

Title: Evaluation of the Teacher Incentive Fund: Implementation and Impacts of Pay-for-Performance After Two Years

Authors and Affiliations: Hanley Chiang, Alison Wellington, Kristin Hallgren, Cecilia Speroni, Mariosa Herrmann, Steven Glazerman, Jill Constantine (Mathematica Policy Research)

Abstract Body

Background: Recent efforts to attract and retain effective educators and to improve teaching practices have focused on reforming evaluation and compensation systems for teachers and principals. In 2006, Congress established the Teacher Incentive Fund (TIF), which provides grants to support performance-based compensation systems for teachers and principals in high-need schools. This study focuses on performance-based compensation systems that were established under TIF grants awarded in 2010. The 2010 TIF grants were designed to create comprehensive performance-based compensation systems that could provide (1) incentives for educators to become more effective in improving student achievement in high-need schools, and (2) support for educators to improve their performance.

This is the second of four planned reports from a multiyear study for the U.S. Department of Education. It examines 2010 TIF grantees' programs and implementation experiences and the impacts of pay-for-performance bonuses on educator effectiveness and student achievement after the first (2011–2012) and the second (2012–2013) year of TIF implementation.¹

Purpose of the Study: This study describes the TIF implementation in all 2010 TIF districts and analyzes, in greater detail, the implementation and impacts of pay-for-performance for the subset of districts participating in the random assignment evaluation (evaluation districts).

The study focuses on the following 4 research questions:

1. What are the characteristics of all TIF districts and their performance-based compensation systems? What implementation experiences and challenges did TIF districts encounter?
2. How do teachers and principals in schools that did or did not offer pay-for-performance bonuses compare on key dimensions, including their understanding of TIF program features, exposure to TIF activities, allocation of time, and attitudes toward teaching and the TIF program?
3. How do pay-for-performance bonuses affect educator effectiveness and the retention and recruitment of high-performing educators?
4. What is the impact of pay-for-performance bonuses on students' achievement on state assessments in math and reading?

Setting: The study provides an overview of TIF implementation in 2012-2013 by all 2010 TIF grantees and an in-depth analyses of TIF implementation and effects of pay-for-performance on educator and student outcomes based on the subset of 10 districts in 7 states that are participating in a random assignment evaluation and completed two years of TIF implementation, the 2011–2012 and 2012–2013 school years.

Participants: The analysis of all TIF districts is based on a total 155 TIF districts who responded to the 2013 district survey (response rate of 95 percent). The in-depth analysis of the 10 evaluation districts is based on surveys to districts and to educators in 132 schools, conducted during their first and second years of TIF implementation. All 10 evaluation districts responded to the district survey in both years. A total of 795 teachers and 129 principals are included in the analysis of survey data in Year 1 and 904 teachers and 125 principals in Year 2. (The response rate for the principal survey was 98 percent in Year 1 (2012 survey) and 95 percent in Year 2

¹ Full study is available at: <http://ies.ed.gov/ncee/pubs/20154020/index.asp>

(2013 survey), and for the teacher survey was 92 percent in both years.) Analyses of impacts of performance-pay-bonuses on educator and student outcomes are based on the 132 schools within the 10 evaluation districts and use district administrative records for over 4,300 teachers and 40,000 students in each year.

Program: The 2010 TIF grant application notice included two competitions: a main competition and an evaluation competition. All applicants were required to include four program components: i) measures of educator effectiveness that included student achievement growth and observations of practice; ii) pay-for-performance bonuses designed to incentivize and reward educators *solely* for being effective; iii) extra pay for educators to take on additional roles or responsibilities such as becoming a master or mentor teacher; and iv) professional development to inform teachers about the performance measures and to provide support for improvement based on individual performance measures.

