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

Executive Summary

August 2016

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 Three Years: Executive Summary

August 2016

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

Elizabeth Warner
Project Officer
Institute of Education Sciences

NCEE 2016-4005
U.S. DEPARTMENT OF EDUCATION

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

U.S. Department of Education

John B. King, Jr.

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

August 2016

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.

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:

Wellington, Alison, Hanley Chiang, Kristin Hallgren, Cecilia Speroni, Mariesa Herrmann, Paul Burkander. (2016). *Evaluation of the Teacher Incentive Fund: Implementation and Impacts of Pay-for-Performance After Three Years: Executive Summary* (NCEE 2016-4005). 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, Dylan Ellis, Kai Filipczak, Chris Jones, William Leith, Serge Lukashanets, Mickey McCauley, Jeremy Page, 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. Tora Davis helped with a variety of critical tasks ranging from tracking data collection activities 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, Dylan Ellis, Kai Filipczak, and John Hotchkiss. Leah Hackleman-Good 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, do 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 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 third from the study, describes the programs and implementation experiences of all 2010 TIF grantees in the 2013–2014 school year, the third of four years of implementation for nearly all grantees. The main findings for all districts that received 2010 TIF grants include the following:

- **Overall implementation of TIF requirements among all 2010 TIF districts was very similar in the third year of implementation as in previous years.** Similar to the previous two years, half of TIF districts in the third year reported implementing all four required components for teachers. Nevertheless, most districts (88 percent) reported implementing at least 3 of the 4 required components for teachers.
- **Few TIF districts in the third year reported that key activities related to implementation of their program were a major challenge, and districts were less likely to report major challenges in the third year than in the second year.** No aspect of TIF implementation was a major challenge to more than one-fifth of TIF districts in the third year. Furthermore, fewer districts in the third year than in the second year reported major challenges with program implementation, such as providing feedback on student achievement growth measures or teacher observations, and calculating performance bonuses. Half of the districts in the third year reported that sustainability of the TIF program was a major challenge, a decline from almost two-thirds (64 percent) of the districts in the second year.

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 three 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), second (2012–2013), and third (2013–2014) years. The main findings for the ten evaluation districts include the following:

- **Pay-for-performance had small, positive impacts on students' reading and math achievement.** After three years of TIF implementation, average student achievement was 1 to 2 percentile points higher in schools that offered pay-for-performance bonuses than in schools that did not. This difference was equivalent to a gain of about four additional weeks of learning.

- **Few evaluation districts structured pay-for-performance bonuses to align well with TIF grant guidance.** The grant notice provided guidance about how to structure pay-for-performance bonuses to be substantial, differentiated, and challenging to earn. At least half of the evaluation districts each year met the guidance for awarding differentiated performance bonuses for teachers. However, in each year, no more than 20 percent of districts awarded bonuses for teachers that were substantial or challenging to earn.
- **Teachers' understanding of performance measures continued to improve between the second and third year of implementation, but many teachers still did not understand that they were eligible for a bonus or underestimated how much they could earn.** A higher percentage of teachers in the third year reported being evaluated on student achievement growth than in the second year, and a higher percentage of teachers in the second year reported being evaluated on at least two classroom observations than in the first year. In schools that offered performance bonuses, about 60 percent of teachers (62 percent in Year 2 and 57 percent in Year 3) correctly reported being eligible for a performance bonus—implying that about 40 percent were unaware they were eligible. Similar to previous years, teachers believed that the maximum bonus they could earn was no more than two-fifths the size of the actual maximum bonus that districts awarded.

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 2012.¹

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.

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 bonuses.** 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 in size, 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.

¹ The 2015 reauthorization of the Elementary and Secondary Education Act renamed TIF the Teacher and School Leader Incentive Grants program. This program will provide grants to eligible entities to develop, implement, improve, or expand performance-based compensation systems or human capital management systems in schools.

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 provide professional development to help educators understand the measures being used to evaluate their performance as well as provide feedback based on their actual performance ratings to help improve their instructional practices.

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 bonus), 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 all four years of TIF implementation.

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 third 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. The second report used 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). The report also captured the views, attitudes, and behaviors of educators as they evolved over two years of implementation (question 2) and presented initial impacts of pay-for-performance on educator effectiveness and student achievement after the first and second years (questions 3 and 4). This third report also focuses on implementation of TIF and the effect of pay-for-performance (questions 1 through 4), but includes information after an additional year (2013–2014) of program implementation. It captures educators' views and attitudes that, by the end of the third year, were shaped by two years of pay-for-performance bonuses. The report also presents impacts of pay-for-performance on educator effectiveness and student achievement after three years of program implementation.

