

**An Impact Evaluation of the
Teacher Incentive Fund (TIF)**

Part A

July 22, 2014

Submitted by:
Institute of Education Sciences
IES/NCEE
U.S. Department of Education
555 New Jersey Avenue, NW
Washington, DC 20208
Project Officer: Elizabeth Warner

CONTENTS

PART A. JUSTIFICATION.....	1
1. Circumstances Necessitating the Collection of Information.....	1
2. Purposes and Uses of Data.....	7
3. Use of Technology to Reduce Burden.....	9
4. Efforts to Avoid Duplication of Effort.....	10
5. Methods to Minimize Burden on Small Entities.....	10
6. Consequences of Not Collecting Data.....	10
7. Special Circumstances.....	11
8. <i>Federal Register</i> Announcement and Consultation.....	11
9. Payments or Gifts.....	12
10. Assurances of Confidentiality.....	12
11. Justification for Sensitive Questions.....	14
12. Estimates of Hours Burden.....	14
13. Estimates of Cost Burden to Respondents.....	15
14. Annualized Costs to the Federal Government.....	15
15. Reasons for Program Changes or Adjustments.....	16
16. Plans for Tabulation and Publication of Results.....	16
17. Approval Not to Display the OMB Expiration Date.....	19
18. Explanation of Exceptions.....	19
REFERENCES.....	20
APPENDIX A: DISTRICT SURVEY	
1. District Letter	
2. District Questionnaire	
APPENDIX B: PRINCIPAL SURVEY	
1. Principal Letter	
2. Principal Questionnaire	
APPENDIX C: TEACHER SURVEY	
1. Teacher Letter	
2. Teacher Questionnaire	
APPENDIX D: PRINCIPAL AND TEACHER ADMINISTRATIVE DATA REQUEST LETTER	
APPENDIX E: STUDENT RECORDS DATA COLLECTION	
1. Cover letter	
2. Instructions for Providing Student Records	
APPENDIX F: DISTRICT INTERVIEW PROTOCOL	
APPENDIX G: CONFIDENTIALITY PLEDGE	

TABLES

Table 1. Data Collection Needs.....	4
Table 2. Schedule of Major Study Activities.....	7
Table 3. Research Questions and Data Collection Methods.....	8
Table 4. Technical Working Group Members.....	11
Table 5. Estimated Response Time for Data Collection.....	15

SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT SUBMISSION

This package requests clearance from the Office of Management and Budget (OMB) to continue data collection activities for a rigorous evaluation of the Teacher Incentive Fund (TIF). This evaluation includes TIF grantees who were awarded funds from the American Recovery and Reinvestment Act (ARRA) of 2009 and the U.S. Department of Education's (ED) fiscal year (FY) 2010 appropriation. The Institute of Education Sciences (IES), within ED, began the evaluation under contract with Mathematica Policy Research and its partners Chesapeake Research Associates and faculty and staff at the Peabody College of Education at Vanderbilt University to conduct the evaluation. A follow-up contract is scheduled to be awarded by September 2014 to complete the evaluation effort.

The main objective of the evaluation is to estimate the impact of differentiated performance-based incentive pay (DPBIP)¹ on student achievement and the mobility and retention of teachers and principals. The evaluation design is an experiment in which researchers randomly assigned schools within a district to either a treatment or control group. The treatment schools implemented educator DPBIP as part of a performance-based compensation system (PBCS). Control schools implemented the same non-differentiated components of the PBCS program and a one percent across-the-board bonus, but did not implement any type of DPBIP for the duration of the TIF grant. We will compare student achievement and other outcomes between the treatment and control schools to estimate the impact of DPBIP compared to the one percent bonus.

The Notice of Final Priorities (NFP) for the TIF 2010 grants, published in the *Federal Register* on May 21, 2010, announced two competitions for grants to be awarded in 2010—the TIF main competition and the TIF evaluation competition; applicants applied to one or the other competition. The evaluation competition resulted in 12 districts participating in the random assignment evaluation.

This submission requests clearance to continue data collection for a fourth year (during the fifth and final year of the TIF cohort 3 grant) using instruments (with minor additions) previously cleared under OMB # 1850-0876. This package requests clearance to collect data that will support completion of the full-scale study during the fifth and final year of the TIF 2010 grants. The minor additions includes seven new questions in section D of the district survey, four new questions in the principal survey (question D13, and questions 1 through 3 in section F), three new questions in the

¹ For this document, DPBIP refers to the differentiated incentive pay portion of a grantee's PBCS. DPBIP programs provide bonuses for highly effective teachers and principals, where effectiveness is based on student achievement growth, observations, and any other criteria included in the district's PBCS.

teacher survey (question E11 and questions 1 and 2 in section F), and five new questions in the district interview (question 3, 21, and questions 36-38 in section G).

Part A. Justification

1. Circumstances Necessitating the Collection of Information

a. Statement of Need for a Rigorous Evaluation of TIF

The specific legislation necessitating this data collection is the ARRA, Division A, Title VIII, Pub. L. 111-5 and Departments of Labor, Health and Human Services, and Education, and Related Agencies Appropriations Act, 2010, Division D, Title III, Pub. L. 111-117. The ARRA requires that ED, to the extent possible, conduct a rigorous national evaluation to assess the impact of PBCS, supported by ARRA funds, on student achievement and educator recruitment and retention in high-need schools and hard-to-staff subjects. This evaluation would meet this requirement.