The 2010 TIF applicants that applied to the evaluation competition (the evaluation districts) were eligible for additional funding and more intensive implementation support in exchange for participating in a random assignment study of the pay-for-performance component of TIF. Evaluation grantees also received more specific guidance about how to structure pay-for-performance bonuses. They received examples of pay-for-performance bonuses that were *substantial* (with an average bonus worth 5 percent of the average educator's salary), *differentiated* (with at least some educators expecting to receive a bonus worth three times the average payout), and *challenging* to earn (with only those performing significantly better than average receiving bonuses). Although applicants had discretion over the proposed structure of the pay-for-performance bonus, these examples provided additional guidance to evaluation grant applicants and might have influenced how they designed their performance-based compensation systems.

Research Design: This study used an experimental study design to assess the impacts of pay-for-performance on educator and student outcomes. Before random assignment, evaluation districts chose which schools to include in the evaluation. Because a primary objective of the study was to measure the impact of pay-for-performance on student achievement on state assessments in high-need schools, every participating school had to have (1) at least half of its students receiving free or reduced-price lunch and (2) at least one grade level tested by state assessments (3rd to 8th grade). Mathematica Policy Research then conducted the random assignment of schools.

Elementary and middle schools within the evaluation districts were assigned randomly to treatment and control groups using a matched-pair randomization approach designed to maximize the balance between the two groups on observable characteristics.. As shown in figure 1, treatment and control schools were expected to implement the same required components of the district's TIF performance-based compensation system described before, except for the pay-for-performance bonus component. As a result, the study measured the impact of pay-for-performance bonuses implemented within the context of broader performance-based compensation systems. The study was not designed to measure the impact of implementing a TIF grant or the multiple components of a performance-based compensation system.

(Insert figure 1)

Treatment and control groups were expected to differ only in the opportunity for educators to receive pay-for-performance bonuses. Teachers and principals in treatment schools were eligible to earn a pay-for-performance bonus; teachers and principals in control schools received

an automatic bonus worth approximately 1 percent of their annual salary. The 1 percent bonus ensured that all educators in evaluation schools received some benefit from participating in the study: either the opportunity to earn a pay-for-performance bonus or the automatic bonus. Therefore, the impact of pay-for-performance estimated in this study potentially reflected two key differences between treatment and control schools: (1) bonuses in treatment schools were differentiated based on performance; and (2) bonuses in treatment schools were larger, on average, than in control schools.

Data Collection: Data for this report came from multiple sources. These sources included district surveys (administered to all 2010 TIF districts), surveys of principals and a sample of teachers in the 132 study schools within the 10 evaluation districts, and administrative data from the evaluation districts on educators' job assignments, performance ratings, and performance bonus awards, as well as student achievement scores on state assessments in math and reading (grades 3 through 8). These data sources enabled us to examine implementation broadly in all TIF districts and, within evaluation districts, to report on more detailed aspects of implementation and the impacts of pay-for-performance on educator and student outcomes.

Data analysis: Within the evaluation districts, we assessed the impacts of pay-for-performance on several educator and student outcomes, including educators' attitudes and behaviors (measured by survey responses), educator effectiveness (measured by performance ratings that educators received from their districts), and student achievement (measured by scores on state assessments in math and reading). Because schools were randomly assigned to offer pay-for-performance bonuses, differences in outcomes between the groups can be attributed to the impact of pay-for-performance. We estimated these impacts using a linear regression that accounted for the random assignment design—in particular, the assignment of schools rather than individuals to the treatment and control groups, as well as the pairing of schools before random assignment. To improve precision, regressions also controlled for pre-implementation school averages of student characteristics. We estimated regressions separately by year and used weights for educators' or students' data to give each school equal weight, so that the estimates reflected the impact of pay-for-performance on an average study school after one and two years of TIF implementation. We present the average outcomes for the treatment group as regression-adjusted means. That is, we present the raw (unadjusted) average outcomes for the control group, and we compute the regression-adjusted treatment group mean as the sum of the control group mean and the estimated impact.

