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

Executive Summary

September 2015

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: Implementation and Impacts of Pay-for-Performance After Two Years

Executive Summary

September 2015

Hanley Chiang
Alison Wellington
Kristin Hallgren
Cecilia Speroni
Mariesa Herrmann
Steven Glazerman
Jill Constantine
Mathematica Policy Research

Elizabeth Warner
Project Officer
Institute of Education Sciences

NCEE 2015-4021
U.S. DEPARTMENT OF EDUCATION

U.S. Department of Education

Arne Duncan

Secretary

Institute of Education Sciences

Ruth Neild

Deputy Director for Policy and Research

Delegated Duties of the Director

National Center for Education Evaluation and Regional Assistance

Joy Lesnick

Acting Commissioner

September 2015

The report was prepared for the Institute of Education Sciences under Contract No. ED-04-CO-0112-0012. The project officer is Elizabeth Warner in the National Center for Education Evaluation and Regional Assistance.

IES evaluation reports present objective information on the conditions of implementation and impacts of the programs being evaluated. IES evaluation reports do not include conclusions or recommendations or views with regard to actions policymakers or practitioners should take in light of the findings in the reports.

This report is in the public domain. Authorization to reproduce it in whole or in part is granted. While permission to reprint this publication is not necessary, the citation should be:

Chiang, Hanley, Alison Wellington, Kristin Hallgren, Cecilia Speroni, Mariesa Herrmann, Steven Glazerman, Jill Constantine. (2015). *Evaluation of the Teacher Incentive Fund: Implementation and Impacts of Pay-for-Performance After Two Years, Executive Summary* (NCEE 2015-4021). 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, Margaret Sullivan, 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 with assistance from Michael Brannan, Dylan Ellis, Chris Jones, Serge Lukashanets, Jeremy Page, Nina Pudukollu, 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, Chris Jones helped with a variety of critical tasks ranging from communicating with TIF districts to 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, John Hotchkiss, and William Leith, with expert guidance from Mary Grider. Cindy George and John Kennedy oversaw the editing of the report, and Jill Miller carefully and patiently prepared the report for publication.

THIS PAGE IS INTENTIONALLY BLANK

EXECUTIVE SUMMARY

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. It examines grantees' programs and implementation experiences and the impacts of pay-for-performance bonuses on educator effectiveness and student achievement.

This report, the second from the study, describes the programs and implementation experiences of all 2010 TIF grantees in the 2012–2013 school year, the second of four years of implementation for nearly all grantees. The main findings for all districts that received 2010 TIF grants include the following:

- **Full implementation of TIF continues 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.
- **Near the end of the second year of implementation, most districts reported that sustainability of their TIF program was a major challenge; however, few reported other key activities related to their program were a major challenge.** By the end of 2012–2013, 65 percent of TIF districts reported that sustainability of the program was a major challenge. In contrast, no more than one-third of districts reported that other activities related to their program (such as incorporating student achievement growth into teacher evaluations or conducting observations) were a major challenge.

This report also provides detailed findings from a subset of 2010 TIF grantees, the evaluation districts, that participated in a random assignment study of the pay-for-performance component of TIF. For the ten evaluation districts that completed two years of TIF implementation, the report provides an in-depth analysis of TIF implementation and the impacts of pay-for-performance bonuses on educator and student outcomes after the first (2011–2012) and second (2012–2013) years. The main findings for the ten evaluation districts include the following:

- **Few evaluation districts structured pay-for-performance bonuses to align well with TIF grant guidance.** The grant notice provided guidance, although not specific requirements, about how to structure pay-for-performance bonuses to be substantial, differentiated, and challenging to earn. At least half of the evaluation districts (70 percent in Year 1 and 50 percent in Year 2) met the guidance for awarding differentiated performance bonuses for teachers. However, in each year, no more than 30 percent of districts awarded bonuses for teachers that were substantial or challenging to earn. Likewise, no more than 30 percent of districts awarded bonuses for principals that were differentiated, substantial, or challenging to earn.

- **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.** Teachers had a better understanding of how their performance was evaluated in the second year than in the first. For example, about 85 percent of teachers reported being evaluated on at least two classroom observations in the second year compared to about 75 percent of teachers in the first year. In schools that offered pay-for-performance bonuses, teachers' and principals' understanding of their eligibility for bonuses also improved (by 13 and 35 percentage points, respectively). However, many teachers in these schools (38 percent in the second year) still did not understand that they were eligible for a bonus. They 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.
- **Pay-for-performance had small, positive impacts on students' reading achievement; impacts on students' math achievement were not significant but similar in magnitude.** After two years of TIF implementation, the average reading score was 1 percentile point higher in schools that offered pay-for-performance bonuses than in schools that did not. This difference was equivalent to a gain of about three additional weeks of learning.

TIF Grants and Requirements

From 2006 to 2012, the U.S. Department of Education awarded about \$1.8 billion to support 131 TIF grants. Sixteen grants were awarded in 2006, 18 in 2007, 62 in 2010, and 35 in 2012.

