

Changing the Education Workforce? The Relationships Among Teacher Quality, Motivation, and Performance Pay

tcrecord.org/Content.asp

by [Daniel H. Bowen](#) & [Jonathan N. Mills](#) - 2017

Background/Context: *With a growing body of evidence to support the assertion that teacher quality is vital to producing better student outcomes, policymakers continue to seek solutions to attract and retain the best educators. Performance-based pay is a reform that has become popular in K-12 education over the last decade. This strategy potentially produces positive impacts on student achievement in two ways: better alignment of financial incentives with desired outcomes and improved the composition of the teacher workforce. While evaluations have primarily focused on the former result, there is little research on whether the longer term implementation of these policies can attract more effective teachers.*

Purpose: *In this study we aim to provide evidence for potential long-term impacts that performance-based pay can have on the composition of the teacher workforce by addressing two questions: Does performance-based pay attract fundamentally different individuals, as measured by their risk preferences, to the teaching profession? Are stated preferences for a particular pay format correlated to measures of teacher quality?*

Research Design: *We apply methods from experimental economics and conduct surveys with 120 teachers from two school districts who have experienced performance pay. We compare the risk preferences of teachers hired under the two pay formats to test the hypothesis that performance-based pay attracts individuals with different characteristics to the profession. We also analyze teachers' survey responses on their preferences for performance-based pay to determine their relationships to two measures of teacher quality: student test-score gains and principal evaluations.*

Conclusions/Recommendations: *We find mixed results regarding the ability of performance-based pay to alter the composition of the teacher workforce. Teachers hired with performance-based pay in place are no different from their colleagues. However, teachers claiming to seek employment in districts with performance-based pay in place appear significantly less risk averse. Surprisingly, additional analyses indicate that teachers' value-added scores and performance evaluations do not predict a positive disposition towards merit pay. Thus, while these results indicate the possibility for performance-based pay to attract different individuals to teaching, they do not provide evidence that such change would necessarily improve the composition of the workforce. Policymakers should take this potential tradeoff into consideration when considering the expansion of performance pay policies.*

INTRODUCTION

Teacher quality is the most critical school-level input when it comes to improving student success. *High quality* teachers, as determined by measures such as value-added scores, are more likely to have their students attend college and enroll in higher-caliber institutions (Chetty, Friedman, & Rockoff, 2014). In terms of learning gains, more effective teachers produce as much as an additional year of student learning relative to less-effective educators (Hanushek, 2011). Moreover, an increase in teacher quality of one standard deviation for one academic year, according to value-added measures, would potentially increase an individual's cumulative lifetime earnings by \$39,000 (Chetty et al., 2014). These findings, which testify to the value of high-quality instruction, have led education reformers to explore policies to improve the

Related Articles

[This Month's Issue](#)



Related Discussion

[Post a Comment](#) | [Read All](#)

About the Author

- Daniel Bowen
Texas A&M University
[E-mail Author](#)
DANIEL H. BOWEN is an assistant professor in the Department of Educational Administration and Human Resource Development at Texas A&M University. His research primarily employs field experiments

to assess policies that influence teacher quality in addition to the role of "non-core" subjects and school-sponsored extracurricular activities, such as field trips, the arts, and school-

composition of the education workforce, both in terms of retaining and attracting effective teachers (Ballou & Podgursky, 1995; Finn, 2001; Fox, 1984).¹

One strategy for improving the caliber of the teacher workforce that has become increasingly popular with policymakers is shifting from the traditional, “step-and-lane” salary schedule to differentiated or “merit” pay (Wallis, 2008). Despite the numerous challenges that stem from constructing assessments and measures to accurately identify and quantify teacher effectiveness (Ballou & Springer, 2015), education policymakers and administrators continue to push forward with efforts aimed at improving the quality of the teacher workforce using financial incentives (Lavy, 2007).

The theoretical argument for performance-based incentives is that compensating teachers according to measurable outcomes better aligns pay and motivation in ways that should ultimately increase student achievement (Dee & Keys, 2004). However, while the primary aim and focus of performance-based pay so far has been to incentivize the current workforce to increase their efforts in the classroom, proponents of this reform contend that it also has the potential to, over time, attract and retain higher-quality employees to the education sector (Balch & Springer, 2015; Podgursky & Springer, 2007).

The theory and supporting empirical research from the field of labor economics suggest that the parameters and expectations of a career, especially with regard to anticipated income and compensation formats, have powerful influences over one’s chosen profession (Salop & Salop, 1976). From this perspective, advocates for strengthening the connection of performance-

based measures to teacher compensation believe that such reforms will better attract and retain individuals who are likely to be more effective in the classroom (Lavy, 2007). Under the step-and-lane system districts primarily compensate teachers according to their obtained degrees, additional certifications, and years of experience. This suggests that, all else equal, a “bad” teacher can make the same salary as a “good” one (Stronge, Gareis, & Little, 2006). In contrast, by linking wages to outcome measures of teacher quality, one might expect performance-based compensation to do a better job of attracting and retaining more-talented individuals to the profession (Lazear, 2000).

Public management scholars, however, contend that individuals drawn to careers in the public service sector have motivations that are typically at odds with the use of explicit, extrinsic rewards embedded in performance-based pay (Francois, 2000; Organ, Podsakoff, & MacKenzie, 2006; Piliavin, Grube, & Callero, 2002). Education is arguably not unique in this regard since many—if not most—teachers are entering the profession in part due to intrinsic, public-service motivations. Survey data support this position; “The consistent pattern has been that altruistic, service-oriented goals and other intrinsic sources of motivation are the primary reasons entering teacher candidates report for why they chose careers in teaching” (Brookhart & Freeman, 1992, p. 46). This driving motivation for entering the teaching profession may require important considerations for determining how to best reward and compensate educators. Specifically, external rewards crowd out this intrinsic public service motivation in two ways. First, intrinsically motivated individuals are likely to perceive rewards as external controls. Second, external rewards can reduce self-esteem when employees perceive them as strategic approaches that disregard their intrinsically-driven motivations (Frey, 1997).

sponsored sports. Recent publications: Bowen, D. H., Buck, S., Deck, C., Mills, J. N., & Shuls, J. V. (2015). Risky business: An analysis of teacher risk preferences. *Education Economics*, 23(4), 470–480.

Bowen, D. H., Greene, J. P., & Kisida, B. (2014). Learning to think critically: A visual art experiment. *Educational Researcher*, 43(1), 37–44.

- Jonathan Mills
Tulane University

E-mail Author
JONATHAN N. MILLS is a postdoctoral fellow at the Education Research Alliance for New Orleans at Tulane University. His current research focuses on the effects of school choice programs on student achievement and nonacademic outcomes, as well as the benefits and unintended consequences of college financial aid programs.
Recent

Changing how teachers are compensated could therefore attract and retain more effective individuals. However, such reforms may also have the unintended consequence of reducing the appeal of the profession to more intrinsically motivated individuals. This deterrence effect could offset the benefits of attracting individuals who potentially perform better on measures rewarded by performance pay systems, such as value-added assessments and principal evaluations. Given its theoretical potential, policymakers should consider empirical evidence regarding this tradeoff when pursuing dramatic changes to teacher compensation systems.

The purpose of this study is first to evaluate whether there is empirical evidence to support the claims for performance-based pay's potential to substantially alter the makeup of the teacher workforce. Furthermore, we

conduct analyses to examine if tying salaries to explicit performance-based measures conflicts with teachers' intrinsic public service motivations.

Few, if any, studies have examined this question to date, largely because performance-based pay interventions are typically implemented on too small a scale or over too limited time period to examine lasting impacts on the teacher workforce. In instances where field studies are not possible, laboratory-style research can help provide controlled settings for testing particular underlying hypotheses. Therefore, in order to examine whether performance-based pay can alter the composition of the workforce, we compare the revealed risk preferences of 120 teachers working in two school districts. These districts implemented performance-based pay policies a few years prior to data collection for this study, thereby providing the opportunity to compare the preferences of teachers seeking employment at these schools when performance-based pay were in place to those seeking employment when it was not.

We use a behavioral measure that captures individual risk preferences to identify changes in these districts' workforce compositions. There is empirical evidence to suggest that individuals who choose to work in the public sector (and specifically, education) tend to be significantly more risk averse (Bowen, Buck, Deck, Mills, & Shuls, 2015; Dohmen & Falk, 2011). Since performance-based pay intensifies employee wage uncertainty, our primary hypothesis is that increases in salary ambiguity will deter (attract) more risk-averse (risk-tolerant) individuals from seeking and obtaining employment at schools with these policies.

To test whether performance-based pay impacts intrinsic public service motivation, we analyze teachers' survey responses regarding the appeal of using performance pay to two measures of teacher quality as defined by these particular merit pay programs: their students' value-added scores and principal evaluations. From a financial standpoint, teachers with higher scores on these measures have more to gain from the implementation of a performance-based pay system. Therefore, we hypothesize that such teachers should be significantly more willing to support performance pay.

