much smaller raises thereafter. In contrast, despite the returns to education that are standard in the single salary schedule, substantial evidence exists that attainment of postbaccalaureate degree credits or master's degrees has little bearing on teacher effectiveness (Clotfelter, Ladd, & Vigdor, 2007; Rivkin et al., 2005), suggesting that an optimal salary schedule would not reward such attainment.

Consequently, we focus on the structure of returns to experience. To facilitate this analysis, consider three highly simplified options for an optimal salary schedule structure, depicted in Figure 1: (a) a linear salary schedule, in which a teacher receives an initial starting salary w_0 and a constant increase π for each year of teaching experience accumulated; (b) a frontloaded salary schedule, in which a teacher receives an initial starting salary w_0 followed by an increase π^N for each year of experience up to year N and then an increase of π^V for each year of experience after year N, with, $\pi^N > \pi^V$; and (c) a backloaded salary schedule, in which a teacher receives an initial starting salary w_0 followed by an increase π^N for each year of experience up to year N and then an increase of π^V for each year of experience after year N, with $\pi^N < \pi^V$. Following Monk and Jacobson (1985), we will call teachers who have not yet reached year N *novice* teachers and teachers who have remained in the district through and beyond year N *veteran* teachers.

Under a linear salary schedule, a teacher with t years of experience will earn $W_t = w_0 + \pi \times t$. Returns to experience are constant at each step in the

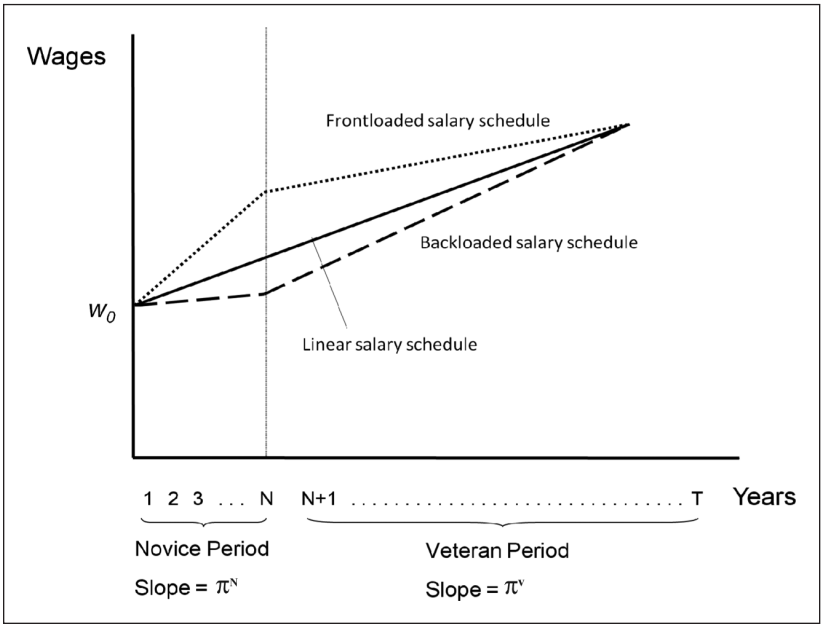

Figure 1. Backloaded vs. frontloaded vs. linear salary schedules

salary schedule. In a frontloaded salary schedule, novice teachers receive larger returns for each year of experience gained than do senior teachers. That is, they earn $\pi^N \times t$ in their earliest years of teaching ($1 \leq t \leq N$) and $\pi^V \times t$ in later years ($N < t$). Since $\pi^N > \pi^V$, novice teachers gain more in salary from year to year than do senior teachers. Under a backloaded salary schedule, $\pi^N < \pi^V$, and senior teachers receive larger experience returns.

Which of the three schedule structures benefits districts the most depends on the incentives implicit in each choice. Human capital theory, which suggests that pay should increase as workers gain skills that make them more valuable to the employer, would imply a frontloaded salary schedule, since effectiveness gains from experience are concentrated early in the career. However, other factors may also be important to districts in their choice of compensation structures. First, districts want to attract high-quality teachers. Higher salaries positively affect the attractiveness of a career in teaching, the attractiveness of teaching jobs in particular districts, and the length of stay in teaching jobs (Dolton, 1990; Hanushek & Pace, 1995; Manski, 1987; Murnane, Singer, Willett, Kemple, & Olsen, 1991; Zabalza, Turnbull, & Williams, 1979).

As higher salaries allow districts to compete with potential teachers' alternative wage offers, districts can purchase greater quality in the local teacher labor market and retain teachers longer by offering higher wages. However, there is evidence that as teachers consider their initial job placements, they respond more to starting salaries than to future rewards, perhaps because they discount future earnings or because they factor in some probability of leaving the profession (Zabalza et al., 1979). The impact of higher salaries earlier in the career likely is even more important for high-quality teacher candidates, who are more likely to command higher opportunity wages in the nonteaching labor market (Goldhaber & Brewer, 1997). Districts get more purchasing power in the market for teacher quality—particularly with regard to teacher recruitment—from offering the largest returns to experience early in the teaching career. Stated differently, if prospective teachers place greater weight on returns to experience they will receive in their novice years than in their veteran years, the return to the district for a high π^N is greater than the return for a high π^V, and the differential increases the more prospective teachers discount future returns to experience.

A second consideration concerns retention. Quit probabilities for teachers are highest in their first few years of teaching, then drop dramatically until teachers near retirement age (Hanushek & Rivkin, 2007; Ingersoll, 2001; Murnane & Olsen, 1989a, 1989b). Teachers who weather the initial years in the classroom are much less likely to leave for a variety of reasons: because they have revealed themselves to be especially committed to the profession, because they have found teaching to be a good fit for their skills and temperament, because they have built up teaching-specific capital that does not transfer easily to other work environments, because they are anticipating the receipt of increasing pension benefits the longer they stay in the district, and so forth. This U-shaped relationship between experience and attrition implies that districts need not continue offering the same set of rewards to keep teachers later in the teaching career as were necessary to retain them through the earliest years when the job was most difficult and they had the greatest set of outside options. In fact, Murnane et al. (1991) found that after 6 to 8 years in teaching, salary ceases to have more than a negligible effect on a teacher's probability of quitting. Once again, the district earns a greater return on π^N than on π^V.

Given evidence in support of frontloading, are there ways in which districts might benefit from a backloaded schedule instead? Ballou and Podgursky (2002) consider the implication from imperfect monitoring models in the labor economics literature that late-career returns may induce higher effort from early-career teachers who do not want to be dismissed before they can realize

those benefits. They dismiss this justification, however, on the grounds that dismissal rates for teachers after tenure, which typically is granted by Year 5 in most states and as early as Year 2 in some states, are very low, suggesting that teachers are not likely to make strong linkages between their performance and the possibility of involuntary turnover. To this point, a recent study by Weisberg, Sexton, Mulhern, and Keeling (2009) found that fewer than 2% of teachers receive "unsatisfactory" ratings on the evaluations that would be cause for dismissal. A study by Heutel (2009) finds further evidence in favor of dismissing this imperfect monitoring justification for backloading. Ballou and Podgursky (2002) also consider extensions of the human capital theory and turnover arguments presented above as potential rationales for accelerated gains late in the career but find theoretical support for neither.

All together, this evidence suggests that frontloaded salary schedules are more likely to be optimal than are linear or backloaded schedules for the average school district. Frontloading offers higher returns to the district because it is more likely to attract high-quality teachers, who are more likely to have higher paying alternatives available to them outside the district, and retain them through the steepest section of the experience-attrition curve by offering them greater incentives to return year-to-year. The frontloaded schedule is also cost-effective in the sense that it does not continue to offer these larger yearly rewards beyond the point in the teacher experience profile when they are no longer likely to significantly influence retention decisions. Setting $\pi^N > \pi^V$ allows the district to shift funds away from ineffective salary premiums toward more productive uses.