The 2010 TIF grants differed from prior TIF grants by providing more detailed guidance on the measures used to evaluate educators and on the design of the pay-for-performance bonuses. The 2010 grants required performance-based compensation systems implemented in districts to include four components. This study focuses most heavily on examining the implementation and impacts of one of those requirements: pay-for-performance bonuses.

Required Program Components of the Performance-Based Compensation Systems

The four required TIF components are:

1. **Measures of educator effectiveness.** Grantees were required to measure the effectiveness of teachers and principals using students' achievement growth and at least two observations of classroom or school practices. They had discretion to include additional measures.
2. **Pay-for-performance bonus.** Grantees had to offer bonuses to educators based on how they performed on the effectiveness measures. The bonuses aimed to incentivize educators and reward them for being effective in their classrooms and schools. Bonuses had to be substantial, differentiated, challenging to earn, and based solely on educators' effectiveness.
3. **Additional pay opportunities.** The performance-based compensation system had to include pay opportunities for educators to take on additional roles or responsibilities.

These roles might include becoming a master or mentor teacher who directly counsels other teachers or develops or leads professional development sessions for teachers.

4. **Professional development.** TIF grantees were required to support teachers and principals in their performance improvement efforts. Support included providing information about measures on which educators would be evaluated and more targeted professional development based on an educator's actual performance on the effectiveness measures.

The TIF Grant Competition

The 2010 TIF grant notice differed from the other rounds in that it included a main and an evaluation competition (Max et al. 2014). By holding two separate competitions, the U.S. Department of Education identified a group of grantees that, by virtue of having applied for an evaluation grant, had indicated their interest and willingness to participate in a more in-depth evaluation of their TIF grants.

A key difference between the non-evaluation and evaluation grantees is that applicants for the evaluation grants 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 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.

Applicants for evaluation grants had to meet the same requirements for the performance-based compensation system as non-evaluation grantees and some additional requirements. One important requirement was that evaluation grant applicants had to agree to participate in a random assignment evaluation of pay-for-performance bonuses. Schools within a district were randomly assigned to implement either all four required components of the performance-based compensation system, including pay-for-performance bonuses (the treatment group), or all components *except* pay-for-performance bonuses (the control group).

The TIF Study

The purpose of this multiyear study is to describe the program characteristics and implementation experiences of 2010 TIF grantees and estimate the impact of pay-for-performance bonuses within a well-implemented, performance-based compensation system. Because educators' understanding of and responses to this policy can change over time, this study plans to follow the grantees for the full duration of the five-year grants.

The study is addressing four 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?

This report is the second of four planned reports from the study. The first report (Max et al. 2014) addressed the first two research questions based on information from the 2011–2012 school year. This second report uses information from the first (2011–2012) and second (2012–2013) years of TIF implementation to describe the ways in which evaluation districts structured the components of their programs and communicated information about those components (question 1). This report also captures the views, attitudes, and behaviors of educators as they evolved over two years of implementation (question 2) and presents initial impacts of pay-for-performance on educator effectiveness and student achievement after the first and second years (questions 3 and 4).

Districts in the Study

Although this report provides the greatest amount of information on the evaluation districts, it also provides a broad overview of TIF implementation by all 2010 grantees in the 2012–2013 school year. This analysis was based on 155 districts that participated in TIF in 2012–2013.

This report's in-depth analyses of TIF implementation and the effects of pay-for-performance on educator and student outcomes were based on information from the evaluation districts. Of the 13 evaluation districts, 10 completed two years of TIF implementation—2011–2012 and 2012–2013—during the period covered by the report. The remaining 3 evaluation districts completed their first year of TIF implementation in 2012–2013. This report focuses primarily on the 10 districts for which data were available on two years of TIF implementation. Focusing on districts that completed two years of TIF implementation enabled us to examine changes in educators' perceptions and practices from the first to the second year and assess whether impacts on educator and student outcomes also evolved during that time.