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 2013–2014 school year. This analysis was based on 144 districts that participated in TIF in 2013–2014.

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 three years of TIF implementation—2011–2012, 2012–2013, and 2013–2014—during the period covered by the report. The remaining 3 evaluation districts completed two years of TIF implementation—2012–2013 and 2013–2014. This report focuses primarily on the 10 districts for which data were available on three years of TIF implementation. Focusing on districts that completed three years of TIF implementation enabled us to examine changes in educators' perceptions and practices from the first to the third year and assess whether impacts on educator and student outcomes also evolved during that time.

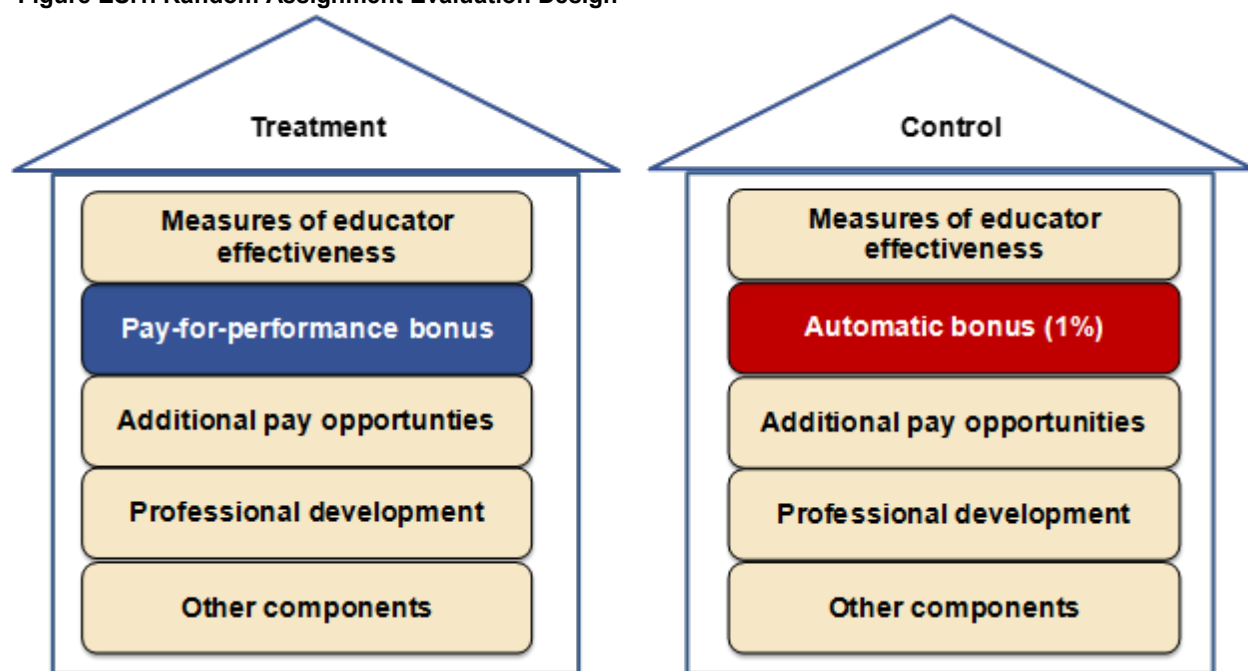
Experimental Study Design

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 three 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 at least half of its students receiving free or reduced-price lunch and 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, 2012–2013, and 2013–2014.

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 of 2012, 2013, and 2014. 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—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 144 TIF districts sought to determine whether they could put into place such a comprehensive system, and whether they faced particular challenges doing so.

Most districts implemented each of the four individual required components of TIF, but were least likely to report offering targeted professional development and evaluating principals using both student achievement growth and at least two observations of school practices. In the third year of implementation (2013–2014), nearly all the districts (over 95 percent) reported offering teachers and principals bonuses based on their performance and offering educators opportunities to earn additional pay (88 percent; Table ES.1). Fewer districts reported offering the required professional development to their teachers (70 percent), using both student achievement growth and classroom observations to measure teacher effectiveness (81 percent), and using both student achievement growth and observations of school practices to measure principal effectiveness (69 percent).