If frontloaded salary schedules are, in fact, more consistent with school districts' goals of recruiting high-quality teachers to the district and rewarding them for the substantial gains in effectiveness they make early in their careers, we suggest that districts who chose to implement backloaded schedules fail to allocate their resources optimally. As Vigdor (2008) points out, more frontloaded schedules not only align district and teacher goals by linking compensation increases with effectiveness gains but also can do so at a cost savings because of reduced premiums on experience later in the career. These surplus funds can be used in further pursuit of district goals of increasing student achievement. Moreover, while the misallocation of resources backloaded salary schedules represent may affect district performance directly, this misallocation may also suggest the presence of more general problems with resource allocation and other management choices within the district.

We run three sets of analyses to determine the prevalence of frontloading, the relationship between frontloading salary schedules and student

achievement, and why districts may instead choose to backload their salary schedules. Our first set of analyses investigates descriptively the degree to which districts in our national sample frontload their salary schedules. This analysis builds on studies by Monk and Jacobson (1985) and Lankford and Wyckoff (1997), which considered the structure of teacher salary schedules in one state (New York) in an earlier time period. More recent work by Vigdor (2008) explores the shape of salary schedules in North Carolina. The second set of analyses looks for support for the hypothesis that frontloaded salary schedules enhance district productivity by testing the association between salary schedule structure and student performance. Finally, once we establish the positive association between frontloading and student achievement, we examine alternative predictors of backloading to answer the question of why so many districts employ salary schedule structures that do not appear consiswith the goals of recruiting and retaining high-quality teachers while holding down salary costs.

Operationalizing Salary Schedule Structure

Our first step is the development of a measure of salary schedule structure. Earlier researchers operationalize backloading as occurring when districts provided a disproportionately large share of salary increases to veteran teachers *over time* (Lankford & Wyckoff, 1997; Monk & Jacobson, 1985). Monk and Jacobson (1985) determine the degree to which a district frontloads or backloads its salary increases by regressing the logged salary of teachers in the same experience cohort in a district on a year indicator, whose coefficient then reflects the average annual percentage salary increase for that experience level. Lankford and Wyckoff (1997) use a similar method, but with individual teacher salaries, regressing individual teacher salary on the years of experience (expressed in a piecewise function) and a host of other relevant teacher characteristics. Both sets of researchers found a high prevalence of backloading. However, given that the SASS data limit us to using only a cross-section of data, we are not able to pursue the same longitudinal regression-based approach to measuring backloading. Rather, we use the district-reported salaries associated with teachers at different experience levels to indicate the structure of the district salary schedule *at a given point in time*. Instead of examining how specific steps or groups of steps on salary schedules increase or decrease over time, we compare the yearly returns to experience gains for veteran teachers relative to those for novice teachers. *Yearly returns to experience*

gains means the additional pay teachers receives from one year to the next to reward them for the additional year of experience accumulated. These yearly experience gains determine the slope of the salary schedule between any two years on the schedule. Our measure of salary schedule structure in effect compares the slope of the schedule late in the teaching career to the slope early in the teaching career, highlighting the degree to which districts structure salary schedules to provide premiums to veteran teachers versus novice teachers. We refer to the relative returns to experience late in the teaching career as the *relative experience premium*.

To construct a measure of the relative experience premium, we first denote the average yearly return to experience for veteran (experienced) teachers as π^V and the average yearly return to experience for novice (inexperienced) teachers as π^N. In reality, these returns may vary from year-to-year throughout the salary schedule, but data limitations lead us to focus on average returns. For districts with frontloaded schedules, we should observe a steeper slope in returns to experience in the early years of teaching and a subsequent flattening of the wage profile later. For districts with backloaded schedules, we should observe flatter slopes early in the career but steeper ones as the teacher gains experience. Figure 1 again illustrates these patterns.

To assess the degree to which a district's schedule is more frontloaded or backloaded, we compare the relative magnitudes of π^V and π^N. We express the relative difference between π^V and π^N as the percentage increase a veteran teacher would see in returns to experience beyond the increases given to novice teachers. That is,

$$RELATIVE\ EXPERIENCE\ PREMIUM = \frac{\pi^V - \pi^N}{\pi^N} * 100 \qquad (1)$$

This measure captures frontloading versus backloading in terms of the extra gains to veteran teachers beyond what is given to novice teachers. For districts with linear salary schedules—that is, districts who give constant returns to experience, or identical steps, throughout the schedule—this measure will be zero, since $\pi^V = \pi^N$. For districts who backload the salary schedule ($\pi^V > \pi^N$), this value will be positive, while in districts that frontload ($\pi^V < \pi^N$), it will be negative.

Unfortunately, available data from SASS do not allow us to calculate π^V and π^N precisely or throughout teachers' entire careers because we observe salaries for teachers only at Years 0, 10, and 20. Thus, we obtain rough estimates for π^V and π^N by calculating the average yearly salary returns before

and after the 10th year of teaching. In other words, $\pi^{V} = \frac{salary_{20} - salary_{10}}{10}$ and $\pi^{N} = \frac{salary_{10} - salary_{0}}{10}$, where subscripts denote years of experience. Inserting these definitions into Equation 1, we have,

$$RELATIVE\ EXPERIENCE\ PREMIUM = \frac{\frac{salary_{20} - salary_{10}}{10} - \frac{salary_{10} - salary_{0}}{10}}{\frac{salary_{10} - salary_{0}}{10}} * 100 \qquad (2)$$

As an illustration, consider a salary schedule that progresses such that a beginning teacher with a bachelor's degree starts out earning US$25,000 per year. By Year 10, this same teacher earns US$45,000. In Year 20, yearly earnings are US$50,000. In this example, $\pi^{N} = \frac{\$45,000 - 25,000}{10} = \$2,000$ and $\pi^{V} = \frac{\$50,000 - 45,000}{10} = \$500$. The relative experience premium in this case would be $\frac{\$500 - 2,000}{2,000} * 100 = -75\%$. From this calculation, we conclude that a veteran teacher's yearly experience returns in this system decline by 75% relative to what is earned by novice teachers. This hypothetical district has a frontloaded salary schedule.

Our source for the measure of salary structure is the 1999-2000 Schools and Staffing Survey (SASS), which is the most recent year of the survey for which our salary schedule measure can be constructed.[1] Administered by the National Center for Education Statistics (NCES), SASS covers a nationally representative sample of schools. In schools selected for participation, questionnaires are given to multiple actors, and unique respondent identifiers make the responses fully linkable across surveys, which will be an important feature for the school-level analyses conducted later in the article. We generate our measure of salary schedule structure using multiple responses from the District questionnaire. All salary and spending measures used in our analyses are adjusted by the NCES Comparable Wage Index for 1999-2000 to take into account geographic differences in costs of labor and other inputs (Taylor & Fowler, 2006).