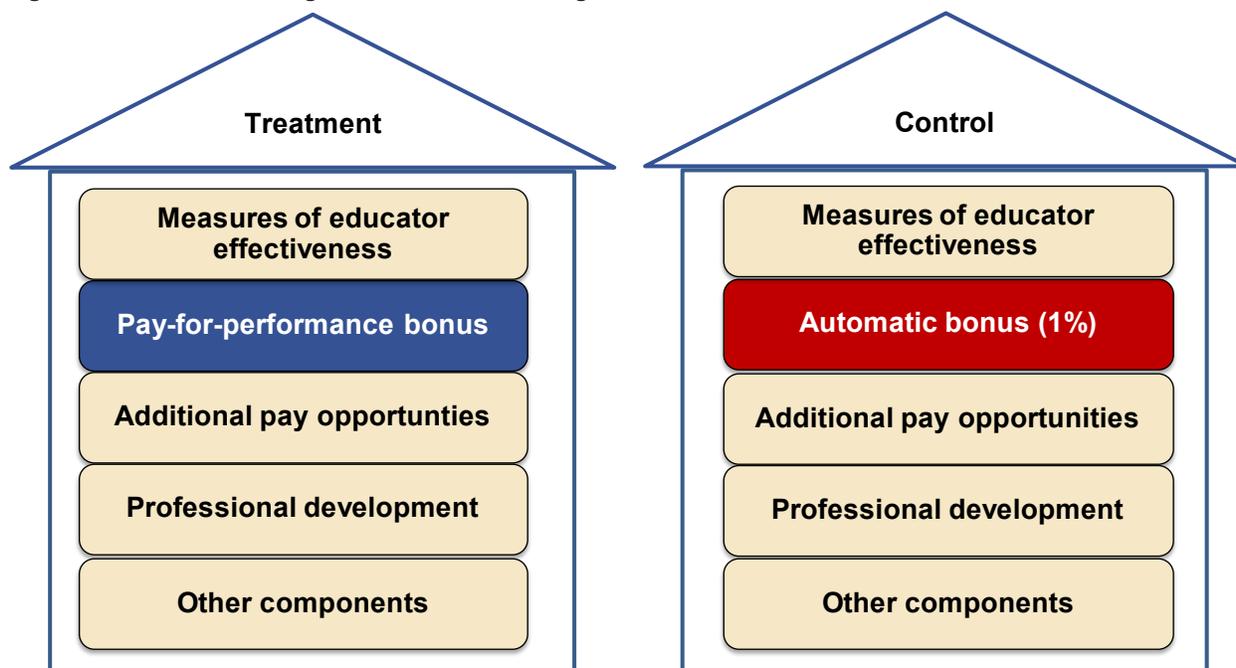
Experimental Study Design

The study used an experimental study design to assess the impacts of pay-for-performance on educator and student outcomes. Elementary and middle schools within the evaluation districts were assigned randomly—that is, completely by chance—to treatment and control groups. As shown in Figure ES.1, treatment and control schools were expected to implement the same required components of the district's performance-based compensation system, 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.

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.

Figure ES.1. Random Assignment Evaluation Design



The key advantage of this study's random assignment design is that, at the beginning of the study, the treatment and control groups were expected to include students and educators with similar characteristics. 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.

Schools in the Study

Analyses of educator and student outcomes were based on 132 schools—66 treatment schools and 66 control schools—that implemented the TIF program for two years. 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).

Data Sources

Data for this report came from multiple sources. 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.

Data on all 2010 TIF districts. The study team collected data on all TIF districts from two sources. First, to compare characteristics of evaluation and non-evaluation districts, the study team used information from the Common Core of Data. Second, to describe broadly the TIF program features that districts reported implementing and the challenges they encountered in implementation, the study team administered a survey to all TIF district administrators in 2011–2012 and 2012–2013.

Additional data on evaluation districts. We obtained more detail on TIF programs and implementation experiences from interviews with district staff and technical assistance documents. To examine educators' attitudes toward their job and the TIF program, the study team administered surveys to all principals and a sample of teachers in treatment and control schools in spring 2012 and spring 2013. We collected districts' administrative records on teachers and principals to describe their performance ratings, bonuses, and additional pay, as well as to examine the impact of pay-for-performance on educator effectiveness. Finally, to assess the impact of pay-for-performance on student achievement, the study team collected districts' administrative records on students enrolled in treatment and control schools.

Methods

The study team used several different methods to describe the implementation of TIF and measure the impact of pay-for-performance on educators' and students' outcomes.

Describing TIF implementation in all 2010 TIF districts. To describe broadly the program characteristics and implementation challenges reported by all 2010 TIF districts, we summarized their responses to the district survey with means or percentages, as appropriate.

Describing TIF implementation in evaluation districts. We conducted a variety of analyses to provide an in-depth description of TIF implementation in the evaluation districts. First, as in the analysis of all 2010 TIF districts, we summarized evaluation districts' survey responses about program characteristics and implementation challenges, but we also supplemented these data with information from telephone interviews and technical assistance documents. Second, to describe educators' actual bonus amounts and performance ratings, we summarized administrative data with means, maximum levels, or percentages of educators receiving particular bonus amounts or ratings. Third, to describe educators' understanding of and experiences with the required TIF components, we summarized educators' survey data, making comparisons between treatment and control schools and across years.

Measuring the impacts of pay-for-performance on educator and student outcomes. 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). For each outcome, we compared the outcomes of educators and students in treatment schools to those of educators and students in control schools. Because the study used random assignment, any differences in educator or student outcomes between the treatment and control groups could be attributed to pay-for-performance and not some other characteristic of the districts or schools.

Detailed Summary of Findings

Programs and Experiences of All 2010 TIF Districts

As a comprehensive program for reforming educator compensation and improving educator effectiveness, TIF programs were designed to have multiple, interrelated components. Our analysis of implementation in all 155 TIF districts sought to determine whether they could put into place such a comprehensive system, and whether they faced particular challenges doing so.

Full implementation of TIF continues to be a challenge, although districts' implementation from the first to the second year improved somewhat. Although 90 percent of all TIF districts in the second year (2012–2013) reported implementing at least 3 of the 4 required components for teachers, about one-half (52 percent) reported implementing all four. This was a slight improvement from the first year (2011–2012), when 85 percent of districts reported implementing at least 3 of the 4 required components and 46 percent reported implementing them all. More than half of the districts (58 percent in Year 1 and 60 percent in Year 2) implemented all required components for principals aside from professional development, a component for which data were not available.