Overall implementation of TIF requirements among all 2010 TIF districts was very similar in the third year of implementation as in previous years. Similar to the previous two years, in Year 3 half of TIF districts reported implementing all four required components for teachers (Table ES.1). Nevertheless, most districts (88 percent) reported implementing at least 3 of the 4 required components for teachers. Likewise, more than half of the districts implemented all required components for principals aside from professional development, a component for which data were not available. Districts' reported implementation of each required component and of all components combined was similar across all three years.

Table ES.1. Districts' Reported Implementation of TIF Required Components for Teachers in Year 3 (Percentages)

	All 2010 TIF Districts	Evaluation Districts
Requirements		
Requirement 1: Measures of educator effectiveness ^a	81	100
Requirement 2: Pay-for-performance bonus	100	100
Requirement 3: Additional pay opportunities	88	100
Requirement 4: Professional development	70	60
Implemented all requirements	50	60
Number of Districts—Range^b	134-144	10

Source: District surveys and district interviews, 2014.

^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.

Few TIF districts reported that key activities related to implementation of their program were a major challenge, and districts were less likely to report major challenges in the third year of implementation than in the second year. No aspect of TIF implementation was a major challenge to more than one-fifth of TIF districts in Year 3. For example, about 20 percent of the districts reported that explaining student achievement growth to teachers or attributing student achievement growth to individual teachers was a major challenge. In addition, compared to Year 2, fewer districts reported major challenges in Year 3. For example, fewer districts reported major challenges with providing feedback on student achievement growth measures (19 versus 30 percent), teacher observations (14 versus 25 percent), or principal observations (4 versus 15 percent). Although concerns about sustainability stand out among the potential challenges, fewer districts in Year 3 than in Year 2 (50 versus 64 percent) reported sustainability to be a major challenge.

Detailed Summary of Findings—2010 TIF Evaluation Districts

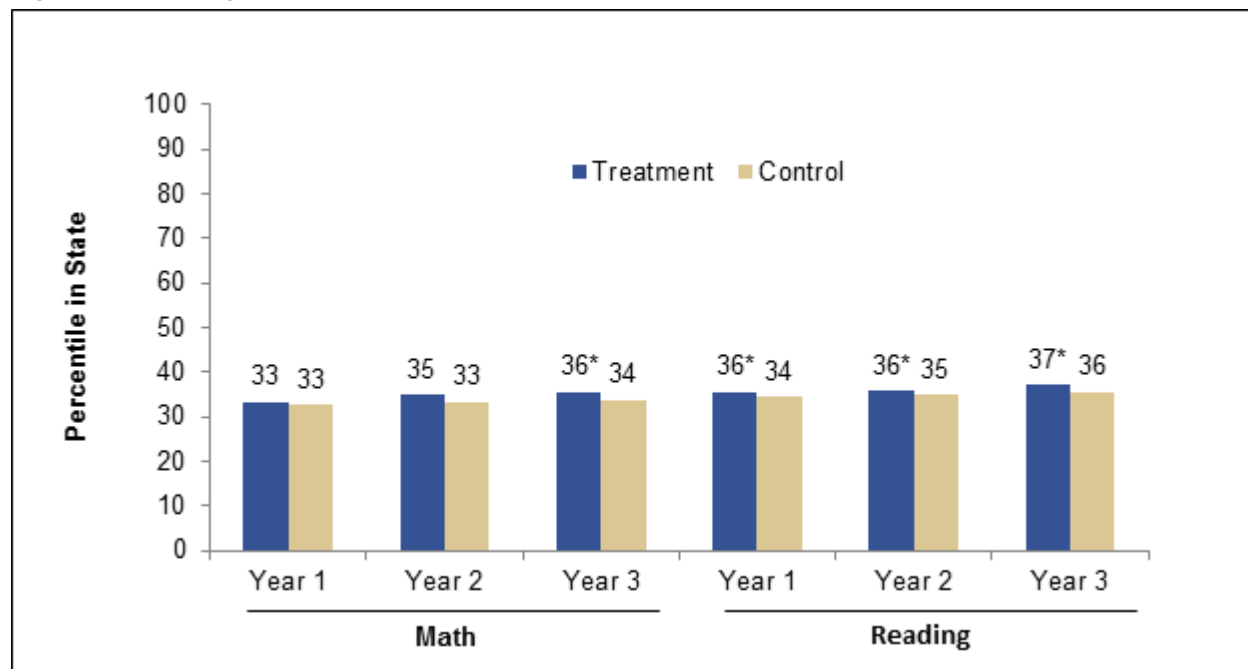
Additional information from the evaluation districts enabled the study team to examine the implementation of pay-for-performance in much greater detail, and measure the impacts of pay-for-performance on educator and student outcomes.² Ultimately, the goal of the TIF grants was to improve student achievement in high-need schools. We first present findings on the impacts of pay-for-performance on student achievement. To put those findings in context, we then present in-depth information on evaluation districts' TIF programs, teachers' and principals' understanding of and experiences with key components of their programs, and impacts of pay-for-performance on educators' satisfaction and effectiveness. Given that districts differed in the design and implementation of their programs, we also present findings on whether those differences were associated with differences in student achievement impacts.