While we do not find that teachers hired during the implementation of a performance pay program differed substantially in revealed risk preferences

recent
publications:
Egalite, A. E.,
Mills, J. N., &
Wolf, P. J.
(forthcoming).
The impact of
targeted
school
vouchers on
racial
stratification
in Louisiana
schools.
*Education and
Urban Society.*

Mills, J. N.
(2013). The
achievement
impacts of
Arkansas
open-
enrollment
charter
schools.
*Journal of
Education
Finance, 38(4),*
320-342.

from those hired at other periods, teachers noting a preference for performance pay are significantly more risk loving. There is no significant difference in individuals' revealed risk preferences coinciding with the

implementation of a performance-based pay system. However, teachers who claimed to have been attracted to performance pay when seeking employment are more risk loving, suggesting that financial incentives could alter the composition of the workforce. We also find evidence of a negative relationship between value-added scores and the extent to which a teacher favors performance-based pay ($p = 0.09$). There appear to be no significant associations between teachers' pay or risk preferences and their year-end principal evaluations.

These findings do not necessarily refute or affirm the claim that performance-based pay reforms can substantially improve the quality of the teacher workforce. Instead, these results indicate that teachers who currently have the most to gain financially from performance pay reforms are no more likely to support these policies than the teachers deemed least effective. This finding could be attributed to performance-based incentives crowding out high-quality teachers' intrinsic motivations or possibly instilling the perception that their efforts are being belittled or controlled with extrinsic rewards. These results should concern policymakers considering the tradeoffs of implementing performance-based pay policies and their impacts on the teacher workforce.

TEACHER PAY IN THE UNITED STATES

The origin of the widely used K-12 single-salary schedule in the United States dates back to the early 20th century. Throughout most of the 19th century teaching was not commonly viewed as a career or professional occupation (Prostik, 1995). Men during this time would often teach as a source of additional income when free from farming duties. Women typically only taught in early adulthood, and the families of the town's children primarily compensated them with room and board (Spring, 2004).

The one-room schoolhouse gradually became obsolete as a result of the growth of cities from the Industrial Revolution, and school districts began to consolidate and bureaucratize the implementation of public education, eventually making graded schools the convention (Reese, 2011). Around this time, school districts began putting into action the use of salary schedules for teacher compensation, a pay format typically consisting of set wage minimums that could fluctuate substantially at the discretion of school-level administrators. While this system had clear advantages over the earlier arrangements, substantial gender-based pay disparities persisted (Prostik, 1995). In response to these wage gaps the Interborough Association of

Women Teachers successfully lobbied for legislation that required "equal pay for equal work in teaching" throughout the State of New York in 1911 (Prostik, 1995, p. 9). Single-salary schedules removed school-level administrator discretion over teachers' salaries by now making wages primarily based on degree attainment and years of service (Lipsky & Bacharach, 1983). During the 1920s and 1930s several big-city districts implemented single-salary schedules and this expansion continued until 97% of U.S. public schools had adopted the single-salary schedule by the middle of the 20th century (Sharpes, 1987).

The single-salary schedule maintained its position as the primary format for K–12 teacher compensation in the United States over the next half century. Nevertheless, even as it was becoming more commonplace, the single-salary schedule has consistently been subject to criticism. School administrators, for example, complained that they could no longer adjust an individual teacher's salary to reward or incentivize performance or address labor market realities (e.g., increasing salaries in response to outside job offers) (Podgursky & Springer, 2007). This concern has been supported with empirical evidence. Single-salary school systems are more likely to face challenges and inefficiencies with teacher recruitment, calling into question whether alternative systems might help districts with their employment challenges; "For these reasons, it is worth exploring different salary systems that would allow districts to address the labor market reality that people with different skills face different opportunity costs in the teaching profession" (Goldhaber & Player, 2005, p. 228).

The push for the more-modern conception of market-based differentiated pay for teachers was first issued in the *A Nation at Risk* (National Commission on Excellence in Education, 1983) report. Shortly after the report's release, President Reagan argued that, "Teachers should be paid and promoted on the basis of their merit and competence. Hard-earned tax dollars should encourage the best. They have no business rewarding incompetence and mediocrity" (Reagan, 1983). In response, several school systems implemented various merit pay policies and programs. In the ten years after the release of *A Nation at Risk*, six states incorporated teacher pay-for-performance incentives, but only one state, Arizona, had gone as far to make students' test achievement a central focus.²

More recently, former Secretary of Education Duncan and President Obama endorsed the development of merit pay systems for teachers and the Department of Education has since established the Teacher Incentive Fund, a grant program supporting the development of performance-based

compensation systems (Resmovits, 2011). In response, more states and large urban school districts have committed to implementing or expanding preexisting teacher merit pay programs (Smarick, 2011).

Performance-based pay incentives have also become increasingly popular among the general public. In a recent national survey, 57% of nonteacher respondents indicated that they were at least "somewhat in favor" of "basing part of the salaries of teachers on how much their students learn." In contrast, only 21% of teachers were at least "somewhat in favor", and 45% were "completely opposed" to such measures (Henderson, Peterson, & West, 2015, p. 10).

There are several plausible explanations for why teachers' views of merit pay do not align with that of the general public. Teachers might worry more about the reliability of tests and their abilities to accurately assess student performance (Ballou & Springer, 2015; Hart, 2015; Koedel, Mihaly, & Rockoff, 2015; Lavigne, 2014). They also could have concerns about the designs of performance incentives and how they may affect a school's environment, such as inducing competition that reduces efforts to collaborate (Goldhaber, DeArmond, Player, & Choi, 2008; Ritter & Jensen, 2010) or enticing teachers to cheat (Jacob & Levitt, 2003). Another consideration is that introducing rewards that are tied to assessments might cause teachers to feel a loss of autonomy in the classroom (DeArmon & Lind, 1999) and to more narrowly focus on

in the classroom (Popkewitz & Lind, 1989) and to more narrowly focus on objectives tied to test scores at the expense of outcomes that fit into a broader mission or understanding of the roles of education (Murnane & Cohen, 1986; Wilms & Chapleau, 1999). However, an explanation for teacher opposition that is often overlooked is the possibility that the design and mission of the K-12 teaching profession has attracted individuals to the field whom are fundamentally resistant to the application of explicit, performance-based incentives (Bowen et al., 2015).

EMPIRICAL EVIDENCE ON TEACHER MERIT PAY

Internationally, performance-based pay programs have generally been successful in incentivizing teachers to increase student achievement (e.g., Lavy, 2009; Muralidharan & Sundararaman, 2011). Moreover, the extent to which a country bases teacher pay on student performance has been positively associated with achievement on the Programme for International Student Assessment (PISA) (Woessmann, 2011). In contrast, effects have typically been more mixed in the United States. Studies have exhibited both null (Fryer, 2013; Goldhaber & Walch, 2012; Goodman & Turner, 2013; Marsh et al., 2011; Springer et al., 2012), and modest positive effects on student test-score gains (Balch & Springer, 2015; Figlio & Kenny, 2007; Fryer, Levitt, List, & Sadoff, 2012; Springer, Ballou, & Peng, 2014).

Teacher performance pay policy and program evaluations have largely focused on short-run academic outcomes, leaving questions of long term impacts on labor markets to the realm of theory (Dee & Keys, 2004; Lazear, 2004; Sawchuck, 2010). Performance-based pay could affect the teacher labor force through two channels:

Agents with different individual characteristics and personality feel attracted by different types of incentives. In this sense providing incentives in firms or organisations has two important effects, an incentive effect *per se* and a selection effect. Importantly, these effects need not be complementary. (Dohmen & Falk, 2010, p. F256)

This suggests performance pay programs could improve the motivation of the current teaching workforce while also attracting better teachers to the workforce. For example, more “high-caliber” college students might consider going into education if they anticipate higher potential earnings as a result of being rewarded for effectiveness rather than experience.

A few studies have attempted to examine the compositional aspect of merit pay. This research has provided preliminary evidence suggesting compensation format has influenced the makeup of the teacher workforce. According to survey data, high-caliber college graduates perceive the K-12 step-and-lane pay scale as a significant deterrent to entering the K-12 education workforce (Auguste, Kihn, & Miller, 2010). Moreover, districts implementing student performance-based pay policies attract teachers from more selective universities (based on the institution’s average incoming freshman SAT score) (Jones & Hartney, 2011).

Examining the relationships between school district environmental factors, teacher perspectives, and the adoption of performance-based pay programs may provide insights into obstacles for implementing such policies. For example, the availability of teacher performance data is positively correlated with the likelihood of a district adopting a merit-based pay policy, but teacher union presence is negatively associated with policy adoption (Goldhaber et al.,

2008). Moreover, union influence is strongly predictive of the likelihood of implementation, size, and scope of teacher performance-based pay programs (Ballou, 2001).

