Evaluation of the Teacher Incentive Fund: Final Report on Implementation and Impacts of Pay-for-Performance Across Four Years: Executive Summary

December 2017

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

U.S. Department of Education

THIS PAGE IS INTENTIONALLY BLANK

Evaluation of the Teacher Incentive Fund: Final Report on Implementation and Impacts of Pay-for-Performance Across Four Years: Executive Summary

December 2017

Hanley Chiang
Cecilia Speroni
Mariesa Herrmann
Kristin Hallgren
Paul Burkander
Alison Wellington
Mathematica Policy Research

Elizabeth Warner
Project Officer
Institute of Education Sciences

NCEE 2018-4005
U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

U.S. Department of Education
Betsy DeVos
Secretary

Institute of Education Sciences
Thomas W. Brock
Commissioner of the National Center for Education Research
Delegated Duties of the Director

National Center for Education Evaluation and Regional Assistance
Ricky Takai
Acting Commissioner

December 2017

The report was prepared for the Institute of Education Sciences under Contract No. ED-IES-14-C-0115. The project officer is Elizabeth Warner in the National Center for Education Evaluation and Regional Assistance.

Chiang, Hanley, Cecilia Speroni, Mariesa Herrmann, Kristin Hallgren, Paul Burkander, Alison Wellington. (2017). *Evaluation of the Teacher Incentive Fund: Final Report on Implementation and Impacts of Pay-for-Performance Across Four Years: Executive Summary* (NCEE 2018-4005). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

This report is available on the IES website at http://ies.ed.gov/ncee.

Upon request, this report is available in alternate formats such as Braille, large print, audiotape, or computer diskette. For more information, please contact the Department's Alternate Format Center at 202-260-9895 or 202-205-8113.

ACKNOWLEDGMENTS

This study would not have been possible without the contributions of many individuals. We are grateful for the cooperation of many TIF administrators, teachers, principals, district leaders, and central office staff who assisted with the study's data collection and provided important information that shaped the study. A dedicated technical assistance team helped TIF districts implement the programs examined in this study. This team was led by Duncan Chaplin and Jeffrey Max and included Lauren Akers, Kevin Booker, Julie Bruch, Albert Liu, Allison McKie, Debbie Reed, Alex Resch, Christine Ross, and Margaret Sullivan from Mathematica and Patrick Schuermann and Eric Hilgendorf from the Peabody College of Education at Vanderbilt University.

Several individuals made enormous efforts to collect data successfully for this study. Sheila Heaviside and Annette Luyegu provided excellent leadership over our administration of teacher, principal, and district surveys, and Kathy Shepperson oversaw the design of key systems for collecting this survey data. Lauren Akers, Nickie Fung, Chris Jones, Margaret Sullivan, Sarah Wissel, and Claire Smither Wulsin patiently conducted and summarized numerous interviews with TIF administrators. Acquiring and processing administrative data required a large effort led by Jacqueline Agufa and Mary Grider with assistance from Michael Brannan, Kai Filipczak, Chris Jones, William Leith, Mickey McCauley, Sarah Osborn, and Juha Sohlberg.

Many people contributed to the analysis and interpretation of the study's data and the production of this report. The study received useful advice from our technical working group, consisting of David Heistad, James Kemple, Daniel McCaffrey, Anthony Milanowski, Richard Murnane, Jeffrey Smith, and Jacob Vigdor. At Mathematica, Jill Constantine and Steven Glazerman helped shape the evaluation and provided expert advice. Sarah Osborn helped with a variety of critical tasks including facilitating the management of the project. The analysis was made possible by an excellent team of programmers, consisting of Raúl Torres Aragon, Michael Brannan, Molly Crofton, Kai Filipczak, and John Hotchkiss. John Kennedy edited the report, and Jill Miller carefully and patiently prepared the report for publication.

THIS PAGE IS INTENTIONALLY BLANK

EXECUTIVE SUMMARY

Research indicates that effective teachers are critical to raising student achievement. However, there is little evidence about the best ways to improve teacher effectiveness, or how schools that serve the students most in need can attract and retain effective teachers. Traditional salary schedules, which pay teachers based on their years of teaching experience and degree attainment, may not reward effective teaching or provide incentives for the most effective teachers to teach in high-need schools.

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 congressionally mandated study, conducted by the U.S. Department of Education's Institute of Education Sciences, examined the implementation of TIF programs in over 130 districts that received grants awarded in 2010. Ten of these districts, the evaluation districts, agreed to participate in a random assignment study of the pay-for-performance component of TIF. Within those districts, this evaluation provided a more in-depth examination of TIF implementation and measured the impacts of pay-for-performance bonuses on educator effectiveness and student achievement.