Most districts implemented each individual required component of TIF, but were less likely to report offering targeted professional development and evaluating teachers and principals using both student achievement growth and at least two observations. In Year 2, nearly all the districts (over 90 percent) reported offering teachers and principals bonuses based on their performance and offering educators opportunities to earn additional pay (Table ES.1). In contrast, approximately three-quarters of the districts reported that they offered the required professional development to their teachers and 80 percent reported using both student achievement growth and classroom observations to measure teacher effectiveness (Table ES.1). Fewer districts (65 percent) reported using both student achievement growth and observations of school practices to measure principal effectiveness. Although all districts were expected to evaluate educators using student achievement growth, districts could choose how to measure student achievement growth to evaluate their educators. In Year 2, almost all TIF districts (about 88 percent) reported using an achievement growth measure based on all students in the school to evaluate teacher and principal effectiveness. Fewer districts (64 percent) reported evaluating teachers based on the achievement growth of only the students in their classrooms.

Near the end of the second year of implementation, most districts reported that sustainability of their TIF program was a major challenge; however, few reported other key activities related to their program were a major challenge. By the end of 2012–2013, 65 percent of TIF districts reported that sustainability of the program was a major challenge. In contrast, fewer than one-third of districts reported that linking student growth data to teachers (30 percent), explaining student achievement growth to teachers (28 percent), and calculating student achievement growth to evaluate teachers (28 percent) were major challenges. Likewise, only one-third of districts (33 percent) reported that providing useful and timely feedback on student achievement measures was a major challenge.

Table ES.1. Districts' Reported Implementation of TIF Required Components for Teachers in Year 2 (Percentages)

	All 2010 TIF Districts	Evaluation Districts
Requirements		
Requirement 1: Measures of educator effectiveness ^a	80	100
Requirement 2: Pay-for-performance bonus	98	100
Requirement 3: Additional pay opportunities	91	100
Requirement 4: Professional development	74	70
Implemented all requirements	52	70
Number of Districts—Range^b	142-155	10

Source: District surveys and district interviews, 2013.

^aTIF districts were required to use student achievement growth and at least two observations by trained observers to evaluate teachers and principals.

^bSample sizes are presented as a range based on the data available for each row in the table.

TIF Implementation in Evaluation Districts

In-depth information from the evaluation districts enabled the study team to examine, in greater detail, whether the components of their programs provided incentives and supports for educators to improve their effectiveness, and whether educators understood those components.

Program Implementation

As a first step, we examined the extent to which evaluation districts implemented the four required components. We also examined the types of measures that districts used to evaluate educators' effectiveness and described educators' actual performance on those measures, focusing on whether different measures were consistent with each other in assessing how well teachers performed.

Most evaluation districts reported implementing all required components for teachers. The only component not consistently implemented was professional development. In Year 2, all evaluation districts reported using measures of effectiveness for teachers and principals that included student achievement growth and at least two observations of classroom or school practices, offering bonuses based on how educators performed on effectiveness measures, and offering additional pay to take on extra roles or responsibilities. Seven of 10 evaluation districts reported providing the required professional development for teachers (Table ES.1).

All evaluation districts reported using the achievement growth of all students in a school to evaluate teachers, and some also chose to evaluate teachers based on the achievement growth of the students they teach. Slightly more than half (60 percent) of evaluation districts reported evaluating teachers based on achievement growth in their classrooms. Within these districts, fewer than half (about 40 percent) of teachers were evaluated on the achievement growth of their students.

Student achievement growth and observation ratings sometimes identified the same educators as high-performing, but many earned higher ratings on observations than on achievement growth. For example, in Year 2 teachers who scored high (in the top quarter of the rating scale) on achievement growth in their schools were twice as likely to score high on classroom