Impacts of Pay-for-Performance on Student Achievement

Pay-for-performance had small, positive impacts on students' math and reading achievement. After three years of implementation, the average student in a control school earned a math score at approximately the 34th percentile of student achievement statewide (Figure ES.2). The average student in a treatment school earned a math score at approximately the 36th percentile—a gain of 2 percentile points. Similarly, the impact on reading achievement after Year 3 lifted the average student in these schools from the 36th to the 37th percentile. These differences translated to a gain of about 4 weeks of additional learning in a typical 36-week school year. These impacts, which represent the cumulative effect of schools' exposure to pay-for-performance for three years, were similar in size to the impacts achieved after two years of implementation.

² 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. Student achievement was measured using students' reading and math scores on state assessments.

Figure ES.2. Average Student Achievement in Treatment and Control Schools (Percentiles)



Source: Student administrative data (N = 40,847 students for Year 1 math; N = 40,708 students for Year 2 math; N = 40,037 for Year 3 math; N = 40,571 students for Year 1 reading; N = 40,390 students for Year 2 reading; N = 39,807 for Year 3 reading).

Figure reads: At the end of Year 1, students in treatment schools earned an average math score at the 33rd percentile in their state, and students in control schools also earned an average math score at the 33rd percentile.

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

TIF Implementation in Evaluation Districts

To understand the impacts of pay-for-performance on student achievement, the study team collected in-depth information about TIF implementation in the evaluation districts. Using this information, we examined the components of their programs to help assess whether they provided incentives and supports for educators to improve their effectiveness. Finally, we examined whether educators understood those components.

Program Implementation

As a first step, the study team examined the extent to which evaluation districts implemented the four required components. These analyses also examined the types of measures that districts used to evaluate educators’ effectiveness and described educators’ actual performance on those measures, focusing on whether educators received similar ratings from different measures and whether performance ratings for the same measure were similar across years.

Most evaluation districts reported implementing all required components for teachers. The only component not consistently implemented continued to be professional development. In Year 3, 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. Six of 10 evaluation districts reported providing the required professional development for teachers (Table ES.1).

When implementing the required effectiveness measures, districts could choose how to evaluate teachers based on student achievement growth. For example, districts could evaluate teachers based on 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.

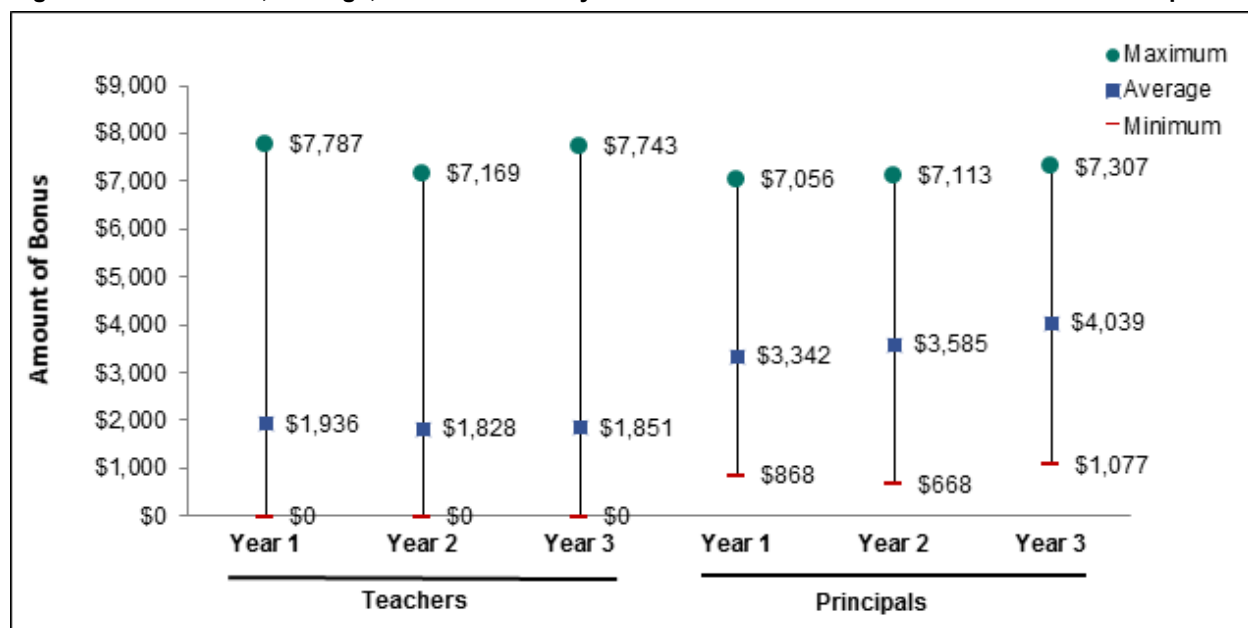
All evaluation districts reported using school achievement growth to evaluate teachers, and some also chose to evaluate teachers based on classroom achievement growth. More than half (70 percent) of evaluation districts reported evaluating teachers based on classroom achievement growth. Within these districts, more than half (about 60 percent) of teachers received classroom achievement growth ratings.