This report, the final report from the evaluation, covers all four years of program implementation (2011–2012 through 2014–2015) under the 2010 TIF grants. The main findings include:

- **Within the ten evaluation districts, pay-for-performance led to slightly higher student achievement in reading and math by the second year of implementation.** Student reading achievement was higher by 2 percentile points at the end of the first year in schools that offered pay-for-performance bonuses than in schools that did not. The total difference remained at 1 to 2 percentile points across the subsequent three years and was statistically significant in most years. From the second year onward, the total difference in math achievement was similar in magnitude, but was statistically significant in only one year. In both subjects, these differences were equivalent to about three to four weeks of learning. Impacts on student achievement also differed across districts and schools, but these differences were mostly unrelated to either district or school characteristics.

- **Most 2010 TIF districts implemented each individual component of the comprehensive, performance-based compensation system required under their grant, and about half implemented all four components for teachers.** Districts were required to (1) use measures of both student achievement growth and observations of classroom or school practices to evaluate teachers' and principals' effectiveness, (2) offer educators bonuses based on their performance, (3) offer educators opportunities to earn additional pay for taking on extra roles or responsibilities, and (4) provide professional development to help educators understand the measures on which they were evaluated and improve their performance on those measures. Starting from the first year and continuing through all subsequent years, nearly all the districts reported offering pay-for-performance bonuses (over 90 percent) and additional pay opportunities (over 85 percent). Fewer districts, but still over half, implemented the required measures of teacher effectiveness (about 80 percent) and principal effectiveness (about 60 to 70 percent) and offered the required professional development to their teachers (59 to 74 percent). The percentage of districts that implemented all four components for teachers was similar across the four years (45 to 52 percent).

- **Many 2010 TIF districts reported that sustainability of their program was a major challenge, and slightly fewer than half planned to offer pay-for-performance bonuses after their grant ended.** In each year, about half or more of the districts reported that sustainability of the TIF program was a major challenge (63 percent in the second year, 48 percent in the third year, and 58 percent in the fourth year). Consistent with these concerns, slightly fewer than half (47 percent) of the districts planned to offer bonuses to teachers based on their performance in the 2015–2016 school year, the year after their grant ended. However, most districts planned to continue other key components of their program, including measures of teacher effectiveness similar to those used in TIF (at least 80 percent), additional pay for taking on extra roles or responsibilities (74 percent), and professional development based on teachers' actual performance ratings to help improve their instructional practices (90 percent).

Background on TIF and This Study

TIF Grants and Requirements

Across four rounds of TIF grants (in 2006, 2007, 2010, and 2012), the U.S. Department of Education awarded about $1.8 billion to help states and districts create comprehensive, performance-based compensation systems for teachers and principals.[1]

Box ES.1. Required Components of 2010 TIF Programs

Measures of educator effectiveness based on student achievement growth and at least two observations of classroom or school practices

Pay-for-performance bonuses that were substantial in size, differentiated, challenging to earn, and based solely on educators' effectiveness

Additional pay opportunities for educators to take on extra roles or responsibilities, such as becoming a master or mentor teacher who directly counsels other teachers

Professional development to help educators understand how they were evaluated and to provide feedback based on educators' actual performance ratings to improve their instructional practices

The 2010 TIF grant notice differed from the other rounds in that it included a main and an evaluation competition. Applicants for evaluation grants had to meet the same requirements for the performance-based compensation system as non-evaluation grantees (Box ES.1), but were subject to some additional requirements and guidance. One important requirement was that evaluation grant applicants had to agree to participate in a random assignment evaluation of pay-for-performance bonuses in which only half of their participating schools would offer those bonuses. In addition, evaluation grantees received more specific guidance about the structure of their 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 bonus), and *challenging* to earn (with only those performing significantly better than average receiving bonuses).

[1] The 2015 reauthorization of the Elementary and Secondary Education Act renamed TIF the Teacher and School Leader Incentive Grants program.

The TIF Study

This study addressed four research questions:

1. What were the characteristics of all TIF districts and their performance-based compensation systems? What implementation experiences and challenges did TIF districts encounter?
2. How did teachers and principals in schools that did and 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 did pay-for-performance bonuses affect educator effectiveness and the retention and recruitment of high-performing educators?
4. What was the impact of pay-for-performance bonuses on students' achievement on state assessments in math and reading?