Local educational agencies (LEAs) use TIF grants to implement performance-based teacher and principal compensation systems in high-need schools. ARRA requires that the funding be used to promote effective school reform in several priority areas. These priorities include increasing teacher effectiveness, achieving equity in the distribution of high-quality teachers, and turning around the lowest performing schools. TIF requirements address these priorities.

Teacher quality is a critical input to student learning, but little is known about how to develop a strong teacher workforce (Rivkin et al. 2005; Rockoff 2004). Researchers have examined strategies to identify, attract, retain, and develop good teachers, including alternative preparation (Decker et al. 2004; Constantine et al. 2009), certification (Tuttle et al. 2009), and in-service training and professional development (Glazerman et al. 2006, Garet et al. 2008, Yoon et al. 2007). However, little is known about incentive compensation programs that tie teacher pay to student performance. Do these programs boost student achievement by attracting and retaining effective teachers and motivating all teachers to improve performance? Which types—for example, school- or individual-based programs or mixed programs (a combination of the two)—are most effective? And what challenges do districts face in implementing these programs?

To assess the overall effectiveness of TIF projects and the effectiveness of particular program models and features, ED has contracted for an evaluation of DPBIP that will be implemented by the 2010 grant recipients. This evaluation will provide important evidence on how changes to the traditional compensation systems for educators may be able to (1) improve student performance in high-need schools and/or (2) bring about desirable changes, such as the presence of more highly effective educators in high-

need schools. Results of this evaluation will provide educators, policymakers, and researchers with critical information on educator compensation reform, the effect of performance-based educator compensation on student achievement, and other aspects of PBCs associated with student achievement.

b. Research Questions

The study's research questions are:

- What is the effect on student achievement of a performance based bonus compared to an across-the-board 1% annual bonus?
- Are there differences in the composition and effectiveness of teachers and principals between these two methods of paying teachers and principals? Are there any differential effects on recruitment and retention of teachers and principals?
- Is a particular type of performance based bonus model—for example, school- or individual-based or mixed programs—associated with greater gains in student achievement? Are other key program features correlated with student and educator outcomes?
- What are the experiences and challenges of districts when implementing these programs?

c. Study Design

To answer the first research question, this study uses an experimental design—study schools within a district were randomly assigned to either a treatment or control group. Random assignment is considered the “gold standard” for social policy evaluations. More than any other approach, it minimizes the chance that any observed differences in outcomes between the study groups are due to unmeasured, pre-existing differences between members of these groups. In the random assignment design, the simple difference between outcomes in treatment and control schools within each district is an unbiased estimate of the impact of the district's DPBIP component.

Both treatment and control schools implemented the same non-DPBIP components of their program. However, only treatment schools include a DPBIP component, while control schools provide an across-the-board one percent educator bonus. Control schools are not permitted to implement a DPBIP component for the duration of the TIF grant.

Treatment schools must implement both teacher and principal DPBIP components that measure effectiveness using gains in student academic achievement and classroom evaluations conducted multiple times during each school year. Teacher incentive models may be individual-based, group-based, or mixed models.

Since we will not randomly assign schools to specific program features (program features differ among grantees), the study will use nonexperimental analyses to address the other research questions. To the extent possible, the study will examine the correlation between different types of DPBIP models and student and educator outcomes. The ability to separately analyze different DPBIP models will depend on the number and type of model(s) implemented by the grantees. The study will also examine the association of other key program features, such as how heavily the DPBIP model weights growth in student achievement with student achievement and educator outcomes.

The study includes approximately 175 schools and their students. It is designed to detect student achievement gains of 0.11 of a standard deviation. Although this may be a larger effect than can be obtained in the first year or two of the program, if DPBIP is effective in retaining and attracting effective teachers as well as improving performance among all teachers, improvement in student achievement should increase over time as educators observe bonuses received by colleagues. In addition, relatively small gains could be realized each year, contributing to larger effects after three or four years of implementation.

As part of the continued evaluation, and to address the research questions, the contractor will:

- Collect principal and teacher contact information for the study team to contact respondents who may change schools during the course of the study.
- Collect student records data to estimate the impact of DPBIP on student achievement.
- Collect administrative data on principals and teachers to track their mobility and recruitment.
- Use principal and teacher surveys to describe their understanding of and experiences with DPBIP, supplement district mobility data, and obtain background information.
- Use district surveys and interviews to describe experiences and challenges of districts when implementing the incentive programs.

d. Data Collection Needs

This study includes several data collection efforts, described below and summarized in Table 1 below. Data will be collected from the districts and schools participating in the evaluation.

District survey. We will administer a survey in winter 2015 to all 2010 TIF main and evaluation districts (Appendix A). The survey will focus on districts' experiences over the longer period and their plans for sustaining

the incentive policies. The survey seeks to contrast how the districts' programs were planned, implemented, and sustained. This survey has been informed by data collection efforts collected under the prior OMB approved data collection (OMB # 1850-0876).

We will mail the 30-minute hard copy questionnaire to each district representative. The mailing will contain a cover letter and district questionnaire. The letter, which will be on ED stationary and signed by an ED official, will describe the study and its objectives and the need for districts' participation, address issues of confidentiality, and provide a senior study member's contact information for questions or concerns. Districts will be asked to complete a hard copy questionnaire and mail it to the contractor selected for ED-IES-14-R-0019 in a postage-paid envelope.