The structure of the SASS data presents additional difficulties in calculating the measure in Equation 2. In particular, while we can substitute the salary for a teacher with a bachelor's degree and no experience ($salary_{BA,0}$) for $salary_{0}$ and the salary for a teacher with a bachelor's degree and 10 years experience ($salary_{BA,10}$) for $salary_{10}$, a similarly straightforward measure for $salary_{20}$ is not readily available because SASS also incorporates additional

pay from attainment of a master's degree into the measurement for that year. That is, we would prefer to calculate π^V as $\pi^V = \frac{salary_{BA,20} - salary_{BA,10}}{10}$, where the subscripts indicate salaries for two teachers with bachelor's degrees at 20 and 10 years of experience, respectively, but SASS does not include a measure that corresponds to $salary_{BA,20}$. Instead, SASS collects information on the salary of a teacher in a given district with 20 years of experience and a master's degree, $salary_{MA,20}$. Given that the step-structure of salary schedules rewards educational attainment, this value no doubt is higher than the unavailable value $salary_{BA,20}$. However, we can manipulate other available SASS data to estimate and subtract the value of the master's degree to get an estimate of $salary_{BA,20}$. In particular, we are given salary values for teachers at the beginning step of the salary scale with both bachelor's and master's degrees: $salary_{BA,0}$ and $salary_{MA,0}$. Making the assumption that the pay bump associated with a master's degree is constant over time, we can approximate $salary_{BA,20}$ by $salary_{MA,20} - (salary_{MA,0} - salary_{BA,0})$.[2]

As operationalized using SASS, the relative experience premium variable measures the percentage difference in the yearly experience return a teacher receives in her second decade of teaching beyond the yearly experience return a teacher with the same education level in that district receives during the first decade of teaching. Positive values indicate greater returns in the second decade relative to the first decade, or a backloaded schedule. Negative values indicate that returns are greater in the first decade relative to the second. Values that are near zero characterize salary schedules as neither frontloaded nor backloaded. Figure 2 shows the true slopes of salary schedules in frontloaded and backloaded districts found in the SASS data set.

Our measure of salary structure defines backloading as the *extra* yearly returns given to veteran teachers above and beyond those given to novice teachers. It is possible, however, that any model results can be sensitive to the particular characterization of backloading rather than the actual action of backloading. To test that this is not the case, we also operationalize backloading in two alternative ways. First, we generate a measure that is the ratio of the yearly returns to veteran teachers to the yearly returns to novice teachers, or $\frac{\pi^V}{\pi^N}$. This measure defines backloading simply as the yearly returns to veteran teachers relative to novices and does not capture necessarily the premium given to veteran teachers. Next, we generate a measure of the relative experience premium in terms of the percentage increase to veteran as opposed

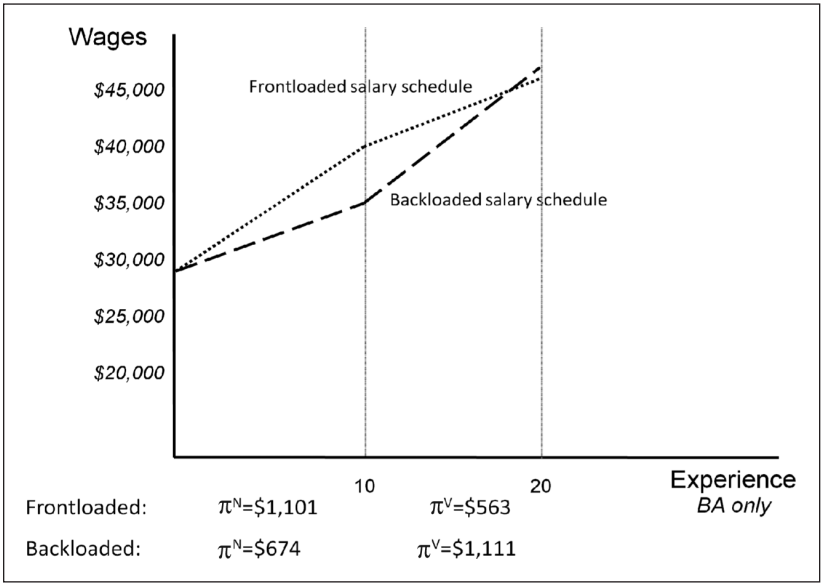

Figure 2. Backloaded vs. frontloaded salary schedules, SASS 1999-2000

to novice teachers, or $\frac{\frac{salary_{20}-salary_{10}}{salary_{10}} - \frac{salary_{10}-salary_{0}}{salary_{0}}}{\frac{salary_{10}-salary_{0}}{salary_{0}}}$. In this instance, we continue to measure the premium given to veteran teachers as in our original measure in Equation 2, but now we define the returns to veteran and novice teachers as the percentage change in salary for veteran or novice teachers relative to their pay at the beginning of their veteran or novice period. For both of these alternative measures, a value of one indicates a neutral salary structure, a negative value denotes a frontloaded structure, and a positive value represents a backloaded schedule. We run all analyses using all three measures and find consistent results (available from the authors upon request). For simplicity and ease of interpretation, we present results only from the measure derived from Equation 2.

Our operationalization of backloading carries with it both advantages and disadvantages relative to measures of backloading used in earlier work by

Monk and Jacobson (1985) and Lankford and Wyckoff (1997). Whereas they capture how veteran versus novice teachers are rewarded over time, we consider how novice versus veteran teachers are rewarded in a given year. This cross-sectional measure captures how an entering teacher views her full earning potential and trajectory: can she expect to be rewarded for her hard work at the beginning of her career or must she wait for the rewards to come later? This question is arguably more important in teachers' decision-making process than is the district's history of changes in raises for early or late career teachers.

However, this method also leaves more room for error: We cannot fully separate out the salary increases due to degree attainment from experience, which may lead to a possible overstating of the returns to pure experience for veteran teachers. Moreover, we cannot make any claims to causality because we are unable to say whether *changes* in salary structure over time from backloading to frontloading or vice versa lead to changes in student outcomes.

We believe that such a correlational study adds to the extant literature even if it cannot provide causal evidence as to the impacts of backloading or frontloading salary schedules. Federal and state policies continue to encourage teacher salary reforms that shift salary schedules away from current structures toward merit pay and other incentive-based reforms. However, there has been relatively little empirical examination of the incentives built into salary schedules as they appear in the majority of districts today. The three empirical studies that do examine the shape of teacher salary schedules all use data from the 1970s through the early 1990s (the most recent longitudinal data set in use in these studies is found in Lankford & Wyckoff, 1997, and runs through the 1994 school year; Monk & Jacobson, 1985; Murnane, Singer, & Willett, 1987). Moreover, they are state based: Two are from New York and one is from Michigan. The earliest of these three also use a data set of the salary schedules of 100 largest American Federation of Teachers (AFT) districts in the country in the 1970s. Revisiting teacher salary schedule structure and its impacts with updated, more representative data can help restart and add to the policy discussion surrounding the relationship between salary structure and important district outcomes.

Do Districts Frontload Salary Schedules?

We argue above that districts maximize the returns to their overall salary expenditures by frontloading their salary schedules to attract and retain high-quality teachers. However, the limited research that exists on the structure of uniform salary schedules suggests that many district salary schedules are not

in fact structured to give the largest returns early in the teaching career. In their study of teacher salary schedules in New York State, Monk and Jacobson (1985) find that the structure of many New York State school district salary schedules changed during the late 1970s and early 1980s such that maximum salaries increased at a faster rate in percentage terms than did starting salaries. Similarly, Murnane et al. (1987) analyze changes in Michigan school districts' teacher salary schedules between 1970 and 1980. They find that real (1970) starting salaries decreased by an average of 20% over the decade, whereas real maximum salaries decreased by only 15%.

Lankford and Wyckoff (1997) find that, over a 25-year span, New York districts increased the real salaries of veteran teachers by an average of almost 15% while increasing the real salaries of novices by only about 3%. Their conservative estimates find that 47% of districts that backloaded actually saw a decline in the real salaries of novice teachers, with an average decline of 7.6% compared with an average increase to veteran teachers of 10.9%. In supplemental analysis using a selected national sample of salary schedules collected by the American Federation of Teachers (AFT) from the largest AFT local unions, Lankford and Wyckoff find that very few of the districts had frontloaded schedules. Vigdor (2008) calculates that North Carolina school districts pay highly experienced teachers an average premium of almost 70% relative to novice teachers.