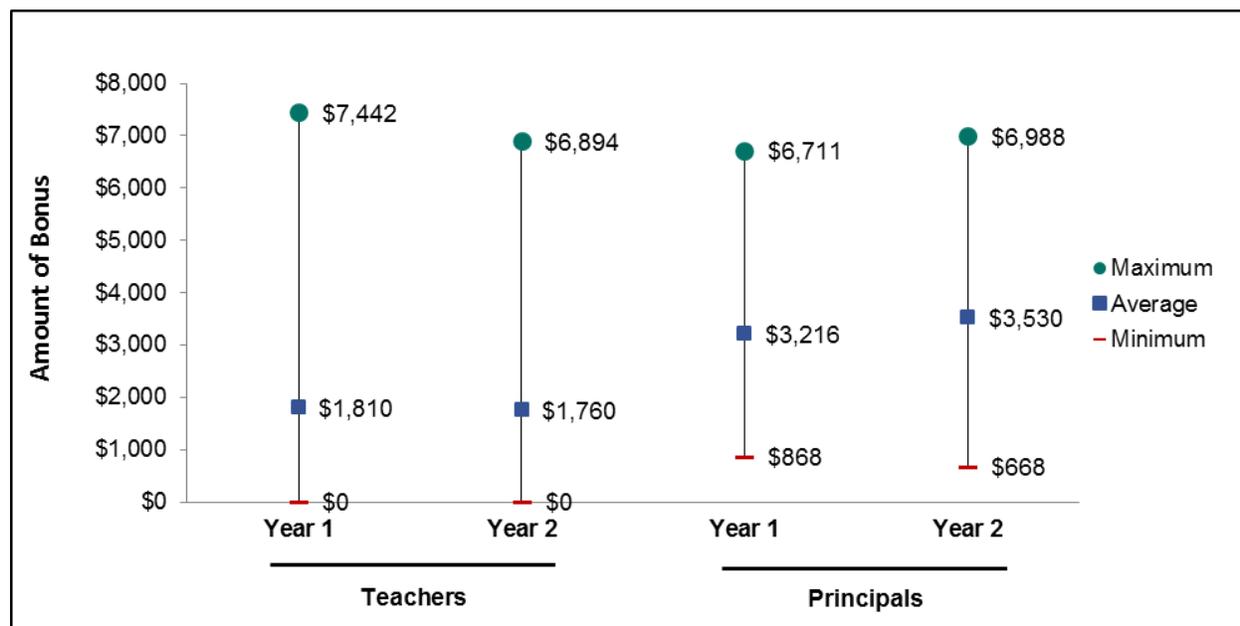
observations compared with teachers who scored low (in the bottom quarter of the rating scale) on achievement growth (38 versus 19 percent). Nevertheless, many teachers (87 percent) who scored low on achievement growth earned at least moderately high ratings on classroom observations, scoring in the top half of the observation rating scale. Likewise, many principals (78 percent) who scored low on achievement growth earned at least moderately high ratings on observations.

Pay-for-Performance Bonuses

The purpose of offering performance bonuses to teachers and principals was to motivate them to improve and reward educators for being effective in their classrooms and schools. To achieve this objective, the TIF notice required that the bonuses had to be substantial, differentiated, and challenging to earn. In this section we examine how well the evaluation districts met this TIF grant guidance.

At least half of districts met the TIF grant guidance for awarding differentiated pay-for-performance bonuses for teachers, but not the guidance for awarding bonuses that were substantial or challenging to earn. On average across evaluation districts, the maximum bonus (\$7,442 in Year 1 and \$6,894 in Year 2) was more than three times the average bonus (\$1,810 in Year 1 and \$1,760 in Year 2), consistent with the example of a differentiated bonus provided in the TIF grant notice (Figure ES.2). However, the average bonus was about 4 percent of the average teacher’s salary—less than the 5 percent guidance for substantial bonuses specified in the TIF grant notice. Fewer than one-third of the districts met the guidance for bonuses that were challenging to earn. Across districts, on average, more than 60 percent of teachers in treatment schools received a bonus in each year.

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



Source: District administrative data (N = 2,189 teachers in Year 1; N = 2,207 teachers in Year 2; N = 65 principals in Year 1; and N = 68 principals in Year 2 in treatment schools).

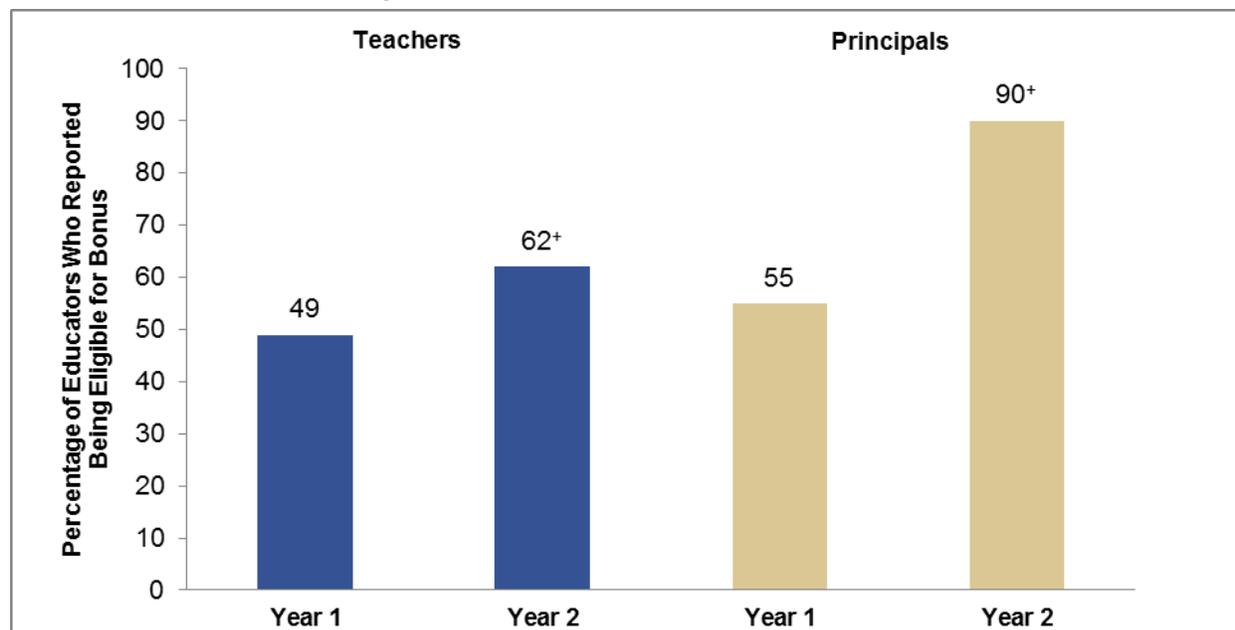
Figure reads: In Year 2, on average across the evaluation districts, the minimum pay-for-performance bonus for teachers was \$0, the average pay-for-performance bonus for teachers was \$1,760, and the maximum pay-for-performance bonus for teachers was \$6,894.