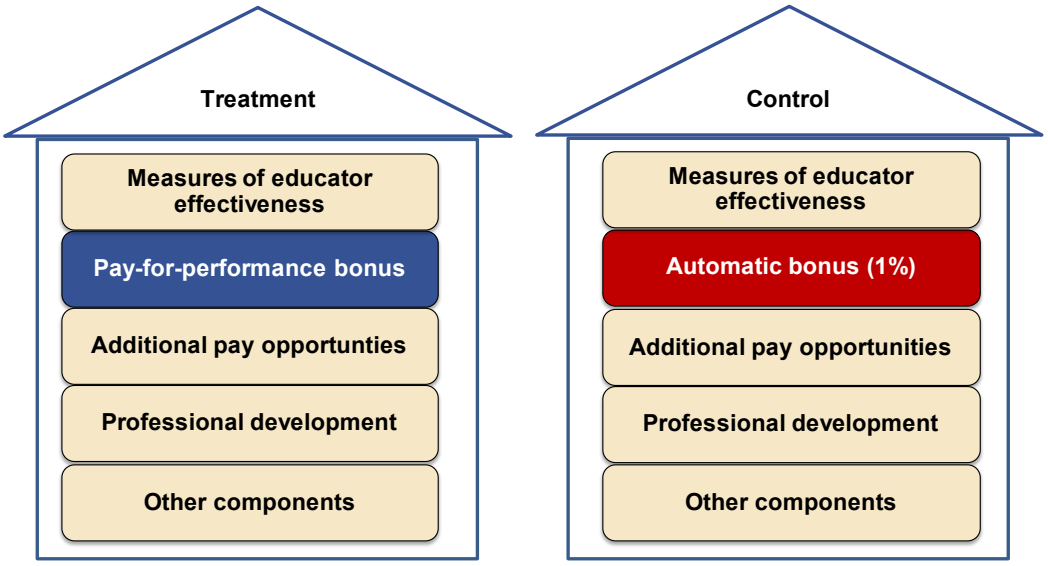
This report addresses each of the research questions with information from all four years of TIF implementation (2011–2012 to 2014–2015). Previous reports from this study (Max et al. 2014; Chiang et al. 2015; Wellington et al. 2016) covered prior years within the grant period.

Districts and schools in the study. The study provides a broad overview of program implementation by all 2010 TIF districts based on districts that had a TIF program in place in each year—153 districts in the first year, 155 in the second year, 144 in the third year, and 139 in the fourth year. This report's in-depth analyses of TIF implementation and the effects of pay-for-performance were based primarily on the ten evaluation districts and 131 schools in grades 3 to 8 that completed all four years of TIF implementation.

Compared with the average U.S. district, TIF districts were larger, were more likely to be urban and located in the South, and had a higher proportion of students who were racial or ethnic minorities and eligible for free or reduced-price lunches (Max et al. 2014). On average, evaluation districts were larger than non-evaluation TIF districts and had smaller percentages of students who were white (38 versus 50 percent). Evaluation districts were also more likely than non-evaluation districts to be in urban areas (69 versus 29 percent) and less likely to be in states with right-to-work laws (54 versus 67 percent).

Random assignment study design. To assess the impacts of pay-for-performance on educator and student outcomes, the study team assigned schools randomly—that is, completely by chance—to offer pay-for-performance bonuses (treatment group, 65 schools) or not (control group, 66 schools). As shown in Figure ES.1, treatment and control schools were expected to implement all other components of TIF. Because the two groups were expected to differ only in the opportunity for educators to receive pay-for-performance bonuses, differences in outcomes between the groups could be attributed to the impact of pay-for-performance. Specifically, the study measured the impact of pay-for-performance bonuses implemented within the context of broader performance-based compensation systems.

Figure ES.1. Random Assignment Evaluation Design



Data sources. Data for this report came from multiple sources (Table ES.1), all with response rates greater than 85 percent. The 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.

Table ES.1. Main Data Sources for This Report

Data Source	How the Study Used this Data Source
Data Collected from Evaluation and Non-Evaluation Districts	
District survey	Describe broadly the districts' programs and implementation challenges
Data Collected from Evaluation Districts Only	
District interview	Obtain more detail on districts' programs and implementation challenges
Principal and teacher surveys	Describe educators' understanding of TIF and measure impacts of pay-for-performance on educators' attitudes
Student administrative data	Measure impacts of pay-for-performance on student achievement
Educator administrative data	Describe teachers' and principals' performance ratings and bonuses; measure impact of pay-for-performance on educator performance ratings

Analysis approach. To describe districts' implementation of TIF, we calculated averages (such as the average of the largest bonuses districts awarded) and percentages (such as the percentage of districts that used specific effectiveness measures). To measure the impacts of pay-for-performance on educator and student outcomes, we compared outcomes in treatment and control schools.