Most teachers received similar performance ratings in the third year of implementation as they did in the second year, with many teachers receiving higher ratings on classroom observations than on student achievement growth. More than half of teachers received similar ratings, based on a 1-to-4 rating scale, in Years 2 and 3. For example, 58 percent of teachers received a similar rating based on classroom observations, and 56 percent received a similar rating based on student achievement growth in their schools. However, in both years, teachers often earned higher ratings on classroom observations than on student achievement growth. For example, in Year 3, slightly more than half (53 percent) of teachers received a higher rating on classroom observations than on student achievement growth in their schools.

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 in size, differentiated, and challenging to earn.

The highest-performing teachers earned a pay-for-performance bonus about four times the average bonus. Yet, most teachers received a bonus, which, on average, was smaller than suggested by the TIF grant guidance. On average across evaluation districts, the maximum performance bonus for teachers (\$7,743 in Year 3) was about four times the average bonus (\$1,851 in Year 3), consistent with the example of a differentiated bonus provided in the TIF grant notice (Figure ES.3). However, more than 70 percent of teachers received a performance bonus, suggesting that bonuses were not challenging to earn. Moreover, the average bonus for teachers was about 4 percent of the average teacher salary—less than the 5 percent guidance for substantial bonuses specified in the TIF grant notice. For principals, bonuses were closer to the grant notice's example of a substantial bonus but were not very differentiated or challenging to earn. The average performance bonus in Year 3 (\$4,039) was slightly less than 5 percent of the average principal salary, the maximum bonus (\$7,307 in Year 3) was less than twice the average bonus, and at least three-fourths of principals received a bonus.

Figure ES.3. Minimum, Average, and Maximum Pay-for-Performance Bonuses for Teachers and Principals

Source: Educator administrative data (N = 2,183 teachers in Year 1, N = 2,193 teachers in Year 2, and N = 2,260 teachers in Year 3; N = 65 principals in Year 1, N = 68 principals in Year 2, and N = 65 principals in Year 3).

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 \$1,936, and the maximum pay-for-performance bonus was \$7,787.

Teachers' and Principals' Understanding of and Experiences with 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, and educators needed to know how to improve their performance. Educators' understanding of the components and how to improve their practices determines how the program can influence educators' behaviors and, ultimately, student achievement.

Most teachers understood that they were evaluated based on student achievement growth and classroom observations, and teachers' awareness of the use of these performance measures continued to improve between the second and third year of implementation. More than 75 percent of teachers in the third year reported being evaluated on student achievement growth, and over 85 percent reported being evaluated on at least two classroom observations. Furthermore, the percentage of teachers who reported being evaluated on these measures continued to increase. For example, a higher percentage of teachers in Year 3 reported being evaluated on student achievement growth (84 percent of treatment teachers and 78 percent of control teachers) than in Year 2 (78 percent of treatment teachers and 72 percent of control teachers), and a higher percentage of teachers in the second year reported being evaluated on at least two classroom observations (87 percent of treatment teachers and 83 percent of control teachers) than in the first year (74 percent of treatment teachers and 76 percent of control teachers).