Teachers’ and Principals’ Understanding of Key Components

In addition to determining how to implement the required components of TIF, districts had to effectively communicate information about those components to educators. Educators’ understanding of the components determines how the program can influence educators’ behaviors and, ultimately, student achievement.

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. Teachers had a better understanding of how their performance was evaluated in Year 2 than in Year 1. For example, about 85 percent of teachers reported being evaluated on at least two classroom observations in Year 2, compared to about 75 percent of teachers in Year 1. In treatment schools, teachers’ and principals’ understanding of their eligibility for bonuses also improved (by 13 and 35 percentage points, respectively; Figure ES.3). However, many teachers in treatment schools (38 percent in Year 2) still did not understand that they were eligible for a bonus. They 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 (Figure ES.4). Principals also continued to underestimate the potential amount of performance bonuses they could receive, but their expectations were better aligned with actual bonus amounts than were teachers’ expectations. In Year 2, principals in treatment schools, on average, reported that the maximum pay-for-performance bonus they could receive was 87 percent of the actual maximum bonus awarded to principals.

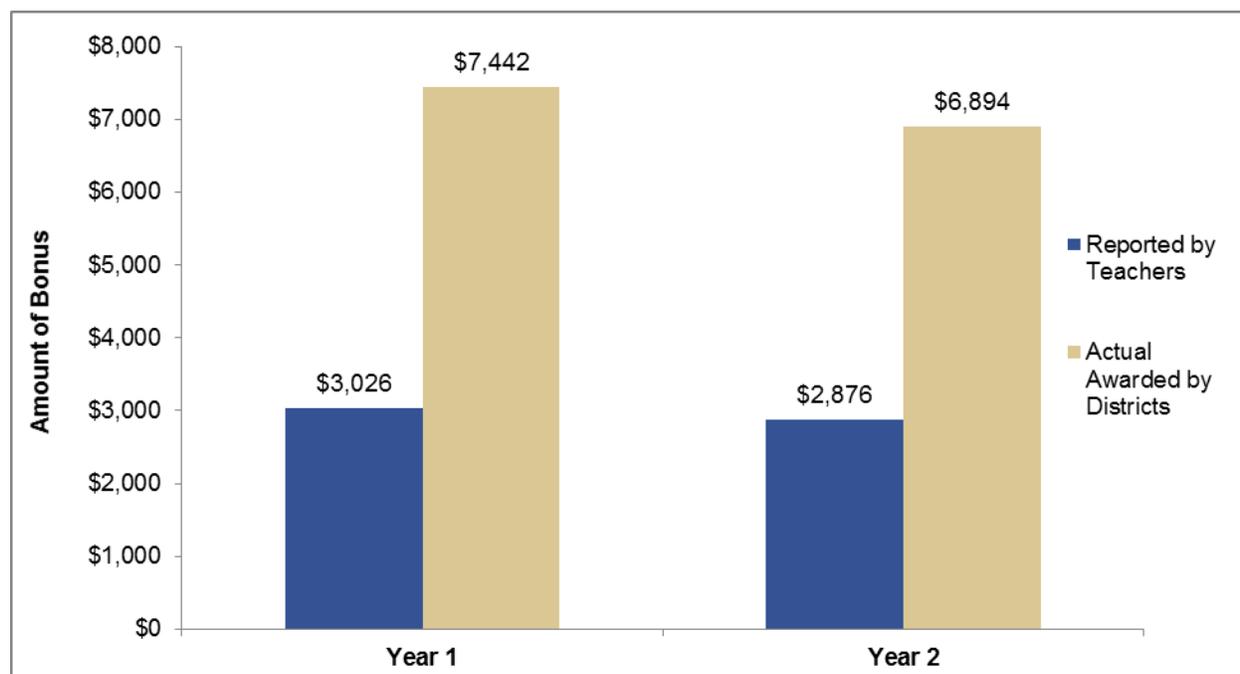
Figure ES.3. Teachers and Principals in Treatment Schools Who Reported Being Eligible for Pay-for-Performance Bonuses (Percentages)



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).

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

*Difference between 2011–2012 and 2012–2013 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 in Years 1 and 2

Source: Teacher surveys (2012 and 2013) and educator administrative data (N = 223 teachers in Year 1; N = 232 teachers in Year 2; N = 10 districts).