There is also growing evidence of heterogeneous preferences for performance pay among teachers with common factors underlying these different perspectives. For example, teachers are more reluctant to support pay reform policies when a greater amount of future earnings are tied to performance (Nadler & Wiswall, 2011). In addition, incoming teachers are typically more receptive to performance-based pay (Milanowski, 2007). These factors suggest teachers may not be inherently opposed to the concept of merit pay, but framing—particularly with regard to the weighting and timing of such programs—seems to play a major role in the likelihood of gaining approval.

THE ROLES OF LABOR ECONOMICS AND AGENT MOTIVATION

Different compensation or incentive formats attract employees according to personal preferences, characteristics, and skillsets (Salop & Salop, 1976). Oyer and Schaefer (2005) have illustrated this phenomenon with the common occurrence of new firms offering employees stock options rather than higher fixed wages. This approach tends to entice less experienced employees who are more optimistic about the firm's potential with the drawback of dissuading applicants who need a steadier, more secure paycheck. Empirical evidence supports these theoretical arguments. In a famous case study from labor economics, an auto glass company had a change in management which resulted in a shift from fixed hourly wages to piece-rate compensation. Not only did this company experience a significant increase in the productivity levels of the current workforce, but, over time, less effective employees exited and were replaced with more productive workers. Therefore, as this particular case demonstrates, the format or schedule of compensation can play a very substantial role in the dynamics and, potentially, the effectiveness of a firm or sector's labor force (Lazear, 2001).

Careers in the public sector tend to have tenure or civil service protections that are more prevalent or extensive than in the private sector. Perhaps unsurprisingly, economists have shown that public sector employees tend to be more risk averse (Dohmen et al., 2005; Hartog, Ferrer-i-Carbonell, & Jonker, 2002; Masclot, Colombier, Denant-Boemont, & Loheac, 2009). Risk aversion has been found to be a significant predictor of seeking and obtaining employment in the public sector (Bellante & Link, 1981). Public sector workers

are also significantly more likely to prefer safer, more guaranteed compensation than those who enter the private sector (Buurman, Delfgaauw, Dur, & Van den Bossche, 2012). With regard to teachers in particular, postsecondary students who choose to enter teaching are significantly more risk averse than their colleagues (Bowen et al., 2015) and are significantly more likely to dislike less equitable pay systems (Perez, 2011). Moreover, the Bowen et al. (2015) and Perez (2011) studies were with teachers in training,

possibly suggesting that this relationship can be at least somewhat attributed to selection, rather than being entirely credited as a socialization effect.

Motivations powerfully influence an individual's enjoyment and performance on the job. Self-determination theorists, for example, contend that individuals feel and demonstrate greater efficacy when they have a strong sense of autonomy over what they do (Deci & Ryan, 1995). This argument derives from extensive research on the distinction between extrinsic and intrinsic motivation which is described as responses to the incentives external to the task itself (extrinsic) versus an individual's interest and commitment to the task independent of external rewards (intrinsic) (Koehler & Rainey, 2008, p. 39). In cases where individuals are predominantly intrinsically motivated to perform a particular job or task, the growth or insertion of extrinsic rewards can also "crowd out" intrinsic motivations which could decrease outputs (Frey, 1997). Studies in the field have corroborated this theory, with studies showing that financial rewards occasionally fail to bring about their intended effects (e.g., Frey & Oberholzer-Gee, 1997; Gneezy & Rustichini, 2000; Handgraaf, de Jeude, & Appelt, 2013).

Intrinsic public service motivation stemming from individuals' altruistic desires or needs to perform meaningful social service is often a major draw for individuals who are attracted to work in the public sector (Brewer, Selden, Facer, & Rex, 2000; Rainey & Steinbauer, 1999). While intrinsic public service motivation is associated with stronger commitments to careers (Knoke & Wright-Isak, 1982), it also presents challenges in managing and compensating these employees (Koehler & Rainey, 2008). For example, intrinsically motivated workers often want more job autonomy, but careers in public service require employees to make certain sacrifices to attain goals that serve the community (Koehler & Rainey, 2008). Professional public sector employees also need to be financially compensated, and, like most other employees, want recognition for their good efforts. However, when rewards for employees come in the form of wages that are tied directly to performance (i.e., an extrinsic financial motivator), workers will potentially perceive these incentives as being controlling and demeaning (Benabou & Tirole, 2003; Frey, 1997).

The role of image or reputational motivation and its interactions with intrinsic and extrinsic motivations also raises issues regarding prosocial behavior in public service management (Benabou & Tirole, 2006). Image or reputational motivation is defined as "an individual's tendency to be motivated partly by others' perceptions. Image motivation therefore captures the role of opinion in utility, i.e., the desire to be liked and respected by others and by one's self" (Ariely, Bracha, & Meier, 2009, p. 544). The introduction of extrinsic incentives can diminish the benefits derived from individuals' public or self-perceptions that typically come with a sense of engaging in prosocial behaviors (Benabou & Tirole, 2006). For example, many K-12 teachers likely go above and beyond the minimal requirements for maintaining employment. Without the presence of financial incentives, it is likely that teachers receive some utility from self-image or the public view of their admirable, selfless dedication to education. However, when employers introduce extrinsic rewards, just the plausibility that a teacher is now putting forth extra efforts for personal financial gains can reduce the positive public- and self-image of the profession.

Psychologists and economists have conducted experiments in both field and laboratory settings in attempts to further examine and potentially address ways to ameliorate motivation conflict through different forms and deliveries of incentives. Recent studies have shown that financial incentives induce more

negative feelings and responses with intrinsically motivated workers; however, when incentives are given sufficient weight, these negative consequences are typically offset by the benefits derived from increased levels of productivity (Falk & Kosfeld, 2006). When a drastic increase in extrinsic incentives is not possible, positive feedback, specifically through acknowledgment and praise appears to have better outcomes than small financial incentives (Deci, Koestner, & Ryan, 1999; Handgraaf et al., 2013).

Our study contributes to the literature on teacher merit pay by utilizing sharp discontinuities in two school districts' compensation formats to examine if merit pay alters the composition of their workforces. We also examine the relationships between teacher quality, risk aversion, and perceptions of performance pay to better determine how these personnel aspects interact and potentially influence teachers' views and reactions to merit pay. In the next section we provide an in-depth overview of the data and analytical strategies that we have applied for this study.

EMPIRICAL FRAMEWORK

PROCEDURES

We applied a procedure developed by Holt and Laury (2002) for this study to measure teachers' risk preferences. This tool is commonly used in experimental economics (e.g., S. Anderson, Harrison, Lau, & Rutstrom, 2008; Dohmen, Falk, Huffman, & Sunde, 2010; Eckel & Wilson, 2004). Prior studies have shown that this assessment has high retest reliability (Harrison, Johnson, McInnes, & Rustrom, 2005) and external validation with individuals' engagements in other risky behaviors (e.g., L. R. Anderson & Mellor, 2008).

We slightly modified the Holt and Laury (2002) task from its original format for this study. This version of this task is different in that all payouts were increased by three times the amount of those from the original version. The instrument is presented in Table 1.³ As with the original version, we asked participants to choose between two options, A and B, for 10 different lotteries. For each lottery the two potential payouts for Option A were \$4.80 or \$6.00, and \$0.30 and \$11.55 for Option B. Option A is herein referred to as the "safe" choice because it has substantially less disparity in the payouts than with Option B. As a participant sequentially goes through each of the 10 lotteries the probability of receiving the higher payout for either Option A or B steadily increases by 10 percentage points. We instructed participants to choose A or B for each of the 10 lotteries and notified them in advance that they would only receive the payout from one of these lotteries that would be randomly determined later by a 10-sided die.

Table 1. Lottery Choices

Lottery	Option A	Possible Roll(s) for Payout	Option B
	p		Payout p

1.	\$6.00	1	\$11.55
	\$4.80	2,3,4,5,6,7,8,9,10	\$0.30
2.	\$6.00	1,2	\$11.55
	\$4.80	3,4,5,6,7,8,9,10	\$0.30
3.	\$6.00	1,2,3	\$11.55
	\$4.80	4,5,6,7,8,9,10	\$0.30
4.	\$6.00	1,2,3,4	\$11.55
	\$4.80	5,6,7,8,9,10	\$0.30
5.	\$6.00	1,2,3,4,5	\$11.55
	\$4.80	6,7,8,9,10	\$0.30
6.	\$6.00	1,2,3,4,5,6	\$11.55
	\$4.80	7,8,9,10	\$0.30
7.	\$6.00	1,2,3,4,5,6,7	\$11.55
	\$4.80	8,9,10	\$0.30
8.	\$6.00	1,2,3,4,5,6,7,8	\$11.55
	\$4.80	9,10	\$0.30
9.	\$6.00	1,2,3,4,5,6,7,8,9	\$11.55
	\$4.80	10	\$0.30
10.	\$6.00	1,2,3,4,5,6,7,8,9,10	\$11.55
	\$4.80	---	\$0.30

A participant's risk tolerance is determined by the lottery at which she is willing to switch from Option A (i.e., the safe option) to Option B. The expected payout for either option is calculated by multiplying the possible payouts with their respective probabilities and then summing within the payout options. A

participant who is risk neutral will always choose the option that offers the higher expected payout for a given lottery. With this particular task a risk-neutral participant chooses Option A for the first four lotteries, switches to Option B for the fifth lottery, and then goes with Option B for the remaining lotteries. A risk-loving participant sacrifices the difference between the expected payouts in exchange for the potential to receive the higher payout at some point before the fifth lottery. A risk-averse participant forgoes the difference in expected payouts to reduce the uncertainty that comes with choosing an option with more disparate payout options and sticks with the safe choice beyond the fourth lottery.