Table 1. Data Collection Needs

Instrument	Data Need	Respondent	Mode	Schedule
District questionnaire	Specific program features, changes made to program, and how district obtained buy-in	District staff	Hard copy, phone follow-up	Winter 2015
Principal questionnaire	Background characteristics, mobility, and knowledge and perceptions of incentives	Principals	Web with email, hard copy and phone follow-up	Spring 2015
Teacher questionnaire	Background characteristics, mobility, and knowledge and perceptions of incentives	Teachers	Web with email, hard copy and phone follow-up	Spring 2015
Principal and teacher administrative data letter	Educator retention, school assignment, background characteristics, standardized test scores	District staff	Electronic or hard copy	Summer/fall 2015
Student administrative records letter	Reading and math standardized test score data for current and prior school year Demographic and socioeconomic characteristics	District staff	Electronic or hard copy	Summer/fall 2015
District interview protocol	Detailed information on program, implementation experiences, and other school improvement efforts	District staff	Telephone semi-structured interviews	Spring 2015

Principal survey. A 30-minute web-based survey will be administered to all principals in spring 2015 (Appendix B). The principal survey will ask about their background characteristics, mobility, the school's hiring practices, and knowledge and perceptions of incentives.

Teacher survey. Administered to a sample of teachers, the teacher survey (Appendix C) will be similar to the principal survey regarding mode of administration and length of questionnaire. As with principals, we will administer surveys to the same teachers as in prior data collection efforts even if they have left the school, as well as new teachers in study schools. The survey will collect information on teachers' educational and professional

background, professional development experiences, teaching and leadership responsibilities, satisfaction with various aspects of their schools, salary and other sources of compensation, and understanding of their school's PBCS.

For both principal and teacher surveys, we will first contact the sample members by email or cover letter (if email is not available or invalid). The initial correspondence will include a description of the study and survey, a link to the website address and instructions on accessing the survey, and a unique username and password. The email will explain the importance of participation, address confidentiality, and provide a toll-free telephone number and email address for questions or concerns. Nonrespondents, whom we will contact by email, telephone, or a remailing, will have the additional option of providing answers either over the telephone or by completing a hard copy version of the questionnaire.

Principal and teacher administrative data. In fall 2015, we will collect data from districts on the hiring, movement between schools, and attrition of principals and teachers participating in the study. We will also attempt to obtain information about the start and end dates of school assignments for these staff, as well as any available background characteristics such as age, sex, race/ethnicity, certifications, degrees, years of teaching experience, and scores on licensure or certification tests. In addition, we will collect several indicators of teacher and principal effectiveness and data on the actual payouts received by staff in recognition of their accomplishments. We will collect these data by the following means:

- Annual listings of principals and teachers (with personnel ID code, school, and grade if applicable) who are eligible for performance pay and the maximum amounts for which they are eligible.
- Annual listings of principals and teachers (with personnel ID code, school, and grade if applicable) who actually receive performance pay and the amounts that they receive.
- Annual data on performance measures received by principals and teachers in treatment and control schools (with personnel ID code, school, and grade if applicable). To the extent possible, performance measures should be separated into those based on observations of classroom or school practices, student achievement and growth, and other performance criteria.

Although we prefer to receive the data in an electronic format, we will use data in whatever form is most convenient for each district. We will send letters to the districts, specifying the specific data elements requested (Appendix D).

Student records data. We will request standardized math and reading test scores for all students in study schools in Summer/Fall 2015. We will also request scores from the year prior to the current study year if those scores

have not been previously obtained. In addition to test scores, we will request that the district data on student characteristics such as sex, race/ethnicity, date of birth, grade, whether they are repeating a grade, eligibility for free- or reduced-price lunch, English language learner status, and mobility within the district. Where possible, we will also request student achievement scores in math and reading, linked to the appropriate teacher. We will send the district a letter specifying the data requested (Appendix E).

District interviews. In spring 2015, we will conduct semi-structured telephone interviews with a district official who is familiar with the TIF evaluation grant program. The interview protocol is designed to collect detailed information on each district in a format that will allow for standardized follow-up questions depending on the response given to a specific item. The interview will address topics such as program implementation experiences, other ongoing school improvement efforts, and plans for performance-based pay. The protocol for the administration in 2015 is included in Appendix F.

e. Study Activities and Data Collection Timeline

This clearance request pertains to the administration of the district survey (Appendix A), principal survey (Appendix B), and teacher survey (Appendix C); collection of the district administrative records on principals, teachers, and students in the study (Appendices D and E); and administration of a district interview (Appendix F). The evaluation will be completed in eight years. Table 2 shows the schedule of data collection activities and the overall evaluation timeline.

Table 2. Schedule of Major Study Activities (including those previously conducted under OMB # 1850-0876)

Activity	Fall 2011	Spring 2012	Winter 2013	Spring 2013	Fall 2013	Winter 2014	Spring 2014	Fall 2014	Winter 2015	Spring 2015	Fall 2015	Spring 2016	Spring 2017
Provide technical assistance to grantees	X	X	X	X	X	X	X	X	X	X			
Collect principal and teacher contact information	X		X		X			X					
Conduct district survey	X		X			X			X				
Conduct principal survey		X		X			X			X			
Conduct teacher survey		X		X						X			
Collect principal and teacher records data from districts	X		X		X			X			X		
Collect student records data from districts			X		X			X			X		
Conduct district interviews		X		X						X			
Prepare first report				X									
Prepare second report							X						
Prepare third report										X			
Prepare fourth report												X	

NOTE: Bolded X signals data collection for this OMB request.

2. Purposes and Uses of Data

Data for the evaluation of TIF programs will be collected and analyzed by the contractor selected under ED-IES-14-R-0019. The data to be collected will be obtained from participants' contact information, district administrative

records, TIF district interviews, and surveys of teachers, principals and districts. The data will be used to address the research questions as shown in Table 3.