Figure reads: In Year 2, teachers in treatment schools reported, on average, that the maximum pay-for-performance bonus they could earn was \$2,876. On average across districts, the actual maximum bonus districts awarded to teachers in Year 2 was \$6,894.

Impacts of Pay-for-Performance on Educators' Attitudes and Behaviors

The ways in which pay-for-performance programs affect educators' attitudes (such as job satisfaction) and behaviors (such as allocation of time) can shape how pay-for-performance affects student outcomes. For example, pay-for-performance could motivate educators to improve their effectiveness if it makes them more satisfied with pay opportunities and the feedback they receive on performance evaluations. However, if the presence of pay-for-performance discourages useful collaboration, lowers morale, or makes a school less appealing to effective educators, it could have a negative effect on the work environment and, ultimately, on student achievement.

Most teachers and principals reported being satisfied with their professional opportunities, how they were evaluated, and their school environment. For example, in Year 2, at least 80 percent of teachers reported being satisfied with their opportunities to enhance their skills, their quality of interaction with colleagues, and colleagues' efforts. The percentage of principals satisfied with aspects of their professional opportunities, evaluation system, and school environment ranged from 61 to 90 percent.

Educators in treatment schools tended to be less satisfied than educators in control schools, with one exception; teachers in treatment schools were more satisfied with their opportunities to earn extra pay. For example, in Year 2, a lower percentage of teachers in treatment schools than control schools were satisfied with the use of student achievement scores to assess their performance (60 versus 69 percent) and with the feedback received on their performance (75 versus

80 percent). In Year 2, principals in treatment schools were less satisfied than principals in control schools with the use of observations to assess their skills (61 versus 85 percent) and the use of student achievement scores to assess performance (66 versus 82 percent.) The one exception to the pattern of lower satisfaction in treatment schools was that more treatment teachers were satisfied with their opportunities to earn extra pay (62 versus 54 percent) in Year 2.

Most teachers had positive attitudes toward the TIF program, but teachers in treatment schools were less likely than teachers in control schools to be positive about TIF. In both years of TIF implementation, about two-thirds of teachers were glad they were participating in TIF and at least half felt TIF was fair. However, treatment teachers in Year 2 were more likely than control teachers to report that TIF reduced their freedom to teach the way they would like (40 versus 30 percent), harmed the collaborative nature of teaching (29 versus 21 percent), and caused increased pressure to perform (65 versus 51 percent).

Impacts of Pay-for-Performance on Educator Effectiveness and Student Achievement

A central objective of the TIF grants is to improve student achievement in high-need schools by increasing educator effectiveness—in particular, by enabling schools to attract and retain more effective educators and motivating educators to improve their effectiveness. This study measured educator effectiveness using the performance ratings that educators received from their districts, and measured student achievement using students' reading and math scores on state assessments.¹

Pay-for-performance led to teachers and principals earning higher effectiveness ratings based on student achievement growth in their schools, but did not affect ratings based on observations of their classroom or school practices. The school achievement growth ratings of teachers and principals in treatment schools were 0.34 points higher than those of teachers and principals in control schools (based on a scale ranging from 1 to 4) in Year 1, and 0.25 points higher in Year 2. In Years 1 and 2, treatment and control teachers earned similar classroom observation ratings, and treatment and control principals earned similar ratings from observations of their school practices.

Pay-for-performance did not enable schools to retain or attract more higher-performing teachers. Teachers who stayed at treatment and control schools over the first two years of TIF implementation were similar in effectiveness, as measured by their classroom observation ratings and classroom achievement growth ratings. Teachers newly hired at treatment and control schools also earned similar ratings.

Pay-for-performance led to more higher-performing principals staying in their schools and more lower-performing principals leaving their schools. School achievement growth ratings were higher among principals who stayed at treatment schools than those who stayed at control schools over the first two years of TIF implementation. Observation ratings were lower among principals who left treatment schools than those who left control schools.