The teachers who participated were made aware of the option to take part in the study the day before or the morning of our arrival to campus. School administrators informed the teachers that the study examined how individuals make economic decisions; that participation was completely voluntary; that the study would take approximately 15 minutes; that they would have the opportunity to participate during their planning periods, lunch break, or at the end of the school day; and that they would be paid for their participation. Upon arrival, the researcher greeted the participant, gave a brief overview of the study, and then provided a hardcopy of the consent form. The researchers reminded the teachers that participation was voluntary and that they could leave at any point in the study. Only three teachers who showed initial interest in partaking ultimately decided to not complete the exercise.

After receiving instructions for the task, we asked participants a task comprehension question to ensure that they understood the procedure. If the teacher then had no further questions, they then marked their preferences for each of the 10 lotteries before proceeding to the other side of the paper to complete a brief survey. This survey included items to assist in collecting background information and other potentially salient characteristics. Information on each participant's age, gender, household income, undergraduate institution, job history, mother's level of education, years of teaching experience, year of hire at their current school, and opinions with regard to performance-based pay were collected. Participants knew that there was a survey involved before starting the task, but they could only view the demographic and opinion items of the survey after completing the behavior task so that these items could not frame or influence their lottery decisions.

When participants were finished, the researcher verified that the participant completed the task and survey.⁴ The researcher then rolled a ten-sided die to establish which of the 10 lotteries would determine the participant's payout. Next, the researcher rolled the die again to determine the participant's payout based on the teacher's choice of either A or B for the randomly selected

lottery. Participants then received their payments, signed a receipt of payment form, and were thanked for participating in the study.⁵ On average, participants received \$6.29 for about 15 minutes of their time.

SAMPLE

One hundred and twenty teachers agreed to participate in the study. We recruited all participants from two school districts which had implemented teacher performance pay programs. One was a charter school district that put performance pay in effect in the 2008–2009 school year and lasted for 4 years until funding for the program was exhausted (i.e., the program was not eliminated due to a lack of school-level support). The other was a traditional public school district that implemented a similar performance pay program in the 2009–2010 school year and was still in operation at the time of the study. Generally speaking, up to \$10,000 of a teacher's salary was based on

Generally speaking, up to \$10,000 of a teacher's salary was based on performance measures.⁶ These performance measures consisted of students' value-added scores at the individual classroom, grade, and school levels as well as principal-conducted year-end evaluations that were based on classroom observations, adherence to school policies, and attendance rates.

Forty-eight of these teachers (40%) started their employments with performance-based pay in place. Of these teachers, 13 (27%) indicated that the merit pay program was a "deciding factor" in their current employment decisions. Teachers' descriptive statistics are provided in Table 2. The teachers sampled were predominantly white females, a reflection of the teacher workforce of the region. In addition, teachers who opted in to the merit pay program were significantly younger, had less teaching experience, and had lower average household incomes, corroborating expectations based on prior literature (e.g., Milanowski, 2007). These findings likely are explained by the fact that most teachers who opted in to the program were hired in more recent recruiting cycles than their colleagues. Teachers who opted in were also more likely to have mothers who obtained at least an undergraduate degree. Teachers at the charter school district appear more likely to have opted in to the performance pay program. However, this relationship could be attributed to the fact that performance pay had been around for the majority of the hiring cycles at this charter school district. Nevertheless, we include an indicator for whether teachers were employed at the charter versus traditional public school district in our analytical models in order to control for the extent to which this difference in district might influence teachers' risk preferences as well as value-added and principal evaluation scores.

Table 2. Sample Descriptive Statistics

Variable	Overall Mean	Opted In	Performance Pay Deciding Factor
Number of Safe Choices	5.4	5.4	4.2
Age	40.2	35.2	32.9
Teaching Experience	11.0	8.0	6.9
Pct. Female	80.0	81.2	84.6
Mother's Education			
Pct. Less Than HS	5.8	2.1	---
Pct. HS Degree	29.2	16.7	7.7
Pct. Some College	23.3	22.9	---
Pct. Undergrad. Degree	26.7	41.7	61.5

Pct. Grad. Degree	15.0	16.7	30.8
Household Income			
Pct. Less Than \$50k	30.0	37.5	53.8
Pct. \$51k-\$80k	27.5	31.3	23.1
Pct. \$81k-\$110k	28.3	22.9	23.1
Pct. More Than \$110k	12.5	8.3	---
Teaching Subject Area			
Pct. General Education	14.5	19.1	23.1
Pct. Core Subject	68.3	76.6	84.6
Pct. STEM	32.5	31.9	30.8
Pct. Humanities	23.1	27.7	38.5
Pct. Special Education	8.5	8.5	7.7
Pct. Other	12.7	8.4	7.7
School			
Pct. Charter	41.7	72.9	92.3
Pct. Traditional Public	58.3	27.1	7.7
Teaching Level			
Pct. Elementary	42.5	50.0	61.5
Pct. Middle	23.3	25.0	30.8
Pct. High	34.2	25.0	7.7

N	120	48	13
---	-----	----	----

Notes: Columns do not represent exclusive groups of teachers. Teaching Subject Area does not sum to 100 percent because categories overlap and some teachers instructed multiple subject areas.

RESEARCH QUESTIONS & ANALYTICAL STRATEGIES

1. Does performance-based pay significantly alter the composition of the teacher workforce in terms of risk preference?

In order to determine whether performance-based pay attracts teachers who are significantly different in terms of their risk preferences the number of safe choices (Option A) that teachers made on the Holt and Laury (2002) task was regressed on their categorical responses to survey items regarding their support for performance pay. The outcome measure for these analyses is a count variable; therefore, a Poisson regression was applied to improve efficiency and reduce the likelihood of biased estimates (Long & Freese, 2006). Characteristics that are likely predictive of risk aversion, independent of the teacher's preferences for performance-based pay, were controlled for with the collected survey data. We also conducted a separate analysis that focuses on those teachers who both opted in and revealed a stated preference for performance pay in their employment decisions. It could be the case that teachers identified as having opted in to schools with these programs were not actively seeking employment at a school that offered performance pay. Therefore, this additional analysis could provide a better sense for whether self-identified, more active performance pay seekers reflect a more substantial shift in teachers' preferences.

2. Are performance pay-preferring teachers assessed as being more effective?

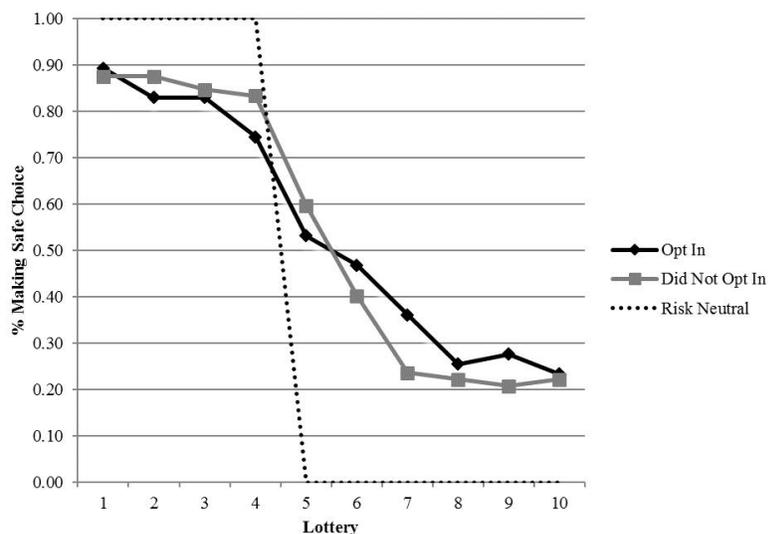
We applied ordinary least squares (OLS) regressions to estimate whether pay preference is predictive of two separate measures of teacher quality. Pay preference was determined both by a teacher's decision to work at a school with performance-based pay in place as well as their surveyed responses to questions with regard to compensation predilections. Our measures of teacher quality were those used to determine compensations under the performance-based pay plans: (1) average student value-added scores on the Northwest Evaluation Association's Measures of Academic Progress assessment and (2) year-end performance evaluations. As with assessing changes in the composition of the workforce, we controlled for other salient characteristics using variables generated from teachers' survey responses.

RESULTS

CHANGE IN THE TEACHER WORKFORCE?