Table 3. Research Questions and Data Collection Methods

Research Question	Data Sources
1. What is the effect on student achievement of a performance based bonus compared to an across-the-board 1% annual bonus?	<ul style="list-style-type: none"> • District administrative records • Principal survey • Teacher survey
2. Are there differences in the composition and effectiveness of teachers and principals between these two methods of paying teachers and principals? Are there any differential effects on recruitment and retention of teachers and principals?	<ul style="list-style-type: none"> • District administrative records • Principal survey • Teacher survey • District survey • District interviews
3. Is a particular type of performance based bonus model—for example, school- or individual-based or mixed programs—associated with greater gains in student achievement? Are other key program features correlated with student and educator outcomes?	<ul style="list-style-type: none"> • District administrative records • Principal survey • Teacher survey • District survey • District interviews
4. What are the experiences and challenges of districts when implementing these programs?	<ul style="list-style-type: none"> • District survey • Principal survey • Teacher survey • District interviews

- **District survey.** We will use the data from the district survey to examine the association between impacts and key program features. Data from the survey will be used to describe districts’ experiences since implementing the TIF program and ascertain their plans for sustaining the program. Data from the district surveys will be used to answer research questions 2, 3, and 4.
- **Principal survey.** The principal survey will be used to assess hiring practices, classroom assignments, knowledge and perceptions of the TIF program in the study schools, how this may change over time, and to supplement administrative data to be obtained from district records. The principal survey can also provide important insight on their motivation for remaining, leaving, or entering a

study school. Data from the principal survey will be used to answer research questions 1, 2, 3, and 4.

- **Teacher survey.** The teacher survey will be used to assess knowledge and perceptions of the PBCS in the study schools and how this may change over time, and to supplement administrative data to be obtained from district records. The teacher survey can also provide important insight on teachers' motivation for remaining, leaving, or entering a study school. Data from the teacher survey will be used to answer research questions 1, 2, 3, and 4.
- **Principal and teacher administrative data.** These data will be used to estimate the impacts of DPBIP on educator mobility and recruitment. The data will also allow us to examine the association between educator characteristics and student and educator outcomes, and to describe the educator sample. These data will be used to answer research questions 1, 2 and 3.
- **Student records data.** We will use existing state or district test score data to estimate the impact of DPBIP on student achievement, the key outcome of interest. Information on students' demographic and socioeconomic characteristics and their achievement test scores prior to the study school year will be used to describe the students in the study and to develop more precise impact estimates. To the extent possible, we will use student-teacher linked data to estimate teachers' value-added score to better understand mobility of high- and low-performing educators. Data obtained from student records will be used to address research questions 1, 2 and 3.
- **District interview.** The semi-structured district interviews will allow us to collect more in-depth information than that collected from the survey, and to follow up for clarification if necessary. We will use this detailed information to more thoroughly understand each program's context, implementation strategy, and challenges. Data from the district interviews will be used to answer research questions 2, 3, and 4.

The overall purpose of this evaluation is to estimate the impacts of DPBIP on student achievement and educator mobility and recruitment in high-need schools. The findings from this study will provide important evidence for school districts and policymakers on the impacts of DPBIP on students, teachers, and school principals. If possible, this evaluation may provide policymakers and school districts with valuable information on the relative effectiveness of individual-based versus group-based compensation systems. The study will also provide important insight into the impacts of other key program aspects of DPBIP models, as well as how districts may overcome common implementation challenges. Study findings will be presented in four

annual reports, beginning fall 2014. In addition, the data collected by the evaluation will be available as restricted-use data files that will serve as a valuable resource for other researchers.

3. Use of Technology to Reduce Burden

The data collection plan is designed to obtain reliable information in an efficient way that minimizes respondent burden. We will set up a toll-free telephone number and email address specific to the study so that participants with questions can easily contact the research team. As much information as possible will be gathered from existing data sources, such as TIF grant application packets submitted by awardees and electronic files provided by districts. If it is too burdensome or not possible for a district to provide data in electronic format, we will provide clear instructions on how to submit copies of the relevant information in hard copy form, to be coded by the study team. Some data, however, can only be obtained directly from principals, teachers, and districts.

A web-based survey will be the primary mode of data collection for teachers and principals in the study. Respondents will also have the option of completing a self-administered hard copy questionnaire or providing answers to a trained interviewer over the telephone. The web-based survey will enable respondents to complete the survey at a location and time of their choice, and its automatic editing system will reduce the number of response errors.

For participants who do not return contact forms, or those whose email addresses are invalid, we will search school or district websites to obtain email addresses. Using email to follow up with nonrespondents will also offer an additional convenient option for respondents. Email reminders will include a link to the survey website and a username-password combination, as well as an attached PDF of the survey if respondents choose to complete a hard copy version.

A district representative familiar with the TIF program will complete questionnaires in hard copy form. For nonresponse follow-up, we will also offer respondents the opportunity to complete the survey over the telephone with a trained telephone interviewer. The study team considered other modes of administering the district survey, such as computer-assisted telephone interview (CATI) or a web-based survey. However, because of the relatively small sample size, the predicted cost of developing these methods outweighed the expected benefits.

We will conduct the district interviews by telephone. This mode of data collection is appropriate for the conversational exchange necessary to obtain answers to the open-ended questions, and to allow probing for more detail than a self-administered survey can provide.

4. Efforts to Avoid Duplication of Effort

The data collection plan avoids unnecessary collection of information by building off of work being conducted under ED-04-CO-0112/0012.