Additional Findings from the Evaluation Districts

To provide context for interpreting the student achievement impacts that we found within the evaluation districts, we examined the districts' implementation of TIF in detail and measured impacts on a variety of other outcomes that could shape the student achievement impacts. In particular, we examined (1) the districts' overall implementation of the TIF required components, (2) the measures used to evaluate educators' effectiveness, (3) the structure of pay-for-performance bonuses, (4)

educators' understanding of and experiences with key components, (5) impacts on teachers' attitudes, and (6) impacts on educator effectiveness.

Overall Implementation of TIF Required Components

Most evaluation districts reported implementing all required components for teachers at the start of their programs. In each year, the only component for teachers that was not implemented by all of the districts was the professional development required by the grant. Starting from the first year, all evaluation districts reported using student achievement growth and at least two observations by trained observers to evaluate teachers, and by the second year, all districts used such measures to evaluate principals as well. In all years, every district also offered bonuses based on how educators performed on the effectiveness measures and offered additional pay to take on extra roles or responsibilities. However, only 6 to 8 of the 10 evaluation districts reported providing the required professional development for teachers, with no clear pattern of increasing or decreasing implementation over the course of the grant.

Measures of Educator Effectiveness

When evaluating educators, districts were required to use student achievement growth and observation measures, but they were given discretion to design the details of these measures. For example, when evaluating teachers based on student achievement growth, districts could measure the achievement growth of the teachers' own students (classroom achievement growth); all students in the same grade, team, or subject area (achievement growth of student subgroups); all students in the school (school achievement growth); or some combination of these measures. When considering their performance ratings from multiple measures, teachers may be more likely to believe the ratings are meaningful and accurate if the ratings are based on their individual performance, are more similar across measures, and are consistent across years.

All evaluation districts reported using observations and achievement growth to evaluate teachers, as required, and most used classroom achievement growth. More than half (60 to 70 percent) of the districts reported evaluating teachers based on classroom achievement growth. Within these districts, about 40 to 60 percent of teachers received classroom achievement growth ratings, typically because they taught grades and subjects tested by state assessments. All districts also used school achievement growth.

Most teachers received similar performance ratings from one year to the next, with many teachers receiving higher ratings on classroom observations than on achievement growth. For example, more than half of teachers received similar ratings, based on a 1-to-4 rating scale, in Year 4 as they did in Year 3 on classroom observations (56 percent), school achievement growth (51 percent), and classroom achievement growth (51 percent). However, in each year, teachers often earned higher ratings on observations than on achievement growth. For example, in Year 4, more than half (56 percent) of teachers received a higher rating on observations than on school achievement growth, and about half (49 percent) received a higher rating on observations than on classroom achievement growth.

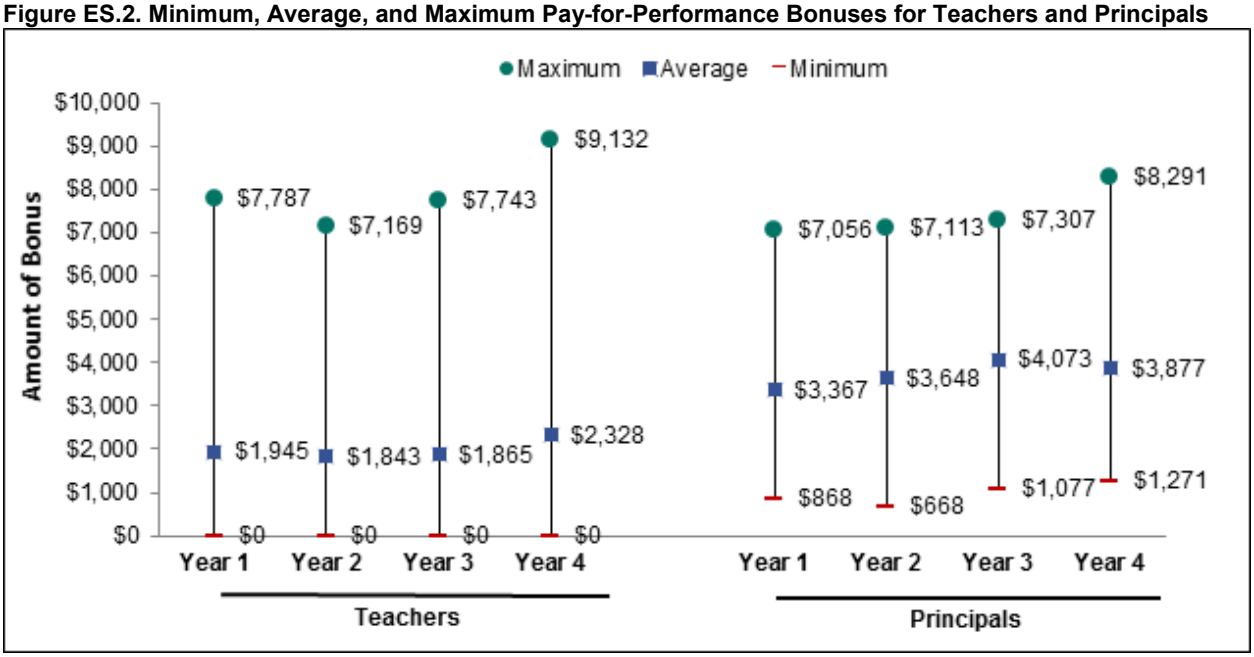
Structure of Pay-for-Performance Bonuses