Figure 1 presents the distributions of the proportion of individuals making safe choices for teachers opting into a performance-based pay system and those who did not. The dotted line serves as a reference, indicating how a perfectly risk-neutral group would respond to the risk-elicitation tool. While both groups have non-trivial number of teachers who always choose the more risky option, the fact that both distributions do not slope as steeply between Lotteries 4 and 5 indicates that both groups are relatively risk averse. In general, the two distributions presented in Figure 1 are noticeably similar. A Kolmogorov-Smirnov test for distributional differences ($p = 0.77$) as well as a Mann-Whitney test of rank-sum differences ($p = 0.98$) fail to reject the hypothesis of no significant difference. In addition, a simple t test for mean differences fails to reject the null hypothesis that the two groups differ in the average number of safe choices ($p = 0.81$). Based on these analyses it appears as though these districts' performance-based programs did not have a significant influence on their workforce compositions in terms of risk preferences.

Figure 1. Percentage Safe Choices in Each Lottery by Opt In Status

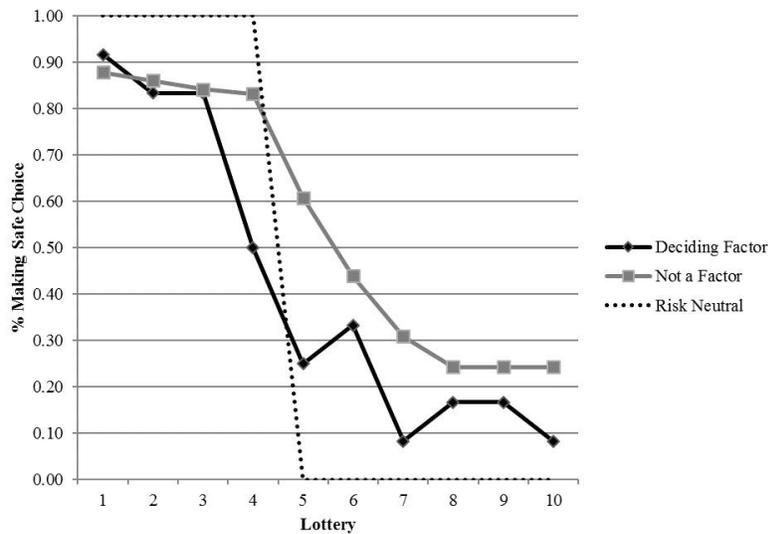


Note: The dotted line is a reference point for the distribution of choices for a hypothetical, perfectly risk neutral group.

When restricting the definition of opting in to just those teachers who claimed that performance-based pay was a deciding factor in their employment decisions, there is evidence that suggests that there is indeed a significant compositional shift. The participants who indicated that performance pay was a deciding factor in their decision to work at their school made 1.33 fewer safe

a deciding factor in their decision to work at their school made 1.55 fewer safe choices on average ($p = 0.06$). The comparison of the full distributions of teachers' choices is presented in Figure 2. In the fifth lottery, where a risk-neutral individual would no longer make the safe choice, only about one fourth of the more-active performance pay seekers make a safe choice while slightly more than 60% of nonactive seekers were still selecting the safer option. A Mann-Whitney rank-sum test of differences indicates a significant difference in risk preferences for these active performance pay seekers ($p = 0.03$). Despite evidence of significantly different means and visual evidence for these teachers having different risk preferences, a Kolmogorov-Smirnov test fails to reject the null hypothesis of no significant difference ($p = 0.48$). However, this failure to reject could stem from power concerns that come with formal tests of normality assumptions in cases of small sample sizes, making these results inconclusive (Razali & Wah, 2011).

Figure 2. Percentage Safe Choices in Each Lottery, Performance Pay as Deciding Factor



Note: The dotted line is a reference point for the distribution of choices for a hypothetical, perfectly risk neutral group.

The analyses presented so far have focused on differences in teacher workforce groups without controlling for additional demographic variables. Table 3 presents the parameter estimates for five separate Poisson regressions where the number of safe choices is the dependent variable. For more straightforward interpretations, estimates are discussed in terms of percent change.⁷ Column 1 indicates that teachers who opt in appear to have similar risk preferences to their colleagues; however, those teachers who opted in and indicated that merit pay was a deciding factor make 29% fewer safe option choices than their colleagues ($p = 0.03$). Controls for gender, age, experience, school district, mother's level of education, total household income, teacher's content subject area, and teacher grade-school level (elementary, middle, or high school) are added in Models 3-5. The finding that performance pay seekers are relatively more risk loving than their colleagues is robust across all models. The only other consistent significant predictor of teacher's risk preference is a continuous variable for the number of years of experience. An additional 10 years of teaching experience significantly increases the number of safe choices anywhere from 13% to 20% on average depending on the model specification ($p < 0.05$).⁸

depending on the model specification ($p < 0.05$).

Table 3. Poisson Regression Results for Number of Safe Choices

Variable	(1)	(2)	(3)	(4)	(5)
Opted In	0.10 (0.09)	0.09 (0.10)	0.06 (0.10)	0.10 (0.10)	0.07 (0.11)
Performance Pay	-0.34**	-0.36**	-0.38**	-0.41**	-0.43**
Deciding Factor	(0.16)	(0.16)	(0.17)	(0.16)	(0.17)
Female		0.04 (0.10)	0.07 (0.11)	0.07 (0.11)	0.08 (0.12)
Age		-0.00 (0.01)	-0.00 (0.01)	-0.00 (0.01)	-0.00 (0.01)
Experience		0.01** (0.01)	0.01** (0.01)	0.01* (0.01)	0.01* (0.01)
Charter		0.09 (0.10)	0.13 (0.11)	0.11 (0.11)	0.15* (0.12)
<u>Additional Controls</u>					
Mother's Education			X		X
Household Income			X		X
Teacher Subject Area				X	X
School Level				X	X
Constant	1.67*** (0.05)	1.59*** (0.23)	1.23*** (0.16)	1.51*** (0.28)	1.22*** (0.39)
N	119	119	117	116	114

Note: Numbers in parentheses below coefficient estimates are standard errors, clustered at the campus level. *, **, and *** indicate significance at the

10%, 5%, and 1% *p*-value levels, respectively, in a two-sided alternative to the null that the coefficient value is zero. Additional controls are insignificant predictors of the number of safe choices with the exceptions of mother's education being less than high school in third model and mother's education being a graduate degree in the fifth model, both of which are negatively associated with number of safe choices ($p < 0.10$).

PAY PREFERENCES AND TEACHER QUALITY

Linking teachers' survey responses on their dispositions towards performance-based pay to both their value-added scores and principal performance evaluations, provides a sense for whether there is a relationship between measures of teacher quality and pay preferences. Table 4 presents average teacher value-added score disaggregated by whether the participant disagreed, agreed, or was neutral for each of the four survey items about their current attitudes towards merit pay. A consistent result is that teachers who generally disagree with merit pay (both for their own compensations and as a general education policy that applies to all teachers) are seemingly more effective in terms of students' average value-added scores than teachers who prefer performance pay or are neutral towards this form of compensation. When provided the hypothetical scenario that a teacher will necessarily be compensated based on student performance, there appears to be no correlation between effectiveness and a preference for compensation being determined by team- or individually based results.

Table 4. Student Value-Added Scores and Pay Preferences

Survey Item	Disagree	Neutral	Agree
-Would seek employment at another performance pay school	+0.12 (27%)	-0.01 (45%)	-0.10 (28%)
-Teachers should be paid based on student performance	+0.12 (34%)	-0.05 (36%)	-0.07 (30%)
-I prefer pay based more on student performance relative to years of experience	+0.23 (42%)	-0.11 (45%)	-0.35 (13%)
-If based on performance, I prefer pay based more on individual versus team performance	+0.06 (34%)	-0.12 (28%)	+0.04 (37%)

Note: Statistics are in terms of standard deviations. Variables have been z-scored such that the overall sample mean is zero with a standard deviation of one. Proportions of respondents who disagree, agree, or are neutral for each survey item are provided in parentheses.

Table 5 displays the relationship between pay preferences and teachers' standardized performances on principal evaluations. Generally the relationship between pay preferences and performance evaluations appears to be the opposite of value-added scores and attitudes toward merit pay. Teachers who are more likely to favor performance pay, both in terms of personal preference and as a matter of public policy, seem to rate higher on their principals' evaluations. There also appears to be less ambiguity in the relationship between support for how to determine teacher performance and principal evaluation scores. Higher-rated teachers are more likely to favor pay based on individual, rather than team, performance.