Findings: The study found that among all 2010 TIF districts,

- **Full implementation of TIF continued to be a challenge, although districts' implementation from the first to the second year improved somewhat.** Although 90 percent of all TIF districts in 2012–2013 reported implementing at least 3 of the 4 required components for teachers, only about one-half (52 percent) reported implementing all four. This was a slight improvement from the first year of implementation, when 85 percent of districts reported implementing at least 3 of the 4 required components and 46 percent reported implementing them all.

For the 10 evaluation districts that completed two years of TIF implementation (the 2011–2012 and 2012–2013 school years), the key findings include the following:

- **Few evaluation districts structured pay-for-performance bonuses to align well with TIF grant guidance.** Overall, the bonuses were not very substantial or challenging to earn in these districts. The average teacher bonus was about \$1,800 (equal to 4% of average teacher salary, less than the 5% recommended in the grant for substantial bonuses). Each year, more than 60% of teachers in the schools that offered pay-for-performance bonuses received one. However, the bonuses were differentiated, based on the Department's guidance. The highest-performing teachers received a bonus of about \$7,000, more than 3 times the average bonus (Figure 2). (Insert Figure 2)
- **Educators' understanding of key program components improved from the first to the second year, but many teachers still misunderstood whether they were eligible for performance bonuses or the amount they could earn.** In schools that offered pay-for-performance bonuses, teachers' and principals' understanding of their eligibility for performance bonuses improved substantially (from 49 to 62 percent for teachers and 55 to 90 percent for principals, Figure 3). However, this also means that 38 percent of teachers in the second year still did not understand that they were eligible for a bonus. Teachers also continued to underestimate how much they could earn from performance bonuses, reporting a maximum bonus that was only two-fifths the size of the actual maximum bonuses awarded. (Insert Figure 3)
- **Pay-for-performance had small, positive impacts on students' reading achievement; impacts on students' math achievement were not statistically significant but similar in magnitude.** Students in treatment schools scored 0.03 standard deviations higher on reading assessments in Years 1 and 2 than students in control schools (Table 1). This difference is equivalent to 1 percentile point of achievement and a gain of about three additional weeks of learning. In math, differences in student achievement between treatment and control schools were also positive and similar in magnitude as those in reading, but not statistically significant. (Insert Table 1)

Conclusions: A primary objective of TIF grants is to raise student achievement in high-need schools. Based on the experiences of ten districts that participated in the national evaluation and completed two years of program implementation, the pay-for-performance component of TIF made a small contribution toward achieving this objective. Pay-for-performance bonuses generated slightly higher student reading achievement, and gains in math were similar in magnitude but not statistically significant.

The driving principle behind TIF is that increasing educator effectiveness is the key to raising student achievement and pay-for-performance bonuses are one way to increase educator effectiveness. We confirmed that the positive impact of pay-for-performance on student achievement was also reflected in positive impacts on educator effectiveness, as measured by the effectiveness ratings that educators received from their districts. Increases in educator effectiveness could have occurred either because teachers and principals improved their own effectiveness or because staffing changes resulted in more effective educators choosing to work at schools with pay-for-performance. We found little evidence for changes in staffing among teachers. Among principals, we found some evidence that pay-for-performance caused more high performers to stay at their schools and more low performers to leave their schools after the first year of TIF implementation.

Evidence from future years will provide more clarity on whether, over a longer period, the impacts of pay-for-performance evolve as educators continue accumulating more understanding of and experience with this program.