To enhance the potential for performance bonuses to motivate teachers and principals to improve, the TIF notice required that the bonuses had to be challenging to earn, substantial in size, and differentiated.

Most teachers and principals received a bonus, a finding inconsistent with making bonuses challenging to earn. The average bonus fell somewhat short of the guidance to make bonuses substantial. The structure of the bonuses was similar across all four years of TIF implementation. In each year, within schools that offered performance bonuses, most teachers (about 70 percent) received a bonus, and the average bonus was about $2,000 (Figure ES.2), representing about 4 to 5 percent of the average teacher salary. Similar to teachers, most principals (more than 70 percent) received a bonus, and the average bonus (about $4,000) was about 3 to 5 percent of their average salary.

Bonuses were differentiated, with the highest-performing teachers earning bonuses significantly larger than the average bonus. In each year, the maximum bonus for teachers (about $7,000 to $9,000) was roughly four times the average bonus (Figure ES.2). For principals, bonuses were less differentiated, with the maximum bonus consistently about twice the average bonus.

Figure ES.2. Minimum, Average, and Maximum Pay-for-Performance Bonuses for Teachers and Principals



Source: Educator administrative data (N = 2,151 teachers in Year 1, N = 2,160 teachers in Year 2, N = 2,236 teachers in Year 3, and N = 2,083 in Year 4; N = 65 principals in Year 1, N = 68 principals in Year 2, N = 65 principals in Year 3, and N = 64 principals in Year 4).

Figure reads: In Year 1, on average across the evaluation districts, the minimum pay-for-performance bonus for teachers was $0, the average pay-for-performance bonus was $1,945, and the maximum pay-for-performance bonus was $7,787.

Educators' Understanding of and Experiences with Key Components

For the components of TIF to lead to improvements in educators' practices, districts had to effectively communicate information about those components to educators, and educators needed to know how to improve their performance.

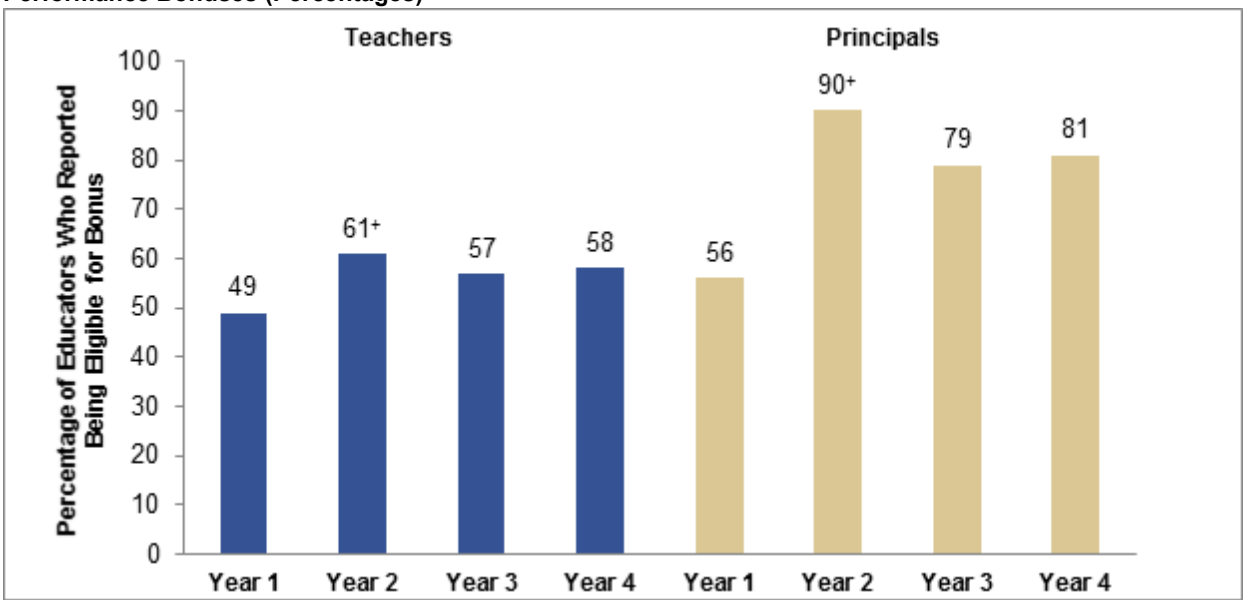
Most teachers were aware of being evaluated based on student achievement growth and classroom observations early in TIF implementation, and their awareness of these performance measures improved over time. At least 70 percent of teachers in Year 1 reported being evaluated on student achievement growth, and more than 70 percent reported being evaluated on at least two classroom observations. Furthermore, the percentage of teachers who reported being evaluated on these measures increased over time, with significant improvements in some years for both treatment and control teachers. By Year 4, 89 percent of treatment teachers and 78 percent of control teachers reported being evaluated on student achievement growth, and more than 85 percent of teachers reported being evaluated on at least two classroom observations. Similar to teachers, about 85 percent of principals in Year 4 reported being evaluated on student achievement growth; however, a smaller percentage of principals (less than 60 percent) reported being evaluated on at least two observations of their school practices.