Table 5. Performance Evaluations and Pay Preferences

Survey Item	Disagree	Neutral	Agree
-Would seek employment at another performance pay school	-0.14 (28%)	-0.03 (45%)	+0.20 (28%)
-Teachers should be paid based on student performance	-0.19 (35%)	-0.10 (38%)	+0.37 (28%)
-I prefer pay based more on student performance relative to years of experience	-0.09 (42%)	+0.00 (45%)	+0.37 (13%)
-If based on performance, I prefer pay based more on individual versus team performance	-0.11 (40%)	+0.05 (34%)	+0.03 (26%)

Note: Statistics are in terms of standard deviations. Variables have been z-scored such that the overall sample mean is zero with a standard deviation of one. Proportions of respondents who disagree, agree, or are neutral on each survey item are provided in parentheses. Proportions do not perfectly align with Table 3 due to 10 teachers having performance evaluations but no value-added scores in our data.

Table 6 presents the results of OLS regressions where value-added score (Models 1–3) or principal evaluation (4–6) serves as the dependent variable. The first three pay preference survey items from Tables 4 and 5 are standardized across respondents, averaged, and then standardized again to create an index score with a mean of zero and standard deviation of one for teacher's disposition towards performance-based pay. A positive score on this index indicates that the teacher had a more favorable than average view of performance-based pay relative to the other participants of this study.⁹ The same control variables from prior analyses were once again included when examining the relationship between pay preferences and both measures of teacher quality. There is a consistent, negative relationship between a teacher's support for performance-based pay and students' value-added

scores, and there is a stable, positive association between merit pay support and principals' evaluations; however, the majority of these findings are not statistically significant. Although the estimate for the performance pay index on value-added scores appears to be quite large, a one standard deviation increase in support for performance pay on this index is related to a one-fourth of a standard deviation decrease in a teacher's average student value-added score, this estimate only achieves a traditional level of statistical significance in the second model ($p = 0.06$).¹⁰ The only other significant

positive predictor of either teacher quality measure is whether the employee worked at the charter, rather than the traditional public school, district. These effects could be due to real differences in teacher quality across districts. However, these differences could also be attributable to omitted variables such as those pertaining to student populations that were conducive to greater standardized test score gains.

Table 6. Regression Results for Pay Preferences and Teacher Quality Measures

Variable	Value-Added Scores			Performance Evaluations		
	(1)	(2)	(3)	(4)	(5)	(6)
Pay Preference Index	-0.18	-0.25*	-0.25	0.11	0.06	0.16
	(0.12)	(0.13)	(0.16)	(0.13)	(0.13)	(0.19)
Female		0.16	0.17		0.38	0.68
		(0.28)	(0.36)		(0.31)	(0.43)
Age		0.03	0.02		-0.01	0.00
		(0.02)	(0.02)		(0.02)	(0.02)
Experience		-0.03	-0.03		0.02	0.02
		(0.02)	(0.02)		(0.02)	(0.02)
Charter		0.58*	0.30		0.71**	0.48
		(0.29)	(0.39)		(0.31)	(0.45)
<u>Additional Controls</u>						
Mother's Education			X			X
Household Income			X			X

Teacher Subject Area			X			X
School Level			X			X
Constant	0.00	-1.13	-1.27	-0.10	-0.46	-1.36
	(0.12)	(0.70)	(0.98)	(0.13)	(0.68)	(1.06)
N	67	67	67	54	54	54

Note: Numbers in parentheses below coefficient estimates are standard errors, clustered at the campus level. *, **, and *** indicate significance at the 10%, 5%, and 1% p-value levels, respectively, in a two-sided alternative to the null that the coefficient value is zero. Additional controls are all insignificant predictors of the number of safe choices with all of these models.

Teachers preferring performance-based pay appear to have received slightly better ratings on their principal-administered annual performance evaluations than those more opposed; however, this correlation is statistically insignificant. It is worth noting that this quality measure was more subjective than the value-added measure. Specifically, principal performance evaluations included criteria such as whether or not teachers adhered to school policies and were cooperative in their work environments. Therefore, the possibility of an endogenous relationship between teachers' perspectives of performance-based pay (reflecting general attitude towards their work environments) and how principals rate them on these evaluations should be taken into consideration.

DISCUSSION

These results provide suggestive evidence that performance-based pay could indeed change the composition of the teacher workforce. Teachers who were employed at these two school districts with merit pay in place appear to be comparable to their colleagues whom were employed with the more traditional step-and-lane pay scale in terms of risk preference. However, teachers who claimed that they were more actively seeking performance-based pay to begin with were substantially more risk loving, suggesting that performance-based pay could attract fundamentally different employees to the teacher workforce.

While these results imply that performance pay has the capability to alter the teacher workforce, additional findings also indicate that this reform might have negative unintended consequences. Teachers who were identified as being more effective are no more likely to support the implementation of performance-based pay policies. This result is surprising as these teachers presumably have the most to gain financially from a performance-based pay system. Additionally, the teachers in this sample knew the information about their performance on these measures. Therefore, not only did these teachers have the most to gain, they had already experienced the financial benefits.¹¹

It is plausible that more effective teachers at these schools have a strong

sense of mission and a commitment to education that is independent of the provision of external rewards. Considering Frey's (1997) theory with regard to inserting extrinsic rewards with intrinsically motivated individuals, the implementation of performance-based pay may have crowded-out or challenged their intrinsic motivations. This crowding-out effect would likely cause teachers to view extrinsic incentives as regulatory controls over their job by subverting the autonomy or the sacrificial nature (real or perceived) of their professions rather than being viewed as rewards for their performances. This potential incompatibility between motivations and rewards could pose issues should teacher merit pay become more common.

There are, however, limitations to this study that inhibit the ability to make more definitive causal claims about the relationship between performance-based pay reform and its potential impacts on the teacher labor market. The limited size of this sample poses concerns regarding the external validity of these findings. Unfortunately, applicants who were not hired or teachers who decided to leave these districts in the years since the implementations of the pay reform were unavailable for surveying. Not having data from all applicants restricts the degree to which we can assess the full extent to which merit pay influenced employee attraction and retention in response to the new compensation format. The screening that took place during hiring cycles may bias these results, though the direction of this bias is uncertain.

CONCLUSIONS

The significant difference in risk preferences between teachers who claim to have prioritized performance-based pay in their employment decision and those who did not suggests that this compensation format reform has the potential to alter the composition of the education workforce. This result corroborates theory and evidence from the aforementioned labor economics literature. However, the incongruity between educator effectiveness and pay preferences could reflect a conflict between the goals and motivations of these teachers in relation to the influences of extrinsic rewards. This result supports theory and evidence from public service management, social psychology, and behavioral economics literature. These findings are not inherently at odds with one another, but they do pose critical challenges that education decision makers must contemplate when designing teacher pay reform policies.

The crucial consideration for policymakers is applying policy levers that optimize the recruitment and retention of highly effective teachers. The appeal of performance-based pay is that personnel recruitment and retention should improve as a result of increasing the lifetime earnings of effective teachers by tying wages to performance measures. However, these benefits come at a cost that has likely been overlooked by education reformers. In addition to the financial and political capital that policymakers need to expend in order to implement such policies, the findings from this study imply that the incorporation of explicit extrinsic rewards may not increase retention and possibly even dissuade current or potential, high-quality educators from the profession because they find intrinsic motivation crowding-out effects to outweigh the financial gains.

At the same time, a finding suggesting that performance-based pay may fail to deliver on its promises to improve efforts to recruit and retain a higher quality

teacher workforce does not necessarily justify maintaining the status quo of step-and-lane pay systems. Labor market realities will likely continue to make it difficult for school districts to employ effective educators, especially in harder-to-staff schools and content areas. Perhaps there is more promise in improving workforce composition with strategies that address these challenges without the use of explicit performance-based incentives (e.g., “combat” pay). Strategies that aim to encourage desired efforts and provide feedback without the negative consequences of extrinsic rewards tied to performance should also be further examined. One such strategy with potential is public acknowledgement and praise (Deci et al., 1999). For example, Handgraaf et al. (2013) tested the effectiveness of this approach with a field experiment where the use of monetary or praise rewards were given either publicly or privately in order to influence employees’ at-work energy conservation. They found that praise rewards were more effective than financial ones and that public recognition was more effective than private feedback. Moreover, workers in the social reward group maintained their efforts long after the completion of the experiment. Although the context of that study is different, it raises the possibility that public recognition and praise could be more cost efficient than merit pay.

While providing more specific, practical solutions to the numerous pressing questions with regard to performance pay and its impacts on the composition of the workforce is beyond the scope of this study, these findings raise vital policy considerations. In particular, this study highlights that, even if performance pay systems can attract different types of individuals to teaching, additional research is still needed to determine whether this type of compositional change is desirable. From a policy perspective, the potential for such unintended consequences suggests a need for incremental change occurring through small-scale, experimental programs rather than large shifts in compensation policy.