Our analysis of the SASS data shows evidence of much more frontloading than prior studies. Table 1 shows that half of districts in the full 1999-2000 SASS sample frontload. The median district frontloads just slightly such that novice teachers see a yearly increase in salary of 1.07% more than veteran teachers' returns to experience, which is a roughly linear schedule. However, the mean district in our sample backloads its salary schedule, with an average relative experience premium in our sample of 47%. Districts that backload tend to do so with larger relative premiums than do districts that frontload; the average backloading district provides 134% greater yearly returns to veteran than to novice teachers, whereas the average frontloading district provides only 37% greater yearly returns to novice than to veteran teachers. Also, there is far greater variation in the extent of backloading than frontloading in the SASS sample. Frontloading districts range from rewarding their novice teachers with 4% greater yearly returns relative to veteran teachers to 88% greater yearly returns. In contrast, backloading districts provide their veteran teachers with 1% to 337% greater yearly returns than are provided to novice teachers.

Table 1. Comparison of Backloaded and Frontloaded Districts, 1999-2000 SASS Sample

	Backloaded districts (N = 2,250)		Frontloaded districts (N = 2,230)		t test
Variable	*M*	*SD*	*M*	*SD*	*p* value
District-level variables					
Relative experience premium (%)	134	43	–37	25	***
Salary, BA, no experience (US$)	25,964.55	4,292.98	25,818.26	3,661.37	
Salary, BA, 10 years experience (US$)	32,166.36	6,341.20	35,703.09	7,444.00	***
Salary, MA, no experience (US$)	28,431.12	4,781.25	28,151.44	4,120.53	
Salary, MA, 20 years experience (US$)	44,986.91	10,934.76	43,104.95	8,901.14	***
π^V	1,035.40	530.75	565.67	292.53	***
π^N	620.18	307.62	968.82	443.29	***
Per-pupil expenditures (US$)	8,804.17	4,523.06	8,477.00	3,669.29	*
District size (enrollment)	3,279	12,957	3,026	15,398	
% Black	0.07	0.17	0.06	0.14	**
% Hispanic	0.07	0.15	0.1	0.2	**
% free/ reduced-price lunch	0.35	0.24	0.4	0.24	***
Urban	0.05	0.22	0.07	0.25	
Suburban	0.42	0.49	0.36	0.48	**
Rural	0.53	0.5	0.57	0.5	*
Fraction population with BA+	0.07	0.04	0.06	0.04	

*p < .10. **p < .05. ***p < .01.

Given that there is substantial variation in the shape of district salary schedules, we explore what sorts of districts choose to backload rather than frontload salary schedules. Table 1 outlines the differences between backloaded and frontloaded districts among potentially important district-level factors. We find that, unsurprisingly, there are significant differences between frontloading and backloading districts along the entire pay schedule. While

backloaded and frontloaded districts provide statistically significantly different base pay for teachers, statistical significance is a byproduct of the large sample size; in substantive terms, the values of the CWI-adjusted base pay are very similar (approximately US$26,000) in the 1999-2000 school year. However, frontloading districts pay teachers with 10 years of experience and a bachelor's degree approximately US$3,500 more than backloading districts do but pay teachers with 20 years of experience and a bachelor's degree approximately US$1,900 *less* than do backloading districts. The average yearly growth in salaries for "novice" (teachers with 10 years or less experience) and "veteran" teachers (teachers with more than 10 years of experience) also significantly differ. Of course, these patterns are expected given that we construct our measure of backloading precisely to reflect differences in districts' compensation structures. Nonetheless, the degree to which novice and veteran salaries differ in districts that backload and frontload is substantively important. These relationships are shown in both Table 1 and Figure 2.

Perhaps of greater interest, Table 1 shows that backloading and frontloading districts also vary by other important district-level characteristics. Backloading districts spend approximately US$300 more per pupil than do frontloading districts. Backloaded districts also have a slightly greater proportion of Black students but a lower proportion of Hispanic students and students who qualify for the free- or reduced-price lunch program. A greater proportion of backloaded districts are found in suburban areas, and fewer are found in rural areas. Given that these differences may affect student performance on standardized tests, we include controls for urbanicity, base salary, per-pupil expenditures, and the minority and poverty makeup of the schools in the school performance models presented in the next section.[3]

The Relationship Between Salary Structure and Student Achievement

To test the hypothesis that frontloading of salary schedules is associated with better district outcomes, we link our salary schedule structure measures to student achievement data that we obtain from the National Longitudinal School-Level State Assessment Score Database. These data, collected by the United States Department of Education and the American Institutes for Research, contain information on how students score relative to statewide benchmarks for multiple grades in five subject areas, including math and reading, on which we focus in our analyses. Although data are available for all 50 states and the District of Columbia, many states are missing data for

one or more of the subject-grade combinations in the 1999-2000 academic year because of individual states' data suppression rules and/or inconsistencies with which states tested or reported scores across grades and subjects.[4] We use data on fourth and eighth grades so that we may analyze effects in elementary and middle schools. Sample sizes in the student achievement analyses range from nearly 800 schools in 15 states to just 380 schools in 12 states.[5] As our use of test score data yields a subsample of the SASS data that is no longer nationally representative, we examine the differences between the remaining schools and districts in the sample and those that are excluded due to lack of test score data. We find that, although there are statistically significant differences in some observable characteristics between the two groups of districts, there are few substantive differences. The notable exceptions are that there are more "regular" (nonspecial) schools in the included sample and that the districts included in our sample have higher proportions of Hispanic students (16% of students in the sample are Hispanic as opposed to only 9% in the excluded districts).[6]

We focus on cut-point assessment data, that is, the percentages of students in the school who achieve basic, proficient, and advanced on state assessments. We make this choice for two reasons. First, these data represent the cut points used in most states' and the national accountability policies, thus tying them to district decision making that stems from accountability policy incentives. Second, scale score data from which we would be able to determine school-level test score gains are rarely available in the 1999-2000 data.[7]

There are important methodological challenges that arise from the use of these cross-state data. States vary both in the kinds of tests administered and in where they set the cut scores. Moreover, they do not always label the cut points the same in their accountability reporting systems. We address these issues using two strategies. First, for states that do not use the easily recognizable labels of *proficient*, *basic*, and *advanced*, we examine the state's data to determine comparable cut points. For example, the 2000 Arizona cut points are labeled *percent approaching standards*, *percent meeting standards* and *percent exceeding standards*, which we translate to *percent basic*, *percent proficient* and *percent advanced*, respectively. Where there was confusion regarding the correct analogous cut points, we went to the state departments of education websites to ascertain the correct comparable labels. Second, we control for unobservable state heterogeneity that may affect our outcome measures, such as test instrument difficulty or choice of test score category cut points, by using state fixed effects in the achievement regressions. Appendix Table A1 provides summary statistics of these assessment data for

fourth and eighth grades in both math and reading for the proficient, basic, and advanced assessment cutoffs.

We estimate a series of equations of the following form:

$$A_{I,D,S} = (\text{RELATIVE EXPERIENCE PREMIUM})_D\beta + S_I\delta + D_D\theta + \gamma_S + \nu_{I,D,S} \quad (3)$$

where A is a measure of achievement in school i. Subscripts d and s denote districts and states, respectively. We estimate Equation 3 using the percentages of students falling beyond each of the three cutoff measures for A. First, we measure achievement as the percentage of students in a given grade who pass the proficiency benchmark on the state exam. We examine this cutoff first because it is usually the most important benchmark for determining compliance with accountability program mandates. However, because we also wish to consider the implications of salary schedule structure for students elsewhere in the achievement distribution, we consider two additional benchmarks as well. In a second set of models, we measure A as the percentage of students scoring *below* the basic proficiency cutoff, which represents the lowest achieving students in a district. In a third set, we measure A as the percentage of students scoring above the *advanced* benchmark or those who achieve the highest marks in the district.