Many teachers in treatment schools did not understand that they were eligible to earn a performance bonus, and their understanding did not improve after the second year of implementation. Although teachers' and principals' understanding of their bonus eligibility improved significantly from Year 1 to Year 2, there was no further improvement beyond Year 2. In Years 2 through 4, about 60 percent of treatment teachers (for example, 58 percent in Year 4) were aware that they could potentially earn a performance bonus, implying that about 40 percent were still unaware (Figure ES.3). Although understanding of eligibility was better among principals than teachers, about 20 percent of principals in Year 4 still did not know they were eligible to earn a bonus based on their performance.

Teachers underestimated how much they could earn from a performance bonus. In each year, teachers in treatment schools believed that the maximum bonus they could earn was no more than 40 percent of the actual maximum bonus districts awarded (Figure ES.4). Principals also underestimated the potential amount of performance bonuses they could receive, but their beliefs better aligned with actual bonus payouts than did teachers' beliefs. Across years, the maximum pay-for-performance bonus that principals reported they could receive ranged from 67 to 91 percent of the actual maximum bonus that districts awarded.

Most teachers reported receiving professional development required under the TIF grant but indicated they received only a few hours of it over the school year. In each year, more than half of teachers reported that they received or expected to receive professional development focused on understanding performance measures used in TIF (ranging from 77 percent of treatment teachers in Year 1 to 53 percent in Year 4). Most teachers (50 to 60 percent) also reported receiving or expecting to receive feedback based on their performance ratings. Of those who expected to receive any professional development on these two topics, the expected amount of time on each topic per year ranged from two to six hours.

Figure ES.3. Teachers and Principals in Treatment Schools Who Reported Being Eligible for Pay-for-Performance Bonuses (Percentages)