5. Methods to Minimize Burden on Small Entities

The primary entities for the study are TIF school districts, schools, principals, and teachers. We will minimize burden for all respondents by requesting only the minimum data required to meet study objectives. Burden on respondents will be further minimized through the careful specification of information needs. We will also keep our data collection instruments short and focused on the data of most interest, and we will speak with relatively few respondents in person. Sample sizes and data requirements for each respondent group were determined by careful consideration of the information needed to meet the study objectives, and were reviewed by the study's technical working group (TWG).

6. Consequences of Not Collecting Data

The data collection plan described in this submission is necessary for ED to conduct a rigorous national evaluation of the TIF and to understand the effectiveness of this education reform strategy. Collecting these data will allow us to finish examining the range of performance-based compensation systems and to answer pressing policy questions about how DPBIP affects student achievement and how grant recipients design, communicate, and implement TIF programs over the full course of the grants.

The consequences of not collecting specific data are outlined below.

- Each wave of the **district survey** targets different aspects of the program: specific features of districts' PBCS, if and how these features changed over time, how districts obtained buy-in, their experiences, and plans to continue their incentive policies. Without administering the district survey, and in multiple waves, we will not be able to capture these key program features and their impact on student achievement and educator mobility.
- Without the **principal and teacher surveys**, we will not know if educators understood the incentive policies, if their choice to stay in, move to, or move from a school was motivated by the incentives. We will also be unable to examine schools' hiring practices and classroom assignments, two factors that may be influenced by the TIF program. Impacts in the second and subsequent years of the implementation of the DPBIP may be larger than those in the first year. Administering the surveys in multiple waves will allow us to examine educators' experiences and perceptions of the programs over time.

- Without **principal and teacher records data**, it will be more difficult to verify educators’ school assignment and track their mobility. Furthermore, without this data we will not be able to compare characteristics between principals and teachers in the treatment and control schools, or to examine whether staff characteristics are associated with student achievement growth or educators’ mobility decisions.
- Without **student records data**, we will have to administer assessments to students in place of using their district or state math and reading test scores to measure student achievement. Without the data on student characteristics, we will not be able to fully describe the study sample and verify the effectiveness of the random assignment.
- Without the **district interviews**, we will not be able to follow up on information obtained from the surveys to obtain a more thorough understanding of the districts’ programs and experiences, or to fully understand any other related school reform initiatives within the district that may affect the impact of DBPIP in the study schools. Multiple waves are necessary as a detailed follow-up to each district survey.

7. Special Circumstances

There are no special circumstances associated with this data collection.

8. Federal Register Announcement and Consultation

a. Federal Register Announcement

The 60-day notice to solicit public comments was published on June 18, 2014. There have been no public comments received to date. The 30-day notice will be published to solicit additional public comments.

b. Consultation Outside the Agency

In executing the evaluation design, the study team has sought input from the technical working group (TWG), which includes some of the nation’s experts in teacher compensation, evaluation methodology, and education policy. We will continue to consult with the TWG throughout the study on other issues that would benefit from their input. Table 4 lists the TWG members.

Table 4. Technical Working Group Members

Name	Title and Affiliation	Expertise
Anthony Milanowski	Assistant Scientist, University of Wisconsin	Teacher compensation
Richard Murnane	Professor of Education, Harvard Graduate School of Education	Teacher compensation and teacher quality

Jacob Vigdor	Professor of Public Policy and Economics, Duke University	Teacher compensation, teacher quality, and evaluation methodology
Dan McCaffrey	Senior Statistician, RAND Corporation	Value added and evaluation methodology
Robert Meyer	Research Professor, University of Wisconsin	Value added
Jeffrey Smith	Professor of Economics, University of Michigan	Teacher quality/methodology
James Kemple	Director of Research Alliance for NY City Schools, Research Professor, New York University	Teacher quality/methodology
David Heistad	Executive Director of Research, Evaluation and Assessment, Minneapolis Public Schools	Program evaluation, value-added in teacher compensation systems
Carla Stevens	Assistant Superintendent, Research and Accountability, Houston Independent School District	Accountability, student assessment, program evaluation, and performance pay models

c. Unresolved Issues

There are no unresolved issues.

9. Payments or Gifts

Incentives for principals and teachers. Incentives proposed here are the same as those cleared under OMB # 1850-0876. Specifically we propose incentives for the principal and teacher surveys to partially offset respondents’ time and effort in completing the surveys. We propose offering a \$20 incentive to an educator each time he or she completes a questionnaire so as to acknowledge the 30 minutes required to complete each questionnaire. This proposed amount is within the incentive guidelines outlined in the March 22, 2005 memo, “Guidelines for Incentives for NCEE Evaluation Studies,” prepared for OMB.

Incentives are also proposed because high response rates are needed to make the survey findings reliable, and we are aware that teachers and principals are the targets of numerous requests to complete surveys on a wide variety of topics from state and district offices, independent researchers, and the Department of Education. Although some districts will have solicited buy-in from teachers to participate in the evaluation, our recent experience with numerous teacher surveys supports our view that obtaining teacher buy-in does not guarantee teachers will devote the time it takes to complete a survey, and monetary incentives increase the likelihood of cooperation of school staff.

The study will not give incentives to districts for completing an interview or a survey, or for providing administrative records data.