Given the limitations of this study, as well as the general lack of research on the compositional effects of performance pay programs, additional analyses are needed to further explore the impacts of performance-based pay reforms. Future research should further examine the extent to which public service motivation affects individuals whom decide to enter the teaching profession. Determining whether there is indeed a robust relationship between teachers’ motivations and effectiveness as well as the likelihood of remaining in the profession would have a major impact on schools’ employment decisions. When examining tradeoffs, it is also worth noting that motivation considerations should not necessarily be taken into consideration without also accounting for student outcomes. Specifically, Fehr and Falk (2002) contend that even if motivational crowding out does occur, the ultimate assessment of the effectiveness of extrinsic rewards should be the net benefit with regard to the outcome of interest. Future studies should assess whether crowd out of intrinsic motivation by extrinsic rewards does indeed come at the expense of student achievement, either as a result of changes in teachers’ outputs or as a consequence of sorting that ultimately reduces the quality of workers entering the profession.

Despite these limitations, this study provides important contributions to the discussion on teacher compensation policies and how they have the potential to affect the composition of the workforce. While it seems plausible that merit pay reforms can alter who is ultimately attracted to the profession, policymakers should further consider the effects with regard to current educators’ motivations and the possibility of unintended consequences with regard to retaining high quality teachers. Determining whether such a shift in the workforce is ultimately beneficial in terms of overall student achievement

the workforce is ultimately beneficial in terms of overall student achievement still requires a longer term evaluation of greater scale and scope. However, since such an evaluation is currently not possible, these results can hopefully help guide policy discussions and provide hypotheses to test in future field experiments that more comprehensively examine the tradeoffs of implementing teacher performance-based pay.

Notes

1. We restrict our definition of teacher quality to being “highly effective” in terms of value-added scores derived from statewide standardized assessments in addition to principal’s evaluations. Although there are important limitations and concerns with the applications of such teacher quality measures (see Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012), our focus is strictly on whether pay reforms would produce the effects that policymakers presumably intend. Therefore, while these quality measures may not serve as undisputed assessments of a teacher’s effectiveness, they do permit the possibility to analyze differences in outcomes that education policymakers seem to have valued.
2. Other states at this time primarily depended on principal evaluations (e.g., Florida, North Carolina, and Texas) (Cornett & Gaines, 2002).
3. This decision was based on guidance we received from colleagues about the expected payout amounts necessary for recruiting participants from our targeted population.
4. One participant had to be excluded from later analyses because she skipped one part of the risk task without the researcher noticing it until after the participant had exited.
5. This payment method replicated the procedure from Holt and Laury (2002).
6. This cap amount fluctuated depending on the teacher’s role and course load. Specifically, teachers who taught core, tested subjects had higher caps than elective course teachers.
7. Percent change is computed as (Long & Freese, 2006).
$$100 \times \frac{E(y|x, x_k + \delta) - E(y|x, x_k)}{E(y|x, x_k)}$$
8. Due to the concern that simultaneously including age and years of experience might pose issues with regard to multicollinearity, models were also run with both variables simultaneously as well as with each one separately. The qualitative interpretations for these variables as well as the independent variable of interest were the same across specifications with the exception of Experience in column (4) having a p value of 0.05, which was significant at $p < 0.10$ rather than $p < 0.05$.
9. Participants’ responses to preference for basing performance on individual versus team performances are not included in the calculation of the index score because this item asks for preferences within the context of a hypothetical performance-based pay system rather than eliciting attitudes about this form of compensation in general.
10. It is worth noting that this relationship likely fails to meet statistical significance in Model 3 due to the added structure provided through the use of additional controls and evinced by the increases in the standard errors for the point estimates.
11. In other words, it is not the case that those teachers deemed more effective on value-added measures were more likely to underestimate their effectiveness and disagree with performance pay based on such a misperception. These teachers already had multiple years of feedback and paychecks that affirmed their effectiveness. but they still were unsupportive of

payments that induce their effectiveness, security, and more appropriate or merit pay.

References

Anderson, S., Harrison, G. W., Lau, M. I., & Rutstrom, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, *76*(3), 583–618.

Anderson, L. R., & Mellor, J. M. (2008). Predicting health behaviors with an experimental measure of risk preference. *Journal of Health Economics*, *27*(5), 1260–1274.

Ariely, D., Bracha, A., Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *The American Economic Review*, *99*(1), 544–555.

Auguste, B. G., Kihn, P., & Miller, M. (2010). *Closing the talent gap: Attracting and retaining top-third graduates to careers in teaching: An international and market research-based perspective*. New York: McKinsey & Company.

Balch, R., & Springer, M. G. (2015). Performance pay, test scores, and student learning objectives. *Economics of Education Review*, *44*, 114–125.

Ballou, D. (2001). Pay for performance in public and private schools. *Economics of Education Review*, *20*(1), 51–61.

Ballou, D., & Podgursky, M. (1995). Recruiting smarter teachers. *The Journal of Human Resources*, *30*(2), 326–338.

Ballou, D., & Springer, M. G. (2015). Using student test scores to measure teacher performance: Some problems in the design and implementation of evaluation systems. *Educational Researcher*, *44*(2), 77–86.

Bellante, D., & Link, A. N. (1981). Are public sector workers more risk averse than private sector workers? *Industrial and Labor Relations Review*, *34*(3), 408–412.

Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies*, *70*(3), 489–520.

Benabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, *96*(5), 1652–1678.

Bowen, D. H., Buck, S., Deck, C., Mills, J. N., & Shuls, J. V. (2015). Risky business: An analysis of teacher risk preferences. *Education Economics*, 23(4), 470–480.

Brewer, G. A., Selden, S. C., Facer, I. I., & Rex, L. (2000). Individual conceptions of public service motivation. *Public Administration Review*, 60(3), 254–264.

Brookhart, S. M., & Freeman, D. J. (1992). Characteristics of entering teacher candidates. *Review of Educational Research*, 62(1), 37–60.

Buurman, M., Delfgaauw, J., Dur, R., & Van den Bossche, S. (2012). Public sector employees: Risk averse and altruistic? *Journal of Economic Behavior & Organization*, 83(3), 279–291.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9), 2633–2679.

Cornett, L. M., & Gaines, G. F. (2002). *Quality teaching: Can incentive policies make a difference?* Atlanta, GA: Southern Regional Education Board. Retrieved from <http://files.eric.ed.gov/fulltext/ED464085.pdf>

Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *Phi Delta Kappan*, 93(6), 8–15.

Deci, E. L., & Ryan, R. M. (1995). Human autonomy. In *Efficacy, agency, and self-esteem* (pp. 31–49). New York, NY: Springer US.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, 125(6), 627–668.

Dee, T., & Keys, B. J. (2004). Does merit pay reward good teachers? Evidence from a randomized experiment. *Journal of Policy Analysis and Management*, 23, 471–488.

Dohmen, T. & Falk, A. (2010). You get what you pay for: Incentives and selection in the education system. *The Economic Journal*, 120(546), F256–F271.

Dohmen, T., & Falk, A. (2011). Performance pay and multidimensional sorting: Productivity, preferences, and gender. *The American Economic Review*, 101, 556–590.

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, *100*(3), 1238–1260.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., & Wagner, G. G. (2005). *Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey*. Retrieved from <https://ideas.repec.org/p/diw/diwwpp/dp511.html>

Eckel, C. C. & Wilson, R. K. (2004). Is trust a risky decision? *Journal of Economic Behavior and Organization*, *55*, 447–465.

Falk, A., & Kosfeld, M. (2006). The hidden costs of control. *The American Economic Review*, *96*(5), 1611–1630.

Fehr, E., & Falk, A. (2002). Psychological foundations of incentives. *European Economic Review*, *46*(4), 687–724.

Figlio, D., & Kenny, L. (2007). Individual teacher incentives and student performance. *Journal of Public Economics*, *91*(5-6), 901–914.

Finn, Jr. C. E. (2001). Getting better teachers—and treating them right. In T. M. Moe (Ed.), *A primer on America's schools* (pp. 127–150). Stanford, CA: Hoover Institute.

Fox, J. N. (1984). Restructuring the teacher work force to attract the best and the brightest. *Journal of Education Finance*, *10*(2), 214–237.

Francois, P. (2000). Public service motivation as an argument for government provision. *Journal of Public Economics*, *78*(3), 275–299.

Frey, B. S. (1997). A constitution for knaves crowds out civic virtues. *The Economic Journal*, *107*(443), 1043–1053.

Frey, B. S., & Oberholzer-Gee, F. (1997). The cost of price incentives: An empirical analysis of motivation crowding-out. *The American Economic Review*, *87*(4), 746–755.

Fryer, R. G. (2013). Teacher incentives and student achievement: Evidence from New York City Public Schools. *Journal of Labor Economics*, *31*(2), 373–407.

Fryer Jr, R. G., Levitt, S. D., List, J., & Sadoff, S. (2012). *Enhancing the efficacy of*

teacher incentives through loss aversion: a field experiment (No. w18237).
National Bureau of Economic Research.

Gneezy, U., & Rustichini, A. (2000). Pay enough or don't pay at all. *The Quarterly Journal of Economics*, 115(3), 791–810.

Goldhaber, D., DeArmond, M., Player, D., & Choi, H. J. (2008). Why do so few public school districts use merit pay? *Journal of Education Finance*, 33(3), 262–289.