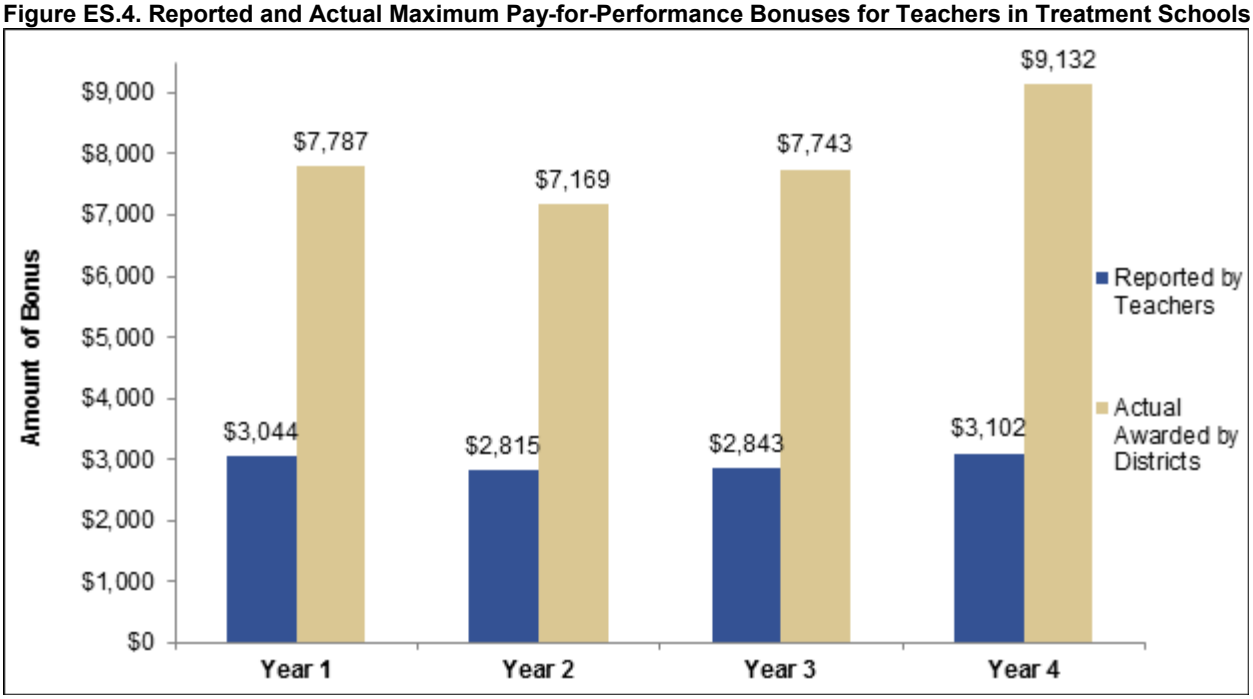


Sources: Teacher and principal surveys, 2012, 2013, 2014, and 2015 (N = 368 teachers in Year 1; N = 439 teachers in Year 2; N = 420 teachers in Year 3; N = 391 teachers in Year 4; N = 63 principals in Year 1; N = 62 principals in Year 2; N = 57 principals in Year 3; and N = 60 principals in Year 4).

Figure reads: In Year 1, 49 percent of teachers in treatment schools reported being eligible for a pay-for-performance bonus.

+Difference with prior year is statistically significant at the .05 level, two-tailed test.

Figure ES.4. Reported and Actual Maximum Pay-for-Performance Bonuses for Teachers in Treatment Schools



Sources: Teacher surveys, 2012, 2013, 2014, and 2015 (N = 218 teachers in Year 1; N = 229 teachers in Year 2; N = 229 teachers in Year 3; and N = 210 teachers in Year 4) and educator administrative data.

Figure reads: In Year 1, on average, the actual maximum pay-for-performance bonus that evaluation districts awarded to teachers was $7,787, and the maximum pay-for-performance bonus teachers reported they could earn was $3,044.

Impacts of Pay-for-Performance on Teachers' Attitudes

To the extent that pay-for-performance enhances or worsens teachers' job satisfaction, their motivation to improve could increase or decrease as a result.

Most teachers were satisfied with their jobs and the TIF program. In each year of TIF implementation, at least two-thirds of teachers in both treatment and control schools reported being satisfied with their job overall and were glad to be participating in TIF. On each specific aspect of their professional opportunities, evaluation system, and school environment, at least half of teachers reported being satisfied.

Although initially less satisfied with their jobs and TIF, teachers in treatment schools were as satisfied as, and sometimes more satisfied than, teachers in control schools by the third year of implementation. In the first two years of TIF implementation, teachers in treatment schools tended to report being less satisfied than teachers in control schools. For example, in Year 2, treatment teachers reported being less satisfied than control teachers with the use of student achievement scores to assess their performance (61 versus 69 percent) and recognition of their accomplishments (61 versus 66 percent). They were also more likely to report that TIF reduced their freedom to teach (39 versus 30 percent) and harmed collaboration (29 versus 21 percent). However, in Years 3 and 4, treatment teachers were no longer less satisfied than control teachers on any of the measured dimensions and, in fact, reported being more satisfied on some dimensions. For example, in Year 4, more treatment teachers than control teachers were satisfied with feedback on their performance (87 versus 77 percent) and believed TIF caused teachers to work more effectively (59 versus 51 percent).

Impacts of Pay-for-Performance on Educator Effectiveness

The rationale behind pay-for-performance is that it can improve student achievement by enhancing educator effectiveness. In light of the positive impacts on student achievement, we examined whether pay-for-performance led to improved classroom practices in ways that were detected by trained observers and were related to higher student achievement.

Pay-for-performance led teachers to earn slightly higher classroom observation ratings by the third year of implementation. Differences between the classroom observation ratings of teachers in treatment schools and those in control schools grew over the four years of implementation and became statistically significant by Year 3. In Years 3 and 4, treatment teachers earned observation ratings that were 0.05 and 0.09 points higher on a 1-to-4 scale than those of control teachers. Pay-for-performance had no impact on the observation ratings of principals.