The primary coefficient of interest is β, which measures the association between the relative experience premium and the student achievement measure. Models also include S_i, a vector of school characteristics, and D_d, a vector of district characteristics to mitigate any biases that result from the possible endogeneity of our relative experience premium measure, to control for nonpecuniary factors that might influence teachers' decisions to stay or to leave a district independent of salary or salary schedule, and to hold constant other contextual and organizational influences on student achievement. School-level control variables are the fraction of Black students in the school, the fraction of Hispanic students in the school, the fraction of students receiving free or reduced-price lunch, and an indicator for whether the school is a regular school (i.e., not alternative, vocational, or focused on special education). These measures are gathered from the SASS questionnaires or from NCES Common Core of Data (CCD) for 1999-2000. Appendix Table A2 provides summary statistics of the school- and district-level control variables included in our analyses.

We also include a lagged measure of student achievement to control for student performance in the previous year, though excluding the lag produces similar results.[8] District-level variables include total district enrollment, the

beginning teacher's salary, indicators for whether the school is in an urban or rural location (suburban excluded), the fraction of the population in the district that has a bachelors' degree or higher, and total per pupil expenditures. This last variable is included to guard against a spurious correlation between achievement and the relative experience premium that may arise if the experience premium is correlated with total resources and resources influence achievement. We are not able to control for teachers' salary expenditures per pupil because the measure is too highly correlated with total per-pupil expenditures. However, we do control for the district's beginning teacher salary since we want to distinguish the effect of salary structure from any effects of overall compensation. As mentioned, salary and spending measures are adjusted by the 1999-2000 Comparable Wage Index to ensure that our results are not biased by unobserved differences in the costs of labor and other inputs across districts. We also include state fixed effects to control for unobservable state heterogeneity that may affect our outcome measures.

Results

Table 2 displays the results for the estimation of Equation 4 by ordinary least squares (OLS) with the percentage passing the proficiency benchmark as the dependent variable. The table has four columns. The first 2 measure proficiency in math in Grades 4 and 8. The next two measure proficiency in reading. All standard errors are clustered at the district level to account for correlation among schools in the same district. Estimates also employ SASS school-level survey weights and state fixed effects.

Across all columns in Table 2, the coefficient on the relative experience premium variable is negative and is statistically significant at least at the .1-level. We also run the same analysis on the proportion of students passing at the proficient level in 3rd, 5th, 6th, 7th, 10th, and 11th grades in reading and in math, which we omit from the article in the interest of space. All of these additional models estimate a negative coefficient for the relative experience premium variable, significant at least at .1-level in the six models representing fifth, sixth, and seventh grades.[9] These findings indicate not only that backloading is negatively associated with the proportion of students reaching the proficiency benchmark but that this relationship is consistent across the elementary, middle, and high school grades. The associations are relatively small in magnitude, however. For example, the coefficient in Model 2 implies that a one standard deviation increase in the relative experience premium—the returns to 1 year of experience for a veteran teacher in her second decade of teaching as opposed to a novice teacher in her first decade of teaching—is

Table 2. The Relative Experience Premium and Percentage Achieving Math and Reading Proficiency, 2000

	Math		Reading	
	(1)	(2)	(3)	(4)
Dependent variable	Percent proficient, Grade 4	Percent proficient, Grade 8	Percent proficient, Grade 4	Percent proficient, Grade 8
Relative experience premium (% change)	−0.003** (0.001)	−0.002* (0.001)	−0.002*** (0.001)	−0.002*** (0.001)
Percent proficiency, 1999	0.580*** (0.045)	0.738*** (0.051)	0.624*** (0.043)	0.651*** (0.044)
District characteristics				
District size (in 1,000s)	−0.004 (0.002)	−0.003 (0.003)	−0.004 (0.003)	0.005 (0.005)
Urban district	−1.087 (1.250)	−1.291 (1.413)	0.170 (1.081)	−2.354* (1.423)
Rural district	1.515 (1.506)	−0.027 (1.413)	−0.840 (1.470)	−1.867 (1.779)
Per pupil expenditures (in US$1,000s)	0.071 (0.200)	−0.217 (0.149)	−0.003 (0.206)	0.183 (0.299)
Beginning teacher's salary (in US$1,000s)	0.389* (0.214)	−0.039 (0.208)	0.212 (0.212)	0.150 (0.215)
Fraction population with BA+	49.959*** (15.864)	20.432 (17.312)	11.959 (13.447)	15.846 (16.935)
School characteristics				
Fraction Black students	−4.318 (2.934)	−8.407*** (3.126)	−7.638** (3.102)	−7.592*** (2.740)
Fraction Hispanic students	−2.672 (2.880)	−3.892 (2.994)	−5.274** (2.260)	−4.927* (2.614)
Fraction students free/reduced-price lunch eligible	−11.627*** (2.937)	−5.330* (3.076)	−9.856*** (2.441)	−5.020* (2.813)
Regular (nonspecial) school	−0.155 (2.488)	2.563 (2.064)	−1.074 (1.567)	1.722 (1.486)
Constant	17.411*** (6.032)	19.056*** (6.782)	25.831*** (6.693)	19.773*** (6.000)
Observations	800	730	830	570
Number of states	15	20	15	16
Adjusted R^2	.849	.909	.845	.884

Note: Clustered standard errors are in parentheses. Models include state fixed effects. All sample sizes rounded to the nearest ten per NCES nondisclosure rules. Models run using 11th-grade data resulted in much smaller samples with consistent results. We run all analyses excluding all potential outlier districts, and our results retain similar magnitudes and directions. Results are available from the authors upon request. States included in Model 1: AR, CA, MA, ME, MO, NY, NC, OH, RI, SC, TX, VT, WA, WI, WY; States included in Model 2: CA, DE, FL, IL, IN, KY, MA, MD, ME, MO, NY, NC, OR, RI, SC, TX, VT, WA, WI, WY; States included in Model 3: AR, CA, CO, FL, KY, ME, NC, OH, RI, SC, TX, VT, WA, WI, WY; States included in Model 4: CA, DE, FL, IL, IN, MD, ME, NC, OR, RI, SC, TX, VT, WA, WI, WY.
*$p < .10$. **$p < .05$. ***$p < .01$.

associated with approximately 1.25 percentage points fewer students meeting the proficiency standard in eighth-grade math, all else equal.

We can compare the findings in Table 2 to those from the models in Table 3, which show the results for the estimation of Equation 4 using the percentage of students who score below the basic benchmark and the percentage who score at or above the advanced benchmark in the top and bottom panels, respectively. The coefficients are in the expected directions in all of the models and are significant in six of the eight.

As expected, we find that schools in districts with greater relative experience premiums for veteran teachers (i.e., those with greater backloading of their salary schedules) have *more* students who score below the basic benchmark on both math and reading assessments. A school with the same characteristics with a relative experience premium measure one standard deviation above the mean—a district with a relatively backloaded schedule—would have 1.67 percent more fourth graders failing to meet even the basic standard on either math or reading than if it was located in a district with a relative experience premium measure one standard deviation below the mean—a district with a relatively frontloaded schedule. A similar comparison for fourth-grade math from Model 7 reveals that the difference between one standard deviation below and above the mean for the relative experience premium measure is associated with almost 1% more students meeting the advanced standard.

Results across Tables 2 and 3 are consistent. Across the board, no matter the grade, more students achieve and pass the proficiency cut point and fewer students fail to pass the basic cut point as districts frontload their salary schedules to a greater extent. Although the results for the advanced level are less precise, the consistent direction of the relative experience premium coefficient suggests that the more frontloaded the salary schedule, the greater proportion of students achieve "advanced" status on state assessment tests, holding school and district characteristics constant.