Many teachers and some principals in schools that offered pay-for-performance bonuses still did not understand that they were eligible for a bonus or underestimated how much they could earn from performance bonuses. By the third year of TIF implementation, about 40 percent

of treatment teachers were still unaware that they could potentially earn a performance bonus (57 percent of treatment teachers reported being eligible for a bonus in Year 3, implying that 43 percent of treatment teachers did not report being eligible for one; Figure ES.4). Although understanding of eligibility was better among principals than teachers, about 20 percent of principals in Year 3 still did not know they were eligible to earn a bonus based on their performance; in fact, fewer principals were aware of their eligibility in the third year of implementation than in the second year (Figure ES.4). Similar to previous years, teachers in treatment schools believed that the maximum bonus they could earn was no more than two-fifths the size of the actual maximum bonus that districts awarded (Figure ES.5).

Most teachers reported receiving professional development on how they were evaluated and how to improve their performance, but indicated they received only a few hours of it over the school year. In Year 3, approximately two-thirds of teachers reported that they received or expected to receive professional development focused on understanding performance measures used in TIF, somewhat fewer (about 58 percent) 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 was three hours over the school year.

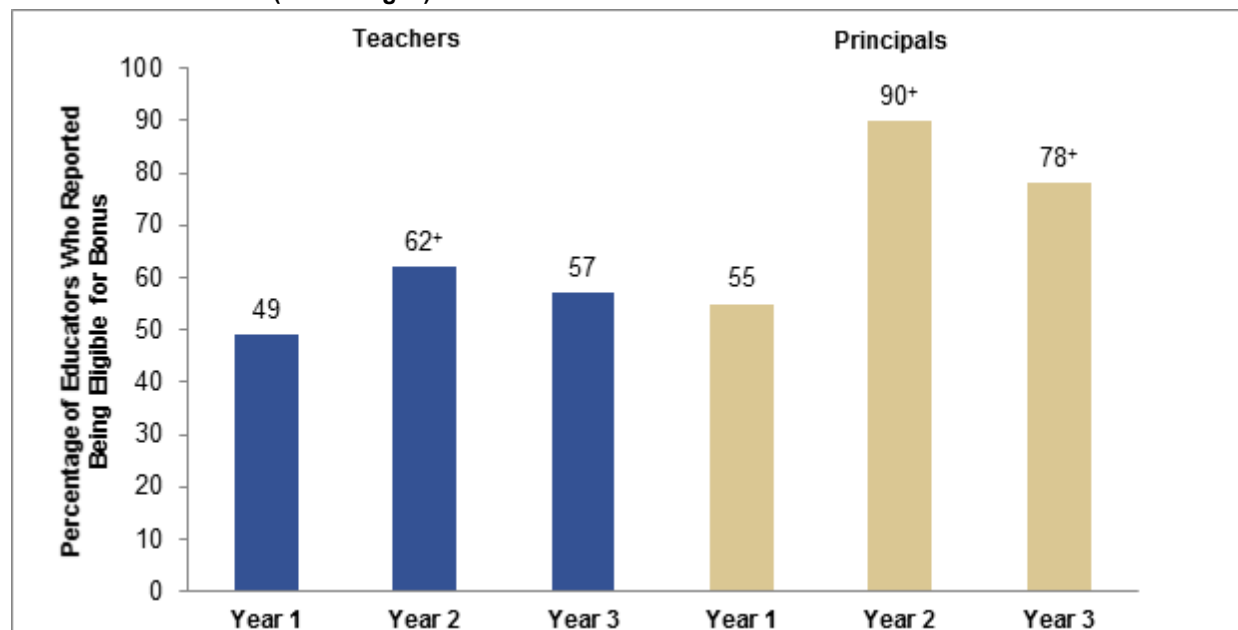
Impacts of Pay-for-Performance on Educators' Attitudes and Behaviors

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 3, about 80 percent of teachers reported being satisfied with their opportunities to enhance their skills, the feedback on their performance, the 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 54 to 96 percent in Year 3.

In contrast to prior years, teachers in treatment schools in the third year of implementation were at least as satisfied as teachers in control schools with their professional opportunities, how they were evaluated, and their school environment. 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, teachers in treatment schools reported being less satisfied than control teachers with recognition of their accomplishments and factors associated with how they were evaluated. Treatment teachers in Year 2 only reported being more satisfied than control teachers with their opportunity to earn extra pay. But in Year 3, treatment teachers reported being more satisfied than control teachers with school morale (62 versus 53 percent), the quality of their interaction with colleagues (83 versus 79 percent), and their opportunities to earn extra pay (61 versus 50 percent). They had similar levels of satisfaction as control teachers did with other aspects of their jobs.