10. Assurances of Confidentiality

The contractor will be required to conduct all data collection activities for this study in accordance with relevant regulations and requirements, which are:

- The Privacy Act of 1974, P.L. 93-579 (5 U.S.C. 552a).
- The Family Educational and Rights and Privacy Act (FERPA) (20 U.S.C. 1232g; 34 CFR Part 99).
- The Protection of Pupil Rights Amendment (PPRA) (20 U.S.C. 1232h; 34 CFR Part 98).
- The Education Sciences Institute Reform Act of 2002, Title I, Part E, Section 183

The research team will be required to protect the confidentiality of all data collected for the study and will use it for research purposes only. The contractor's project director will be required to ensure that all individually identifiable information about respondents will remain confidential. All data will be kept in secured locations and identifiers will be destroyed as soon as they are no longer required. All members of the study team having access to the data will be trained and certified on the importance of confidentiality and data security. When reporting the results, data will be presented only in aggregate form, such that individuals and institutions will not be identified. Included in all voluntary requests for data will be the following statement:

Responses to this data collection will be used only for statistical purposes. The reports prepared for this study will summarize findings across the sample and will not associate responses with a specific district or individual. We will not provide information that identifies you or your district to anyone outside the study team, except as required by law. Additionally, no one at your school or in your district will see your responses. While your participation in this study is voluntary, it is very important that you complete the questionnaire.

For those instruments where data collection is required as a condition of their evaluation grant, all grant required requests for data will include the following statement:

Responses to this data collection will be used only for statistical purposes. The reports prepared for this study will summarize findings across the sample and will not associate responses with a specific district or individual. We will not provide information that identifies you or your district to anyone outside the study team, except as required by law. Additionally, no one at your school or in your district will see your responses. Participation or cooperation with this activity is a condition of your grant (EDGAR: part 75.591, Authority: 20 U.S.C. 1221e-3 and 3474).

The following safeguards are routinely required of contractors for IES to carry out confidentiality assurances, and they will be consistently applied to this study:

- All contractor employees will be required to sign a confidentiality pledge (such as Appendix G) that emphasizes the importance of confidentiality and describes employees' obligations to maintain it.
- Personally identifiable information (PII) is maintained on separate forms and files, which are linked only by sample identification numbers.
- Access to hard copy documents is strictly limited. Documents are stored in locked files and cabinets. Discarded materials are shredded.
- Access to computer data files is protected by secure usernames and passwords, which are only available to specific users.
- Sensitive data is encrypted and stored on removable storage devices that are kept physically secure when not in use.

The contractor will be required to maintain confidentiality including personnel training regarding the meaning of confidentiality, particularly as it relates to handling requests for information, and providing assurance to respondents about the protection of their responses. The contractor will also be required to build-in safeguards concerning status monitoring and receipt control systems.

11. Justification for Sensitive Questions

Some respondents may consider their contact information to be sensitive. This information is necessary in order to limit possible sample attrition that could result from respondents changing schools or professions.

The principal and teacher surveys will ask for demographic information (ethnicity, race, year of birth) and information about respondents' educational and professional background. Data on these topics are important to help us understand if there is an association between student achievement, educator outcomes, and educator characteristics. Questions used to obtain personal background information have been asked frequently in other surveys and were pretested for this study under OMB # 1850-0876, with the pretest sample of teachers and principals reporting no concerns.

To address concerns about disclosing personal information, all cover letters and questionnaires will clearly state that all responses will be treated as confidential, that participation is voluntary, and that failure to provide some or all requested information will not affect the respondent's professional status in any way. The questions will also be worded in a sensitive, nonjudgmental manner.

Some demographic information about the students (for example, qualification for free- or reduced-price lunch or special education status) or their test scores may be sensitive. Demographic information is important to control for any differences in the characteristics of students in the classes that may have arisen by chance. Test score data is essential for this evaluation because student achievement is the primary outcome of interest. These scores will be linked to the data file by each respondent's unique, study-generated identification number. After this linking process, personal identifiers, such as a student's name, school identification number, and date of birth, will be removed.

There are no questions of a sensitive nature in the district survey or interview.

12. Estimates of Hours Burden

Table 5 provides an estimate of time burden for the data collections, broken down by instrument and respondent. These estimates are based on our experience collecting administrative data from districts, administering surveys to school principals and teachers, and conducting telephone interviews with district representatives.

Table 5. Estimated Response Time for Data Collection

Respondent/ Data Request	Number of Targeted Respondent s	Expected Response Rate (%)	Number of Respondent s	Unit Response Time (Hours)	Total Response Time (Hours/Year)	Total Burden Time (Hours)
Districts^a						
Student records data	13	100	13	8.0	104	104
Principal and teacher records data	13	100	13	8.0	104	104
Principals						
Principal surveys	175	90	158	0.5	79	79
Teachers						
Teacher surveys	2000	85	1700	0.5	850	850
Districts						
Surveys	153	80	122	0.5	61	61
Interviews	13	100	13	0.75	10	10
Total			2019			1208

Annual number of respondents and responses for the 3 years of this collection are 673. The total number of burden hours for this revised collection is 1208 hours which will generate a program change of 403 burden hours annually for this revision.

^aDepending on the grantee, administrative records data may be provided by another source, for instance the state or grantee.

The total of 1208 hours covers the data collection conducted during the final grant year of the evaluation, and includes the following efforts: up to 16 hours for each of the 13 districts to collect and assemble administrative records on students, principals, and teachers participating in the evaluation; 30 minutes for 158 principals (90 percent of the anticipated 175 principals in the sample) to complete the principal survey; 30 minutes for 1,700 teachers (85 percent of the anticipated sample of 2000 teachers in the sample) to complete the teacher survey; 30 minutes for 122 district representatives (80 percent of the 153 districts participating in the study) to complete a district survey; and 45 minutes for the 13 districts participating in the evaluation to complete a telephone interview. Annual number of respondents and responses for the 3 years of this collection are 673 (the total number of respondents, 2019, divided by 3). The total burden hours for this collection are 1208, which will generate a program change of 403 burden hours annually for this revision (the total burden hours, 1208, divided by 3).