Goldhaber, D., & Player, D. (2005). What different benchmarks suggest about how financially attractive it is to teach in public schools. *Journal of Education Finance*, 30(3), 211–230.

Goldhaber, D., & Walch, J. (2012). Strategic pay reform: A student outcomes-based evaluation of Denver's ProComp teacher pay initiative. *Economics of Education Review*, 31(6), 1067–1083.

Goodman, S. F., & Turner, L. J. (2013). The design of teacher incentive pay and educational outcomes: Evidence from the New York City Bonus Program. *Journal of Labor Economics*, 31(2), 409–420.

Handgraaf, M. J., de Jeude, M. A., & Appelt, K. C. (2013). Public praise vs. private pay: Effects of rewards on energy conservation in the workplace. *Ecological Economics*, 86, 86–92.

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review*, 30(3), 466–479.

Harrison, G. W., Johnson, E., McInnes, M. M., and Rutstrom, E. E. (2005). Temporal stability of estimates of risk aversion. *Applied Financial Economics Letters*, 1(1), 31–35.

Hart, C. (2015). Evaluating and rewarding teachers. *Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource*, 1–15. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/9781118900772.etrds0121/abstract>

Hartog, J., Ferrer-i-Carbonell, A., & Jonker, N. (2002). Linking measured risk aversion to individual characteristics. *Kyklos*, 55, 3–26.

Henderson, M. B., Peterson, P. E., & West, M. R. (2015). No common opinion on the Common Core: Also teacher grades, school choices, and other findings from the 2014 EdNext poll. *Education Next*, 15(1), 8–12.

Holt, C. A. & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644–1655.

Jacob, B. A., & Levitt, S. D. (2003). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *The Quarterly Journal of Economics*, 118(3), 843–877.

Jones, M., & Hartney, M. (2011). *Show who the money? Does performance pay attract higher quality teachers?* Working Paper.

Koedel, C., Mihaly, K., & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review*. Advance online publication. doi:10.1016/j.econedurev.2015.01.006

Koehler, M., & Rainey, H. G. (2008). Interdisciplinary foundations of public service motivation. In J. L. Perry & A. Hondeghem (Eds.), *Motivation in public management: The call for public service* (pp. 33–55). New York, NY: Oxford University Press.

Knoke, D., & Wright-Isak, C. (1982). Individual motives and organizational incentive systems. *Research in the Sociology of Organizations*, 1(2), 209–254.

Lavigne, A. L. (2014). Exploring the intended and unintended consequences of high-stakes teacher evaluation on schools, teachers, and students. *Teachers College Record*, 116(1).

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

Lavy, V. (2009). Performance pay and teachers' effort, productivity, and grading ethics. *American Economic Review*, 99(5), 1979–2011.

Lazear, E. P. (2000). Performance pay and productivity. *American Economic Review*, 90(5), 1346–1361.

Lazear, E. P. (2004). Output-based pay: Incentives, retention or sorting? *Research in Labor Economics*, 23, 1–25.

Lipsky, D. B., & Bacharach, S. B. (1983). The single salary schedule vs. merit pay: An examination of the debate. *Collective Bargaining Quarterly*, 11(4), 1–11.

Long, I. S., & Freese, I. (2006). *Regression models for categorical dependent*

Wooldridge, J. M. (2009). *Regression models for categorical dependent variables using Stata*. College Station, TX: Stata Press.

Marsh, J. A., Springer, M. G., McCaffrey, D. F., Yuan, K., Epstein, S., . . . & Peng, A. X. (2011). *A Big Apple for educators: New York City's experiment with schoolwide performance bonuses: Final evaluation report*. Santa Monica, CA: RAND Corporation.

Masclot, D., Colombier, N., Denant-Boemont, L., & Loheac, Y. (2009). Group and individual risk preferences: A lottery-choice experiment with self-employed and salaried workers. *Journal of Economic Behavior & Organization*, 70, 470–484.

Milanowski, A. (2007). Performance pay system preferences of students preparing to be teachers. *Education Finance and Policy*, 2(2), 111–132.

Muralidharan, K., & Sundararaman, V. (2011). Teacher performance pay: Experimental evidence from India. *The Journal of Political Economy*, 119(1), 39–77.

Murnane, R. J., & Cohen, D. K. (1986). Merit pay and the evaluation problem: Why most merit pay plans fail and a few survive. *Harvard Educational Review*, 56(1), 1–18.

Nadler, C., & Wiswall, M. (2011). Risk aversion and support for merit pay: Theory and evidence from Minnesota's Q Comp Program. *Education Finance and Policy*, 6(1), 75–104.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: U.S. Department of Education.

Organ, D. W., Podsakoff, P., & MacKenzie, S. (2006). *Organizational citizenship behavior: Its nature, antecedents, and consequences*. Thousand Oaks, CA: Sage.

Oyer, P., & Schaefer, S. (2005). Why do some firms give stock options to all employees?: An empirical examination of alternative theories. *Journal of Financial Economics*, 76(1), 99–133.

Perez, M. (2011). *Inequity aversion differences: Experimental evidence among prospective teachers and lawyers*. Available at SSRN 2018697

Piliavin, J. A., Grube, J. A., & Callero, P. L. (2002). Role as resource for action in public service. *Journal of Social Issues*, 58(3), 469–485.

Podgursky, M., & Springer, M. G. (2007). Teacher performance pay: A review. *Journal of Policy Analysis and Management*, 26(4), 909–949.

Popkewitz, T. S., & Lind, K. (1989). Teacher incentives as reforms: Teachers' work and the changing control mechanism in education. *Teachers College Record*, 90(4), 575–594.

Prostik, J. (1995). *History of teacher pay and incentive reforms*. Washington, D.C.: Educational Resources Information Center.

Rainey, H. G., & Steinbauer, P. (1999). Galloping elephants: Developing elements of a theory of effective government organizations. *Journal of Public Administration Research and Theory*, 9(1), 1–32.

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, 2(1), 21–33.

Reagan, R. (1983, May 21). *Address at commencement exercises at Seton Hall University in South Orange, New Jersey*. Retrieved from <http://www.reagan.utexas.edu/archives/speeches/1983/52183a.htm>

Reese, W. J. (2011). *America's public schools: From the common school to "No Child Left Behind"*. Baltimore, MD: Johns Hopkins University Press.

Resmovits, J. (2011, July 29). Arne Duncan boosts merit pay at teaching conference. *The Huffington Post*. Retrieved from http://www.huffingtonpost.com/2011/07/29/duncan-boosts-merit-pay-a_n_913608.html

Ritter, G. W., & Jensen, N. C. (2010). The delicate task of developing an attractive merit pay plan for teachers. *Phi Delta Kappan*, 91(8), 32–37.

Salop, J., & Salop, S. (1976). Self-selection and turnover in the labor market. *Quarterly Journal of Economics*, 90, 619–627.

Sawchuck, S. (2010, September 9). Merit pay found to have little effect on achievement. *Education Week*. Retrieved from http://www.edweek.org/ew/articles/2010/09/21/05pay_ep.h30.html

Sharpes, D. K. (1987). Incentive pay and the promotion of teaching proficiencies. *Clearing House*, 406–408.

Smarick, A. (2011). Diplomatic mission: President Obama's path to performance pay. *Education Next*, 11(1), 56–63.

Spring, J. (2004). *The American school 1642-2004* (6th ed.). New York, NY: McGraw-Hill Companies, Inc.

Springer, M. G., Ballou, D., Hamilton, L., Le, V., Lockwood, J. R., McCaffrey, . . . & Stetcher, B. (2012). *Final report: Experimental evidence from the Project on Incentives in Teaching*. Nashville, TN: National Center on Performance Incentives at Vanderbilt University.

Springer, M. G., Ballou, D., & Peng, A. X. (2014). Estimated effect of the Teacher Advancement Program on student test score gains. *Education Finance and Policy*, 9(2), 193–230.

Stronge, J. H., Gareis, C. R., & Little, C. A. (2006). Attracting, developing, retaining—and paying—quality teachers. *Teacher pay and teacher quality: Attracting, developing, and retaining the best teachers* (pp. 1–21). Thousand Oaks, CA: Corwin Press.

Wallis, C. (2008, February 13). How to make great teachers. *TIME*. Retrieved from <http://www.time.com/time/magazine/article/0,9171,1713473,00.html>

Wilms, W. W., & Chapleau, R. R. (1999, November 3). The illusion of paying teachers for student performance. *Education Week*. Retrieved from <http://www.edweek.org/ew/articles/1999/11/03/10wilms.h19.html>

Woessmann, L. (2011). Cross-country evidence on teacher performance pay. *Economics of Education Review*, 30(3), 404–418.

Cite This Article as: *Teachers College Record* Volume 119 Number 4, 2017, p. 1-32

<https://www.tcrecord.org> ID Number: 21714, Date Accessed: 9/6/2019 1:59:49 PM

[Purchase Reprint Rights for this article or review](#)