Figure ES.4. Teachers and Principals in Treatment Schools Who Reported Being Eligible for Pay-for-Performance Bonuses (Percentages)

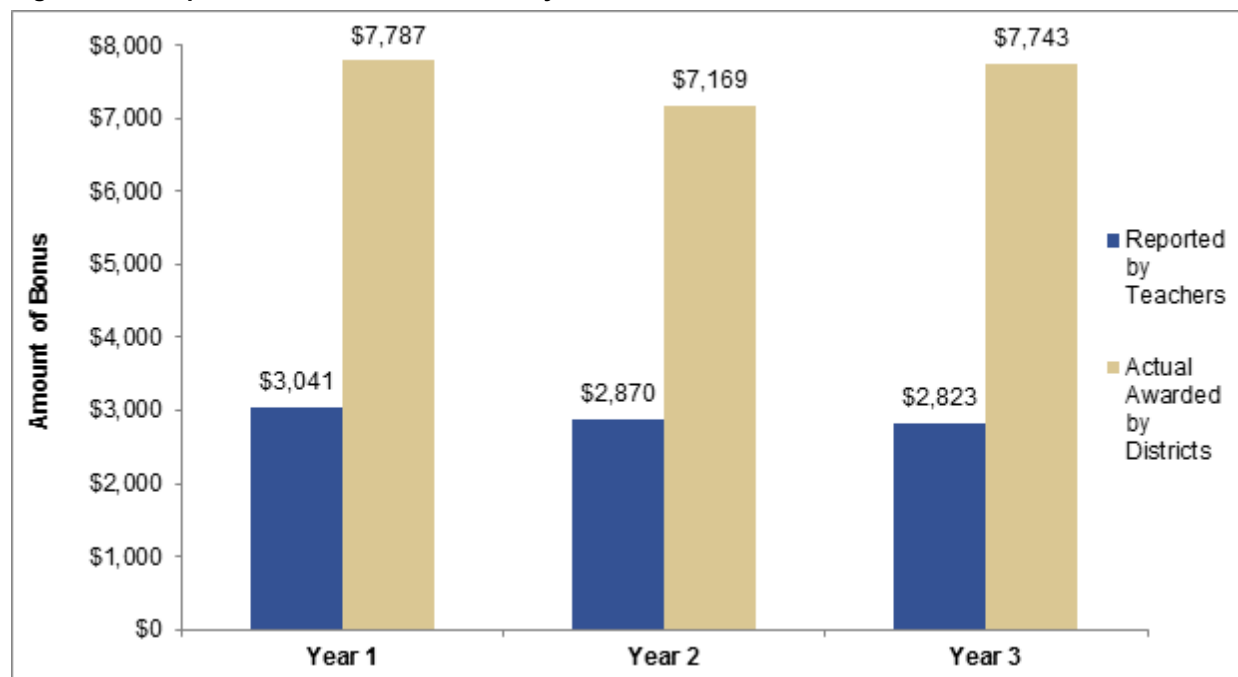


Source: Teacher and principal survey, 2012, 2013, and 2014 (N = 377 teachers in Year 1; N = 444 teachers in Year 2; N = 424 teachers in Year 3; N = 64 principals in Year 1; N = 63 principals in Year 2; and N = 58 principals in Year 3).

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 within treatment status is statistically significant at the .05 level, two-tailed test.

Figure ES.5. Reported and Actual Maximum Pay-for-Performance Bonuses for Teachers in Treatment Schools



Source: Teacher survey (2012, 2013, and 2014) and educator administrative data (N = 223 teachers in Year 1; N = 232 teachers in Year 2; N = 232 teachers in Year 3; N = 10 districts).

Figure reads: In Year 1, on average, the maximum pay-for-performance bonus that teachers reported they could earn was \$3,041, and the actual maximum pay-for-performance bonus that evaluation districts awarded to teachers was \$7,787.

Most teachers had positive attitudes toward their TIF program, and by the third year of implementation, teachers in treatment schools felt at least as positively toward TIF as teachers in control schools did. In Year 3, as in prior years, most teachers were glad to be participating in TIF. However, in contrast to Year 2, treatment teachers in Year 3 no longer felt less favorably than control teachers about the effect of TIF on teacher collaboration, their freedom to teach the way they like, and the use of student test scores to measure student learning. And for the first time, treatment teachers were more likely than control teachers to report that their job satisfaction increased due to the TIF program (39 versus 33 percent in Year 3). However, pay-for-performance continued to cause a higher percentage of treatment teachers than control teachers to feel increased pressure to perform (by 14 and 10 percentage points in Years 2 and 3, respectively).