Why Might School Districts Backload Salary Schedules?

The results of the last section show that districts that structure their salary schedules to reward experience gains for novice teachers to a greater extent than they reward experience gains for veteran teachers have larger student achievement than those who do not, suggesting that districts may experience productivity gains from structuring teacher pay in this manner. Yet we observe

Table 3. Relative Experience Premium and Math and Reading Percentage Below Basic and Advanced, 2000

	Math		Reading	
	(1)	(2)	(3)	(4)
Dependent variable	Percent below basic, Grade 4	Percent below basic, Grade 8	Percent below basic, Grade 4	Percent below basic, Grade 8
Relative experience premium (% change)	0.002** (0.001)	0.001*** (0.000)	0.002** (0.001)	0.001*** (0.000)
Percent below basic, 1999	0.615*** (0.047)	0.697*** (0.049)	0.570*** (0.053)	0.837*** (0.055)
Constant	10.484** (5.045)	17.685*** (4.660)	5.955 (5.314)	–0.189 (4.767)
Observations	510	470	550	310
Number of states	10	13	10	9
Adjusted R^2	.797	.869	.803	.894
	(5)	(6)	(7)	(8)
Dependent variable	Percent advanced, Grade 4	Percent advanced, Grade 8	Percent advanced, Grade 4	Percent advanced, Grade 8
Relative experience premium (% change)	–0.001* (0.001)	–0.001 (0.000)	–0.002*** (0.001)	–0.001 (0.001)
Percent advanced, 1999	0.679*** (0.044)	0.845*** (0.062)	0.612*** (0.046)	0.836*** (0.123)
Constant	5.747 (4.688)	11.533** (5.296)	10.146** (4.401)	3.476 (6.805)
Observations	590	540	620	380
Number of states	12	16	12	12
Adjusted R^2	.833	.849	.834	.880

Note: Clustered standard errors are in parentheses. Models include state fixed effects and all district and school control variables included in the previous models. All sample sizes rounded to the nearest ten per NCES nondisclosure rules. Models run using 11th-grade data resulted in much smaller samples with consistent results. We run all analyses excluding all potential outlier districts and our results retain similar magnitudes and directions. Results are available from the authors upon request.
*$p < .10$. **$p < .05$. ***$p < .01$.

that nearly half of all districts in the national SASS sample backload to some extent. In this section, we consider an explanation for this observation suggested by other studies (Ballou & Podgursky, 2002; Lankford & Wyckoff, 1997; Monk & Jacobson, 1985)—that districts backload in response to political power exercised by rent-seeking veteran teachers or teachers' associations. The idea is that veteran teachers benefit more from backloading and thus might seek to influence district compensation policy in this direction.

We investigate teacher political power using two measures of the presence of teachers' unions in our analyses. First, we use a district indicator from the SASS data set of whether a district collectively bargains its teachers' contract, "meets and confers" with a teachers' association regarding the contract, or does not bargain at all with a teachers' association. Second, we compare the mean magnitude of backloading in states with and without laws that permit collective bargaining. If teachers' unions influence districts to backload salary schedules, then we should see that districts that collectively bargain have salary schedules that are backloaded to a greater degree than those who do not and that states in which collective bargaining is legal also have more backloaded salary schedules, on average, then do states in which collective bargaining is not legally required.

Table 4 shows evidence of exactly these patterns. Using simple *t* tests, the top panel of the table compares the average relative experience premiums of districts that label themselves as "no bargaining," "meet and confer," and "collective bargaining" districts. The bottom panel compares the average relative experience premiums in states that do and do not have collective bargaining laws for teachers. We find a positive relationship between the degree of union representation in districts and the magnitude of the relative experience premiums awarded to veteran teachers relative to novice teachers. Districts with collective bargaining contracts have relative experience premiums more than twice as large as districts without collective bargaining, a difference that is statistically significant at the .05-level. Districts who operate under meet-and-confer rules but who do not employ formal collective bargaining have relative experience premiums that fall in between the other two kinds of districts. Meet-and-confer districts, on average, have salary schedules with relative experience premiums that are about 77% larger than those for no bargaining districts, though this difference is not statistically significant.

In addition, we find that the relative experience premium awarded to veteran teachers is significantly smaller in states that do not allow collective bargaining as compared with states that do permit collective bargaining for teachers. In fact, the relative experience premium in the typical district in a

Table 4. Relative Experience Premiums by Collective Bargaining Status

	M	Difference from "No Bargaining" group	*p* value
District-level status			
No bargaining	19.39	—	
Meet-and-confer	34.41	15.02	.36
Collective bargaining	49.44	30.05	.05
State-level status			
No CB law for teachers	12.57	—	
CB law for teachers	45.74	33.17	.05

Note: $N = 4{,}170$ schools.

state that permits collective bargaining for teachers is 3.6 times as large as the relative experience premium in the average district in a state that does not permit collective bargaining. This difference is statistically significant at the .05-level.

In other words, districts that bargain collectively with teachers' unions are *more* likely to have backloaded salary structures that are inconsistent with behaviors that would attract and retain high-quality faculty. This evidence supports the conclusion that the political power of more experienced teachers, particularly as expressed through teachers' associations and unions, may disrupt districts from being able to choose salary structures that best fit their needs. The distortion of optimal salary expenditures may be one mechanism for explaining the negative relationship other work has observed between teacher union activity and student performance (e.g., Hoxby, 1996).

Conclusions and Policy Implications

This article asserts multiple rationales for why the frontloading of teacher salary schedules is likely to benefit school district outcomes. Relative to a backloaded salary structure, a frontloaded schedule should allow for a more efficient allocation of district budgets by saving funds that might otherwise yield little return when allocated to veteran teachers' salaries and putting those dollars to more productive, student achievement-enhancing uses. Consistent with these claims, we show that schools in districts with frontloaded salary schedules perform better in terms of student proficiency in reading and math at the fourth- and eighth-grade levels. Furthermore, we

provide evidence that students in both the left and right tails of the achievement distribution similarly benefit from the frontloaded salary schedule, with frontloaded districts having fewer students below the basic cutoff and more students above the advanced cutoff. Despite the theoretical appeal of frontloaded schedules, we find that only about half of districts in our sample actually structure their salary schedules in this manner, with the other half rewarding senior teachers to a greater extent than early career teachers. We find support for veteran teacher political power as one explanation for the backloaded salary schedule.

While the degree to which we can draw strong policy implications from these findings is limited by their correlational nature, they do suggest several avenues for further experimentation and analysis. The first is the possibility that districts that currently backload their teacher salary schedules may benefit from shifting to a more frontloaded compensation plan. However, there may be challenges in doing so. In many states and districts, the specific structure of teacher compensation is a direct result of formal or informal bargaining between district representatives and teachers' associations or unions. Given these collective bargaining constraints, drastically reshaping the salary schedule may be difficult, and districts may be required to steepen the pay-experience slope for novice teachers at a faster rate than they diminish returns for later experience gains. Thus, shifting to a frontloaded salary schedule may incur increased salary costs in the short term as districts increase experience returns to novice teachers while less drastically changing the pay structure for more senior teachers. In the long term, however, salary frontloading has the potential to reduce overall teacher salary costs by eliminating low-return rent payments to veteran teachers who are unlikely to leave the profession.