13. Estimates of Cost Burden to Respondents

There are no direct costs for respondents.

14. Annualized Costs to the Federal Government

The estimated annual cost of the study to the federal government is \$1,714,286. The total cost of the seven-year study is \$12 million, which includes recruiting grantees, districts, and schools; designing and administering data collection instruments; processing and analyzing data; and preparing reports.

15. Reasons for Program Changes or Adjustments

There is an overall program change of 403 burden hours attributed to this revision. The program change is a result of the 403 annual burden hours from this revision of the collection being added to the burden hours of 1359 already approved for the current ongoing collection. A change request will be submitted to OMB after all of the currently approved data collection activities are completed.

16. Plans for Tabulation and Publication of Results

a. Tabulation Plans

Our tabulation plans are the same as those cleared under OMB # 1850-0876. Specifically, our tabulation plans include four sets of analyses aligned to the research questions. Random assignment of schools within a district to a treatment group that will implement DPBIP or to a control group not allowed to do so for the duration of the TIF grant is an ideal design for assessing overall effectiveness. Our primary impact analysis will exploit this experimental design to provide rigorous estimates of the impact of DPBIP on student achievement and teacher/principal mobility and recruitment. Additional nonexperimental analyses are designed to estimate the relative effectiveness of individual-based versus group-based or mixed incentive programs, explore the association of other key program features with student achievement and teacher/principal outcomes, and to learn about districts' implementation experiences and challenges.

Estimating the overall impact of DPBIP. What follows is the same as those cleared under OMB # 1850-0876. With this experimental design, the simple differences between mean outcomes in the treatment and control schools should yield unbiased estimates of the impacts of DPBIP. However, the precision of the estimates can be improved by using regression procedures to control for student, teacher, or school baseline characteristics that may explain some of the variation in outcomes not related to the treatment itself. These characteristics may include student controls, such as test scores from the year before TIF implementation; gender, race/ethnicity, free- or reduced-price lunch eligibility, special education status, and English learner status; teacher controls, such as demographic characteristics, age, experience, and educational background; and school-level averages of the student or teacher characteristics. Regression procedures also enable us to

adjust for any differences between treatment and control groups in these baseline characteristics that happen to arise due to chance or sample attrition. The regression model must be flexible enough to include the full range of programs and generate estimates of district-specific impacts, which can then be aggregated to produce an overall estimate. We will therefore estimate variations of the following model for the outcome y_{ijk} of individual (student or teacher) i in school j within district k :

$$(1) \quad y_{ijk} = \mathbf{R}_{jk} \boldsymbol{\alpha} + \sum_{k=1}^K \beta_k (T_{jk} \times G_k) + \mathbf{X}_{ijk} \boldsymbol{\delta} + \mathbf{Z}_{jk} \boldsymbol{\gamma} + u_{jk} + \varepsilon_{ijk}$$

where \mathbf{R}_{jk} is a vector of indicators for combinations of grade levels and randomization strata; $\boldsymbol{\alpha}$ is a vector of grade-by-strata fixed effects; T_{jk} is a treatment indicator; G_k is a dummy variable for district k ; β_k is the impact of DPBIP in district k ; \mathbf{X}_{ijk} is a vector of baseline individual characteristics with coefficient vector $\boldsymbol{\delta}$; \mathbf{Z}_{jk} is a vector of baseline school-level characteristics with coefficient vector $\boldsymbol{\gamma}$; u_{jk} is a random school effect; and ε_{ijk} is a random individual error term. The district-specific impacts of performance pay, β_k , are the key coefficients of interest in equation (1). We will estimate equation (1) with ordinary least squares (OLS) using Huber-White (“sandwich”) standard errors that account for school-level clustering.

Our primary interest is in the overall, average impact of DPBIP in the full study sample. To estimate the average impact of DPBIP on schools in the study, we will take a weighted average of the estimated district-specific effects, $\hat{\beta}_k$, with weights equal to the number of treatment and control schools within each district. The standard error of the average impact estimate can be calculated from the estimated variances and covariances among the district-specific impacts from equation (1).

The evaluation includes four years of analyses, and the following describes the analyses previously outlined in the previously cleared OMB submission. Impacts in the second and subsequent years of the implementation of the DPBIP may be larger than those in the first year for several reasons. First, changes in educator effort and the composition of the teaching staff at treatment schools may be more pronounced after educators observe the payments from earlier years. Also, if educators improve their performance over time, in years 2 through 5 of the grant, some students will have had multiple years of exposure to the treatment. For these reasons, equation (1) will be estimated separately for assessing impacts for each year of implementation, as well as cumulative impacts.

The impact of DPBIP on the outcomes of interest—student achievement and educator mobility and recruitment—will be estimated with a variant of equation (1). Student achievement outcomes are math and reading scores from spring 2012, 2013, 2014, and 2015 state or district assessments. Because tests will differ across states, grade levels, and subjects, we will convert raw scale scores to z-scores (raw scores minus the mean score divided by the standard deviation of scores on that test among students in that grade and state) in order to scale the outcome variable comparably across all students in the sample. Using district records, we will measure teacher retention as a dichotomous outcome for whether or not the teacher returns to work in the grantee site and/or in his or her initial school in fall of 2011 and continue to do so annually through 2015. Because the retention outcome is dichotomous, we will estimate the probit model analog of equation (1). Annual school-level teacher data from study schools in fall, 2011 through fall, 2015 (from district records) and spring 2012, 2013, 2014, and 2015 (from the principal and teacher surveys) will be analyzed as outcomes to examine impacts on the composition of the teaching staff. If available from administrative records, the quality of applicants who apply to teach in study schools for school years 2012-2013, 2013-2014, 2014-2015, and 2015-2016 will also be analyzed, including the total number of applicants, average experience level, percentage of applicants who have teaching experience, and the selectivity of the college from which they graduated. Equation (1) can be aggregated to the school level for the analysis of composition outcomes.