Impacts of Pay-for-Performance on Educator Effectiveness

The ways in which pay-for-performance programs are implemented and their effects on educators' attitudes could lead to changes in educator effectiveness. In fact, 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.

Pay-for-performance had a positive impact on teachers' and principals' performance ratings based on student achievement growth in the first year of implementation, but by the third year educators in treatment and control schools received similar ratings. On measures of school achievement growth, educators in treatment schools earned ratings in Year 1 that were 0.34 points higher on a 1-to-4 rating scale than those of educators in control schools. Likewise, among teachers who were evaluated on classroom achievement growth, those in treatment schools earned ratings that were 0.18 points higher than those of teachers in control schools in Year 1. However, the impacts of pay-for-performance on both school and classroom achievement growth ratings diminished over the three years of TIF implementation. For example, by Year 3, educators in treatment and control schools earned similar school achievement growth ratings, and by Year 2, teachers in treatment and control schools earned similar classroom achievement growth ratings.

Pay-for-performance led to slightly, but not statistically significantly, higher classroom observation ratings for teachers in each year. Although differences between the classroom observation ratings of teachers in treatment schools and those in control schools were not statistically significant, they were positive and similar in all three years and almost significant by Year 3 (p -value = 0.07 in Year 3). In all three years, there were no statistically significant differences between observation ratings for principals in treatment and control schools.

Differences in Student Achievement Impacts Across Districts

The study's main findings on the impact of pay-for-performance on student achievement represent an average impact of pay-for-performance across the 10 evaluation districts. However, these districts differed in many ways, including the design and implementation of their pay-for-performance programs. These differences raise the possibility that the impacts of pay-for-performance could have also differed among districts.

The impacts of pay-for-performance on student achievement differed across 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 varied substantially, but 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 and award bonuses, (2) the size of the average bonus, (3) the level of differentiation of bonuses, (4) the degree to which earning a bonus was challenging, (5) the timing of awarding bonuses based on the prior year, and (6) teachers' understanding of their pay-for-performance eligibility.

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. In fact, fewer districts in the third year of implementation reported major challenges to implementing their TIF program than in the second year. However, many districts still did not put into place all the required components by the end of the third year of implementation.

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 three 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 achievement in reading and math. Most of the impact emerged in the first two years, and did not significantly grow in the third year.

The theory underlying the belief that pay-for-performance bonuses can lead to large impacts on student achievement depends on many factors. First, educators must understand their eligibility for a performance bonus. Yet, near the end of the third year of implementation, many educators continued to misreport their eligibility, and their understanding was no better than it was in the previous year.

Second, pay-for-performance needs to provide educators with the motivation to improve and cause effective educators to want to work in schools offering pay-for-performance bonuses. However, bonuses continued to be small on average and generally not challenging to earn, which may have dampened the motivation for teachers to improve. Furthermore, teachers still underestimated how much they could earn from the bonuses, so they may not have perceived a compelling monetary incentive to become a high performer. On the other hand, in contrast to previous years, by the third year teachers who were eligible for pay-for-performance were at least as satisfied with their jobs as those who were not eligible. However, this improvement in satisfaction was not accompanied by a larger impact on student achievement in Year 3. The improvement in satisfaction may not have been large enough to trigger changes in educator effectiveness, or it may take time for more favorable attitudes to translate into better classroom and school practices.

Third, educators need to know how to change their practices in ways that improve student achievement. We found that pay-for-performance did have small (although insignificant), positive impacts on teachers' classroom observation ratings each year. This suggests that teachers may have changed their practices slightly in response to pay-for-performance. Yet, teachers reported receiving few hours of professional development aimed at helping them improve their practices based on their performance. From this evidence, it is unclear whether teachers could really identify the changes to their practices that would most effectively improve their performance and raise student achievement.

Although the overall impact of pay-for-performance on student achievement was small, impacts were larger in some districts than in others. This raises the question of whether particular ways of designing or implementing their TIF programs could lead to larger impacts. However, none of the characteristics we examined could help explain observed differences in student achievement impacts across districts.

Evidence from the fourth and final year of implementation may provide more clarity on whether an additional year of implementation enhances educators' understanding of and experience with this program, and how impacts of pay-for-performance may evolve.

THIS PAGE IS INTENTIONALLY BLANK