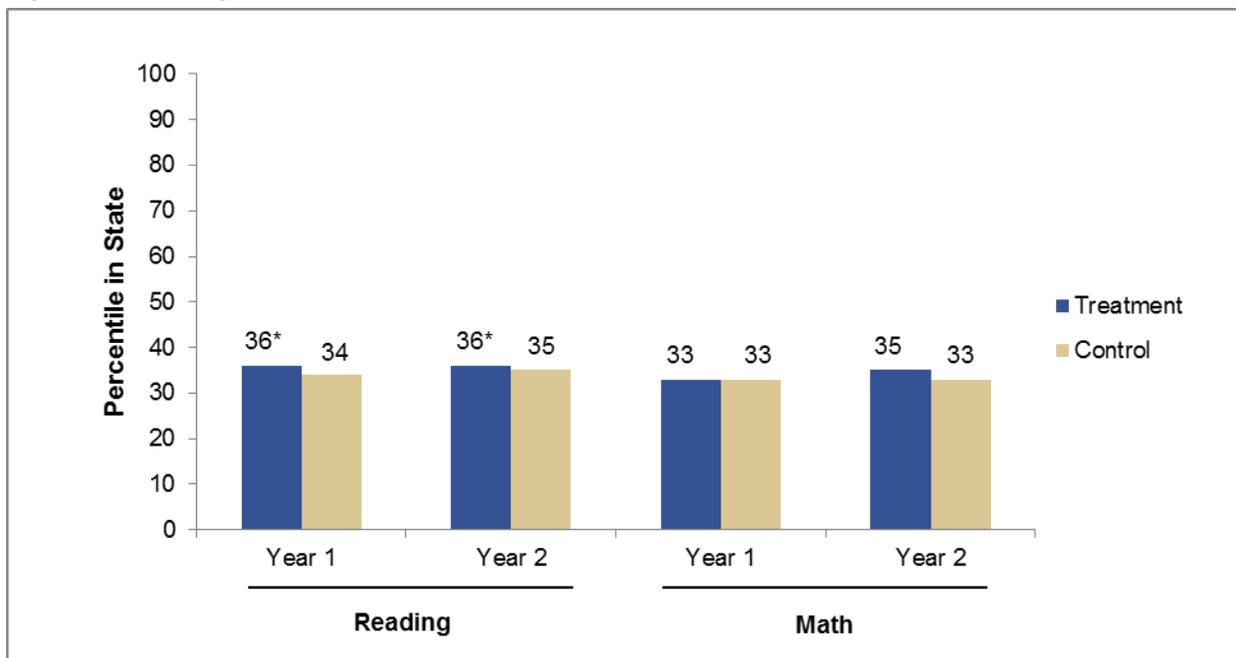
Pay-for-performance had small, positive impacts on students' reading achievement; impacts on students' math achievement were insignificant but similar in magnitude. In Years 1 and 2, the average student in a treatment school earned a reading score at approximately the 36th

¹ This study examined the impacts of pay-for-performance bonuses on the average outcomes of schools that offered those bonuses, but for simplicity we describe these findings as impacts on educators' or students' outcomes.

percentile in his or her state, whereas the average student in a control school scored at approximately the 34th or 35th percentile—a difference of 1 to 2 percentile points (Figure ES.5). This difference translated to a gain of about 3 weeks of additional learning in a typical 36-week school year. In math, differences in student achievement between treatment and control schools after Year 2 were not statistically significant, but were also positive and similar in magnitude to those in reading.

The impacts of pay-for-performance on student achievement differed among districts, but differences in impacts were not related to differences in key program characteristics measured by this study. The impacts of pay-for-performance on reading and math achievement were not related to a variety of program and implementation characteristics, including (1) the use of student achievement growth in teachers’ own classrooms to measure teacher effectiveness, (2) teachers’ understanding of their eligibility for performance bonuses, and (3) the timing of bonus notification and award.

Figure ES.5. Average Student Achievement in Treatment and Control Schools After Years 1 and 2 (Percentiles)



Source: Student administrative data (N = 40,576 students for Year 1 reading; N = 40,391 students for Year 2 reading; N = 40,852 students for Year 1 math; and N = 40,709 students for Year 2 math).

Figure reads: In Year 2, students in treatment schools earned an average reading score at the 36th percentile in their state, and students in control schools earned an average reading score at the 35th percentile.

*Difference between treatment and control schools is statistically significant at the .05 level, two-tailed test.

Concluding Thoughts

Overall, the 2010 TIF districts were able to implement most required components of a comprehensive performance-based compensation system without major, widespread challenges. However, many districts still did not put into place all the required components by the end of the second year of implementation. In addition, TIF districts were expected to sustain their programs beyond the life of the grant, but, midway through their grant, many TIF districts already reported that sustaining their programs would be a major challenge.

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. However, it is unclear whether these staffing changes among principals actually contributed to the positive impacts of pay-for-performance on student achievement. The positive impacts on reading achievement materialized in the first year—before principals had the opportunity to remain at or leave their schools—and the impacts did not increase from the first to the second year. The remaining explanation for why pay-for-performance raised student achievement in the first two years of TIF implementation is that it caused educators to improve their performance.

Many factors could have contributed to the size of the impact of pay-for-performance bonuses on student achievement. For example, the theory behind awarding educators performance bonuses to improve student achievement assumes that the prospect of earning a bonus can motivate educators to change their practices. Although pay-for-performance made teachers more satisfied with their opportunities to earn extra pay, it made teachers less satisfied with factors associated with how they were evaluated, their school environment, and their TIF program. These effects would have had offsetting impacts in shaping educators' motivation to work more effectively or to work in schools that offer performance bonuses. Furthermore, implementation findings indicated that many educators were unaware of important aspects of their program. For example, educators' understanding of program components improved between the first and second years, yet many teachers in treatment schools still did not understand that they were eligible for a performance bonus or underestimated how much they could earn from these bonuses. It is also unclear whether the actual structure of the bonuses would have provided educators with an incentive to modify their classroom or school practices, given that most educators received a bonus and average bonuses were not large.

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.

THIS PAGE IS INTENTIONALLY BLANK