Appendices

Appendix A. References

n/a

Appendix B. Tables and Figures

Figure I. Random Assignment Design

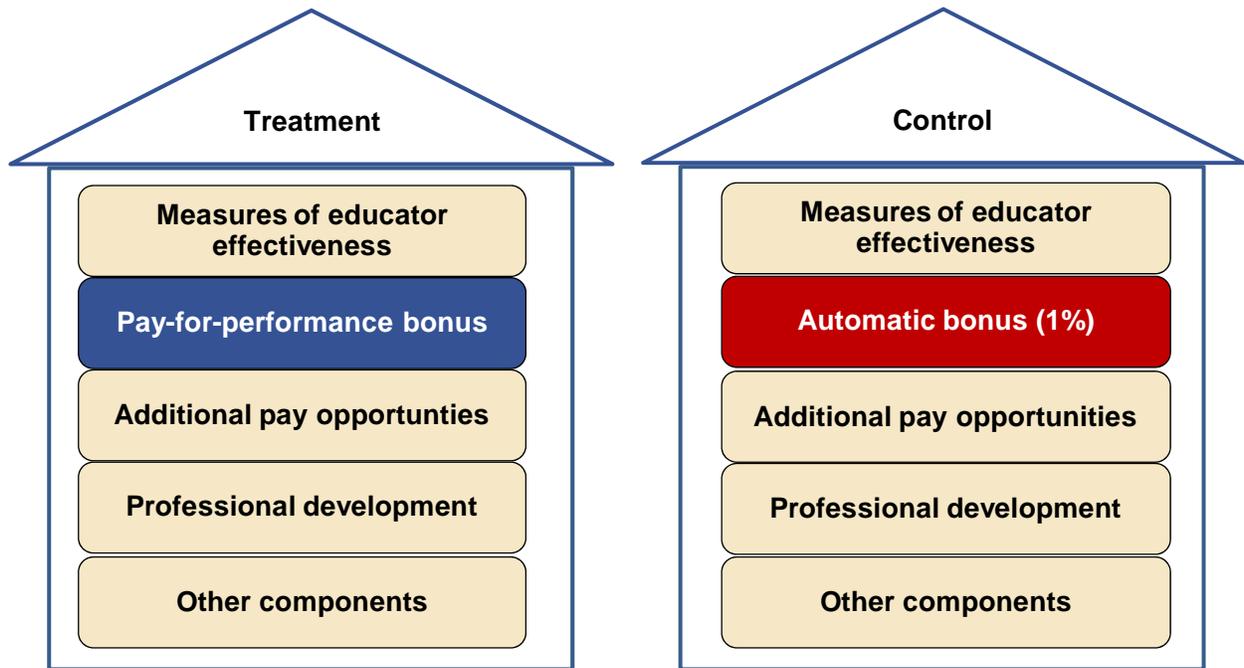
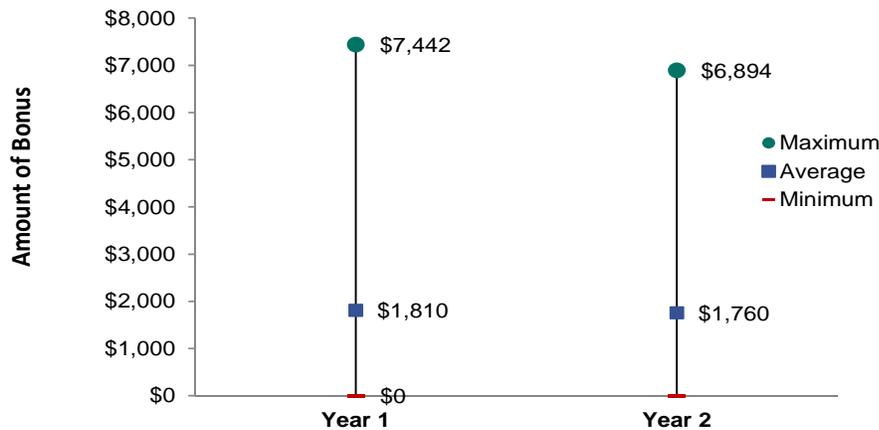