Districts who do not recognize the potential efficiency gains from frontloading may too readily agree to allocate scarce dollars to the veteran teachers whose interests disproportionately are represented in the bargaining or meet-and-confer process. Even in the absence of bargaining, district administrators may see increasing salary gains from experience late in the teaching career as an opportunity to reward teachers for their long service and commitment to the district. Our results suggest that these allocations are inefficient and ultimately associated with lower student performance. Based on these results, we would caution district administrators and bargaining representatives from negotiating gains to veteran teachers into uniform teacher salary schedules at the expense of salary gains to early career teachers.

Frontloading district salary schedules may be a cost-effective and practical reform to promote teacher quality within the confines of the current pay

model until more controversial plans such as individual teacher merit pay can meet substantial implementation challenges and gain greater political traction. Our results suggest that nontrivial gains in school productivity might be realized with a less radical restructuring of many districts' compensation plans. Moreover, we suggest that districts who do implement pay-for-performance plans may see greater gains by implementing those plans in conjunction with a restructuring of the salary schedule base to allocate greater salary gains to the least experienced teachers.

There are numerous ways that districts can achieve a more frontloaded salary schedule besides reconstituting the schedule itself to provide greater increases in salary for early stage teachers. For example, districts can explore the use of loan forgiveness programs that reward new teachers by repaying their student loans for the first multiple years of teachers' careers. Alternately, districts can offer signing bonuses for new teachers who join the district, can place new teachers at higher levels on the salary schedule based on out-of-district experience, and can provide retention bonuses over the first few years of their careers. Such creative ways of rewarding early stage teachers may be more politically feasible than providing raises for novice teachers both because they can be rescinded in times of financial turmoil and because they may be more acceptable to the unions who collectively bargain on behalf of teachers. In addition, districts can use such bonuses or loan forgiveness programs to target specific kinds of teachers, such as those who teach high-need subjects or in hard-to-staff schools, or those who meet specific academic criteria.

Of course, the move to frontloading from a backloaded system presents potential challenges. In particular, as reducing veteran teachers' experience increments to fund increases for novice teachers may be infeasible, especially in the politicized environment in which teacher salaries are decided, moving to a frontloaded structure might mean increasing overall compensation costs. Higher overall compensation, with greater compensation realized earlier in the career, could have implications for pension costs as well.

Our use of the SASS data set in conjunction with school-level assessment scores from across the country allows us to examine the structure of salary schedules in districts across multiple states and thus helps us build on earlier work by extending models that have generally focused on limited samples. Three important caveats are in order. First, the sample of SASS districts that we are able to match to achievement test measures is not representative of all districts in the United States. An analysis of schools and districts excluded from our samples due to a lack of matching test score data shows that our

samples contain districts that look systematically different from the population of U.S. school districts as a whole. Importantly, however, our independent variable of interest, the relative experience premium, does not differ significantly between the in-sample and out-of-sample districts. Also, in general our models include districts that vary in terms of geographic dispersion and collective bargaining status. The use of district and school control variables in the analyses further alleviate concerns about comparability of the schools in our sample and other schools. Nevertheless, some unobservable differences remain without doubt, and our results should be understood in this context.

Second, the SASS data on salary structure are coarse. They only allow us to calculate the relative experience premium for teachers in their first 10 years and in their second 10 years of teaching. Although many teachers continue teaching well beyond their 20th year, we cannot identify how districts shape their salary schedules for teachers' last 10 or 20 years of teaching. As a result, we cannot model the impacts of further backloading in the final years of salary schedules on student outcomes.[10] Similarly, they prevent us from examining more fine-tuned differences in schedules within the first 20 years, which may be important as districts alter schedules by compressing the number of steps over time.

Third, the cross-sectional nature of our data constrains us from generating causal conclusions about the impacts of salary structure on teacher turnover and student achievement. That we have only a single year of data also forces us to make the assumption in our models that teachers are recruited to a district and retained in that district based on the present salary schedule. While this is a strong assumption, it is unlikely that districts shift their salary schedules from backloaded to frontloaded, or even substantially along that continuum, over the course of teachers' decision-making time frames. Nonetheless, these analyses would benefit from longitudinal data of the sorts typically found in individual state data sets. We hope that this study serves as a starting point for future work to delve deeper into the causal relationships between salary schedule structure and important student and district outcomes.

In general, there is a need for more research on the budgetary and other political implications of restructuring salary schedules and using alternatives to single salary schedules. Proposals to implement policies that would change traditional teacher salary schedules must be evaluated not only for their potential success in reducing turnover of high-quality teachers and increasing student achievement but also for their feasibility. For many districts, the

suggestion that they shift to new compensation policies that force them to spend even more of their limited budgets on teacher compensation, especially without evidence of significant returns to such compensation plans, is impractical. Findings from this study may help policy makers and district administrators consider alternatives to their current compensation plans that are more immediately feasible and still help to enhance student achievement.

Appendix

Table A1. Summary Statistics, 2000 Assessment Data

Test score measure	Backloaded districts					Frontloaded districts				
	N	*M*	*SD*	Min	Max	*N*	*M*	*SD*	Min	Max
% proficient										
Math										
4th grade	650	48.68	28.72	0	100	590	52.62	29.90	0	100
8th grade	670	40.77	27.60	0	100	520	42.40	30.82	0	100
Reading										
4th grade	610	58.77	22.98	0	100	560	61.97	22.07	0	100
8th grade	570	52.43	26.11	0	100	430	58.40	27.18	0	100
% below basic										
Math										
4th grade	450	25.73	20.83	0	100	380	21.59	22.31	0	97
8th grade	460	35.91	24.78	0	100	360	32.70	24.67	0	100
Reading										
4th grade	380	21.62	17.41	0	100	340	16.79	15.09	0	87
8th grade	330	24.17	20.49	0	100	240	16.13	18.58	0	100
% advanced										
Math										
4th grade	510	18.83	20.46	0	100	400	18.92	19.07	0	89
8th grade	510	13.08	17.17	0	100	390	14.31	17.78	0	100
Reading										
4th grade	480	18.25	17.01	0	100	380	15.79	15.04	0	81
8th grade	420	18.99	21.37	0	100	310	22.34	20.38	0	87

Note: All sample sizes rounded to the nearest ten per NCES nondisclosure rules.

Table A2. Summary Statistics of School and District Control Variables

	4th grade ($N = 800$)				8th grade ($N = 730$)			
	M	*SD*	Min	Max	*M*	*SD*	Min	Max
District-level variables								
Relative experience premium (% change)	41.11	354.78	−100.00	5,403.66	0.77	548.37	−100.00	6,444.57
District size (in 1,000s)	55.46	177.69	0.07	1,068.23	59.77	156.14	0.07	1,068.23
Urban district	0.35	0.48	0	1	0.25	0.44	0	1
Rural district	0.24	0.43	0	1	0.37	0.48	0	1
Per pupil expenditures (in US$1,000s)	8.50	2.58	4.88	43.34	9.08	3.04	5.00	33.65
Beginning teacher's salary (in US$1,000s)	28.85	3.04	20.32	39.60	28.90	3.25	21.27	40.59
Fraction of population with BA+	0.07	0.04	0.01	0.28	0.07	0.04	0.01	0.28
School-level variables								
Fraction Black students	0.15	0.23	0	1	0.16	0.26	0	1
Fraction Hispanic students	0.21	0.28	0	1	0.15	0.24	0	0.99
Fraction students free/reduced-price lunch eligible	0.47	0.28	0	1	0.41	0.26	0	1
Regular (nonspecial) school	0.95	0.22	0	1	0.91	0.29	0	1