The impacts of pay-for-performance on classroom observation ratings did not appear to explain the impacts on student achievement. Schools that experienced larger impacts of pay-for-performance on observation ratings did not experience larger impacts on student math or reading achievement. Therefore, although classroom observation ratings detected some improvements in practices due to pay-for-performance, they did not identify the improvements that were actually associated with student achievement.

Concluding Thoughts

Overall, the 2010 TIF districts were able to implement most components of a comprehensive, performance-based compensation system, typically starting early in their grant period and lasting throughout all four years of their programs. However, many districts anticipated that sustaining the TIF program beyond the life of their grant would be difficult. In particular, fewer than half of the 2010 TIF districts planned to continue offering performance bonuses.

A primary objective of TIF grants is to raise student achievement in high-need schools. Within the evaluation districts, this study found that the pay-for-performance component of TIF had small, positive impacts on student achievement by the second year of implementation. From that year onward, reading and math achievement was higher by 1 to 2 percentile points in schools that offered performance bonuses than in schools that did not.

To draw lessons from these impact findings for future policies on performance-based compensation, it is useful to consider possible explanations for why performance bonuses within the TIF program had any positive impacts on student achievement and why those impacts were small. According to the rationale behind pay-for-performance bonuses, they can improve student achievement only if educators (1) face a bonus structure that provides meaningful incentives for improvement, (2) understand key components of the program, (3) feel motivated to adjust their practices or their choice of where to work to earn these bonuses, and (4) know how to change their practices in ways that improve student achievement. As we summarize next, some, but not all, of these factors were in place within the evaluation districts.

First, the structure of the bonuses provided a mix of stronger and weaker incentives for teachers to improve. The highest-performing teachers could earn a performance bonus worth about four times the average bonus, which provided an incentive for teachers to demonstrate high performance. However, in each year, the criteria for earning any bonus resulted in about 70 percent of teachers earning bonuses within schools that offered them. Therefore, even some teachers who performed worse than the typical teacher earned a bonus. If failing to receive a bonus represented a clear signal about having room for improvement, the bonus structure gave a minority of teachers this type of encouragement to improve.

Second, educators' misunderstanding of these bonuses may have hampered the degree to which this policy could influence educators' behavior. Across all four years of implementation, many teachers were unaware they were eligible for a performance bonus or underestimated the amount they could earn. These teachers perceived limited or no monetary incentives to improve their performance even though their districts had actually structured the bonuses to provide stronger incentives.

Third, some teachers perceived schools that offered performance bonuses to be a more appealing place to work, potentially enhancing their motivation to remain at these schools and improve their practices. By the third year of TIF implementation, teachers' satisfaction with key aspects of their jobs was either unchanged or improved as a result of pay-for-performance. However, we have no evidence that these favorable impacts on teachers' job satisfaction contributed to improvements in student achievement. In fact, positive impacts on student achievement emerged in the first two years—a period when pay-for-performance was actually *lowering* satisfaction.

Fourth, it is unclear whether teachers knew how to change their classroom practices in ways that could improve student achievement. By the third year of implementation, pay-for-performance led to small increases in teachers' classroom observation ratings. However, schools that experienced larger impacts of pay-for-performance on observation ratings did not experience larger impacts on student achievement. Therefore, we found no indication that changes in teachers' measured practices were the source of the improvements in student achievement. The disconnect between changes in measured practices and changes in achievement could have been due to a number of factors, which this study did not have the data to examine. One possibility is that the amount of targeted professional development teachers received—no more than six hours over the school year—was insufficient to promote changes in practices that were substantial enough to improve student achievement. Another possibility is that the observation measures encouraged teachers to focus on aspects of instruction that were not related to student achievement.

Although the impacts of pay-for-performance were small, the costs of the bonuses were also low enough such that this policy was at least as cost-effective as some alternative policies that have been evaluated. Specifically, a cost-effectiveness analysis suggests that pay-for-performance was more cost-effective than class-size reduction (through four years of program implementation) and about as cost-effective as providing transfer incentives for high-performing teachers to move to low-performing schools (at the end of two years).[2] However, the available evidence cannot predict the policies' cost-effectiveness beyond the limited number of years in which these policies were implemented and evaluated.

[2] For example, after four years, raising student achievement by the actual impact of pay-for-performance required spending $499 per student on pay-for-performance, but would have required spending $767 per student on class-size reduction. However, to achieve the same impact as pay-for-performance did after two years, transfer incentives would require $193 per student, nearly identical to the $196 per-student cost of pay-for-performance.

THIS PAGE IS INTENTIONALLY BLANK

THIS PAGE IS INTENTIONALLY BLANK


DEPARTMENT OF EDUCATION
UNITED STATES OF AMERICA