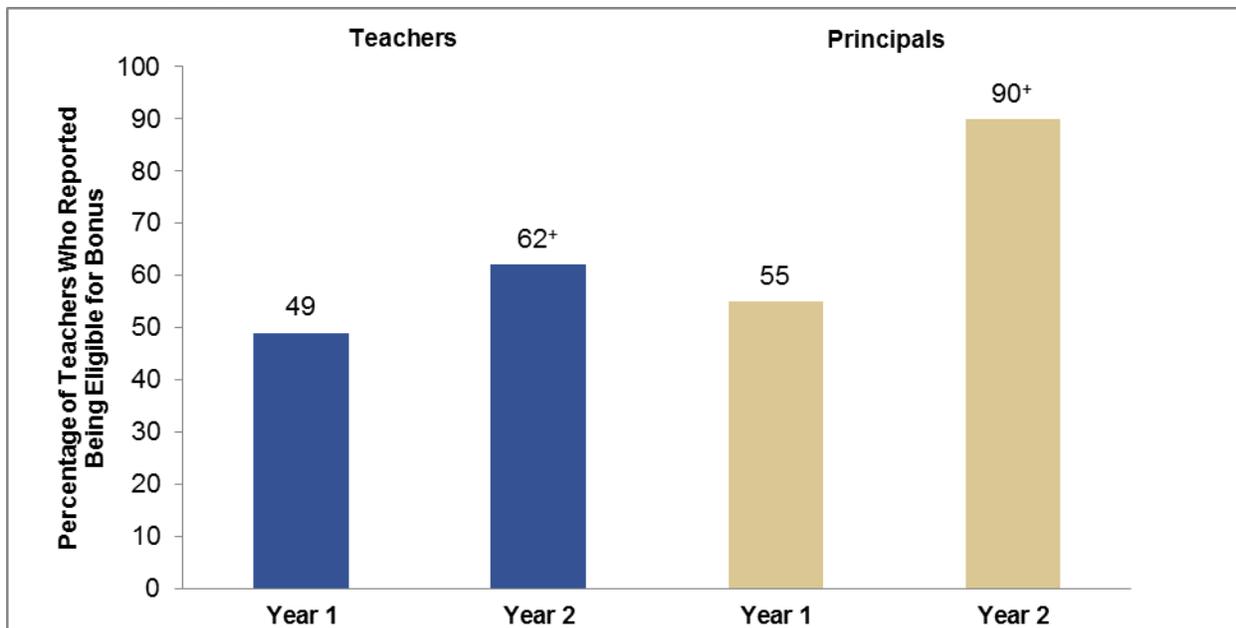
Figure 2. Minimum, Average, and Maximum Pay-for-Performance Bonuses for Teachers in Years 1 and 2



Source: District administrative data (N = 2,189 teachers in Year 1; N = 2,207 teachers in Year 2).

Figure reads: In Year 2, on average across the evaluation districts, the minimum pay-for-performance bonus for teachers was \$0, the average pay-for-performance bonus for teachers was \$1,760, and the maximum pay-for-performance bonus for teachers was \$6,894.

Figure 3. Percentages of Teachers and Principals in Schools that Offered Pay-for-Performance Bonuses Who Reported Being Eligible for Performance Bonuses



Source: Teacher and principal surveys, 2012 and 2013 (N = 377 teachers in Year 1; N = 444 teachers in Year 2; N = 64 principals in Year 1; and N = 63 principals in Year 2).

*Difference between 2011–2012 and 2012–2013 is statistically significant at the .05 level, two-tailed test.

Figure reads: Among teachers in schools with pay-for-performance, 49 and 62 percent reported being eligible for a pay-for-performance bonus in Year 1 and Year 2, respectively.

Table 1. Student Achievement in Math and Reading (Student z-score units)

Year and Subject	Treatment	Control	Impact	p-value	Number of Students	Number of Schools
Year 1						
Math	-0.43	-0.45	0.02	0.335	40,852	132
Reading	-0.37	-0.40	0.03*	0.040	40,576	132
Year 2						
Math	-0.39	-0.43	0.04	0.068	40,709	132
Reading	-0.36	-0.39	0.03*	0.026	40,391	132

Source: Student administrative data.

Note: The difference between the treatment and control estimates may not equal the impact shown in the table due to rounding.

*Impact is statistically significant at the .05 level, two-tailed test.