Note: Samples summarized here correspond to the samples used in the largest-*N* regression model for each grade. Per-pupil expenditures and beginning teacher's salary figures are CWI-adjusted using the NCES 1999-2000 CWI data set. All sample sizes rounded to the nearest ten per NCES nondisclosure rules.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Unfortunately some of the district questionnaire items necessary to construct our salary schedule measures were not included in the more recent version (2003-04) of the SASS data. Perhaps worse, these measures also were excluded from earlier versions of SASS, making it impossible for us to look at the changes in experience premium over time.
2. We assume a uniform rather than a percentage increase associated with a master's degree in our measure of salary schedule. We make this assumption for two reasons. First, it is not uncommon to see uniform pay increases associated with master's degrees in salary schedules. The authors' examination of 490 California school districts' certificated personnel collective bargaining agreements shows that 47% of districts provide a set pay increase for teachers who hold a master's degree. The second reason we use a uniform rather than percentage allocation for raises associated with master's degrees is because of the data limitations of SASS, which do not allow us to directly measure master's degree raises. If the districts included in our sample use percentage rather than uniform pay increases to reward teachers with master's degrees, our measure of salary schedule shape will be incorrect for those districts.
3. We also compared the two types of districts with respect to observable teacher characteristics, including experience, education level, demographic profile, and whether they hold a regular teaching certificate. These comparisons (untabulated) showed no substantive differences in the characteristics of teachers employed.
4. Each state in the NLSLSAS database has its own data suppression rules, rendering some school and district aggregate test scores unreported in the data set. For example, Arkansas required that aggregate test score data cannot be released for students in groups with fewer than five students, and in Delaware data cannot be reported for students in groups with fewer than 15 students.
5. Sample sizes in this study rounded to the nearest ten per National Center for Education Statistics nondisclosure rules.
6. Results are available from the authors upon request.

7. We also run identical analyses using standardized achievement scores as our measure of average student achievement. These specifications yield results consistent with those from our cut-point analyses. However, our estimates suffer from decreased precision due to substantially smaller samples of schools with available scale score data. These results are available from the authors upon request.
8. In the interest of space, we do not present both sets of models. The level models (not including lagged student performance) are available from the authors upon request.
9. Results are available from the authors. The experience premium variable is negative and significant at the .01-level for the fifth-, sixth-, and seventh-grade percentage proficient reading models and for the fifth-grade percentage proficient math model. The experience premium variable is negative and significant at the .05-level for the sixth-grade percentage proficient math model, and at the .10-level for the seventh-grade percentage proficient math model.
10. The truncation of the salary schedule data in SASS should not greatly affect our main results for two main reasons. First, unless districts shift their salary structures enough for teachers with more than 20 years of experience to substantively change π^{Y} such that it is no longer greater than π^{N}, backloading districts will remain backloaders. There is no reason to believe that districts that backload will stop or reverse this practice as teachers gain in experience. Second, the theory of teacher turnover behavior and districts' desires to reward gains in effectiveness is driven by assumptions about teachers in the early stages of their careers. Teachers already in the district for 20 years are highly unlikely to make attrition decisions based on salary concerns, and effectiveness returns to experience plateau after 4 or 5 years (Hanushek & Rivkin, 2007; Murnane et al., 1991; Rivkin et al., 2005). Therefore, the lack of available data about salary schedules for teachers with more than 20 years of experience should not greatly affect the implications of our results.

References

Ballou, D., & Podgursky, M. (2002). Returns to seniority among public school teachers. *Journal of Human Resources*, *37*, 892-912.

Clotfelter, C., Ladd, H., & Vigdor, J. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, *26*, 673-682.

Dolton, P. (1990). The economics of UK teacher supply: The graduate's decision. *Economic Journal*, *100*(4), 91-104.

Figlio, D. (1997). Teacher salaries and teacher quality. *Economics Letters*, *55*, 267-271.

Figlio, D., & Kenny, L. W. (2007). Individual teacher incentives and student performance. *Journal of Public Economics*, *91*, 901-914.

Goldhaber, D. D., & Brewer, D. J. (1997). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *Journal of Human Resources*, *32*, 505-523.

Hanushek, E., & Pace, R. (1995). Who chooses to teach (and why)? *Economics of Education Review*, *14*, 101-117.

Hanushek, E., & Rivkin, S. G. (2007). Pay, working conditions and teacher quality. *Future of Children*, *17*(1), 69-86.

Heutel, G. (2009). Testing implications of a tournament model of school district salary schedules. *Economics of Education Review*, *28*(1), 143-151.

Hoxby, C. M. (1996). How teachers' unions affect education production. *Quarterly Journal of Economics*, *111*, 671-718.

Ingersoll, R. M. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Education Research Journal*, *38*, 499-534.

Imazeki, J. (2005). Teacher salaries and teacher attrition. *Economics of Education Review, 24*(4), 431-449.

Lankford, H., & Wyckoff, J. (1997). The changing structure of teacher compensation, 1970-94. *Economics of Education Review*, *16*, 371-384.

Lavy, V. (2007). Using performance-based pay to improve the quality of teachers. *Future of Children*, *17*(1), 87-109.

Loeb, S., & Page, M. (2000). Examining the link between teacher wages and outcomes: The importance of alternative labor market opportunities and non-pecuniary variation. *Review of Economics and Statistics*, *82*(3), 393-408.

Manski, C. (1987). Academic ability, earnings and the decision to become a teacher: Evidence from the National Longitudinal Study of the High School Class of 1972. In D. Wise (Ed.), *Public sector payrolls*. Chicago, IL: University of Chicago Press; 291-316.

Monk, D. H., & Jacobson, S. L. (1985). The distribution of salary increments between veteran and novice teachers: Evidence from New York State. *Journal of Education Finance*, *11*, 157-175.

Murnane, R. J., & Olsen, R. J. (1989a). The effect of salaries and opportunity costs on duration in teaching: Evidence from Michigan. *Review of Economics and Statistics*, *71*, 347-352.

Murnane, R. J., & Olsen, R. J. (1989b). Will there be enough teachers? *American Economic Review*, *79*, 242-246.

Murnane, R. J., Singer, J., & Willett, J. (1987). Changes in teacher salaries during the 1970s: The role of school district demographics. *Economics of Education Review*, *6*, 379-388.

Murnane, R. J., Singer, J., Willett, J., Kemple, J., & Olsen, R. (1991). *Who will teach? Policies that matter*. Cambridge, MA: Harvard University Press.

Podgursky, M., & Springer, M. (2007). Teacher performance pay: A review. *Journal of Policy Analysis and Management*, *26*, 909-949.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, *94*, 247-252.

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay and student achievement. *Quarterly Journal of Economics*, *125*(1), 175-214.

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

Sawchuk, S. (2009, January 27). School leaders target salary reform toward newer teachers. *Education Week*. Retrieved from http://www.edweek.org/ew/articles/2009/01/28/19salary_ep.h28.html

Taylor, L. L., & Fowler, W. J., Jr. (2006). *A comparable wage approach to geographic cost adjustment (NCES 2006-321)*. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Vigdor, J. (2008). Scrap the Sacrosanct Salary Schedule. *Education Next*, *8*(4), 36-42.

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. New York, NY: The New Teacher Project.

Wright, P. S., Horn, S. P., & Sanders, W. L. (1997). Teachers and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, *11*, 57-67.

Zabalza, A., Turnbull, P., & Williams, G. (1979). *The economics of teacher supply*. New York, NY: Cambridge University Press.

Bios

Jason A. Grissom is assistant professor of public policy and education at Peabody College, Vanderbilt University. His research applies the perspectives of political science, public administration, and economics to the study of K-12 governance, including its policy, management, and political dimensions. Recent work has examined school board decision-making, principal effectiveness, and the labor market for school administrators.

Katharine O. Strunk is an assistant professor of education and policy in the Rossier School of Education, and by courtesy the School of Policy, Planning and Development at the University of Southern California. Her research centers on K -12 education policy, focusing on teacher labor markets, education governance, and district-level reforms.