To better understand mobility of high- and low-performing principals and teachers, for grantees where we can obtain or calculate a measure of staff effectiveness, we will also estimate a model of transitions that includes a teacher or school measure of effectiveness, and interactions of this measure with treatment indicators in the set of independent variables. The coefficients on the effectiveness measure by treatment interactions provide an estimate of whether differences in retention between highly effective and less effective principals or teachers are more or less pronounced in treatment versus control schools. Since high- and low-performing teachers are not being randomly assigned to treatment and control schools, and estimates of their effectiveness may be endogenous if DPBIP induces greater teacher effort, these estimates are nonexperimental and will need to be interpreted with caution. Wherever possible, we will obtain or calculate value-added estimates based on student achievement to measure teacher effectiveness. In addition, if possible, we will also use districts' measures of effectiveness.

Estimating the effectiveness of key program features. What follows is the same as those cleared under OMB # 1850-0876. We will conduct exploratory analyses to assess whether particular features of DPBIP are associated with impacts on student achievement. These analyses will, in particular, examine the relative effectiveness of DPBIP models that place

different weights on individual versus group performance in the determination of incentive payouts. Other programmatic features of interest include the average and maximum size of the incentive payouts and the degree to which the payouts vary across educators.

Since we do not expect that districts will randomly assign specific components of their DPBIP to schools, we will not be able to experimentally assess the relative effectiveness of different DPBIP program features. Instead, we will examine the association between impacts and key program features in a regression framework. We will be careful to note that an observed association between impacts and programmatic features may not necessarily have a causal interpretation.

For these analyses, we will rely on findings from the implementation analysis to examine how the variation in programmatic features is related to the impact. Our basic approach is to regress the estimated district-specific impacts from equation (1) on a measure of a specific programmatic feature.

For the estimated impact $\hat{\beta}_k$ from district k , we estimate:

$$(2) \hat{\beta}_k = \pi_0 + \lambda W_k + \omega_k$$

where π_0 is an intercept, W_k is a measure of a specific programmatic feature with associated coefficient λ , and ω_k is an error term that includes random error in estimating the true impact β_k . Because impacts might be more precisely estimated in some districts than in others, we will weight grantees by the precision of the estimated impacts when estimating equation (2) to account for this source of heteroskedasticity in the error term. For each of the programmatic features described earlier, we will estimate equation (2) with the specified program feature as the only covariate, given the limited number of grantees in the sample.

Understanding the implementation experiences of TIF districts.

What follows is the same as those cleared under OMB # 1850-0876 with minor revisions. Understanding the implementation experiences and challenges of TIF grantees will provide essential information for improving the implementation of future incentive programs and is crucial for the interpretation of the impact findings. We will analyze the implementation data collected from grantee, district, and school documents; district, principal, and teacher surveys; and telephone interviews with districts to report on their incentive policies and experiences. Since the evaluation districts were purposively selected, and the impact estimates cannot necessarily be generalized beyond this sample, we will use the district surveys to construct tables on their incentive policies, comparing the evaluation districts to all recent awardees. We also will use the district surveys and information from telephone interviews to document and analyze implementation challenges. The principal and teacher surveys will provide

critical context to determine if they understood the incentive compensation policy and program in their district and school and adjusted their behavior accordingly. After the initial survey, for each subsequent wave of the principal and teacher surveys, we will construct tables to assess any changes in educators' understanding and behavior.

b. Publication Plans

We will prepare a final report presenting the results of these tabulations. This report will be the final of four reports prepared under this evaluation. The final report, with a projected release date of Spring 2017, will describe districts' implementation strategies and challenges and examine impacts through the fifth and final year of the TIF cohort 3 grants. Reports will be written in a style and format accessible to policymakers and research-savvy practitioners and will comply fully with the standards set by the National Center for Education Statistics.

17. Approval Not to Display the OMB Expiration Date

The study will display the OMB expiration date.

18. Explanation of Exceptions

No exceptions are being sought.

REFERENCES

- Anderman, C., A. Cheadle, S. Curry, P. Diehr, L. Shultz, and E. Wagner. "Selection Bias Related to Parental Consent in School-Based Survey Research." *Evaluation Review*, vol. 19, no. 6, 1995, pp. 663-674.
- Constantine, Jill, Daniel Player, Tim Silva, Kristin Hallgren, Mary Grider, and John Deke. "An Evaluation of Teachers Trained Through Different Routes to Certification." Princeton, NJ: Mathematica Policy Research, February 2009.
- Decker, Paul, Steven Glazerman, and Daniel Mayer. "The Effects of Teach For America on Students: Findings from a National Evaluation." Princeton, NJ: Mathematica Policy Research, June 9, 2004.
- Eaton, Danice K., Richard Lowry, Nancy D. Brener, Jo Anne Grunbaum, and Laura Kann. "Passive Versus Active Parental Permission in School-Based Survey Research: Does the Type of Permission Affect Prevalence Estimates of Risk Behaviors?" *Evaluation Review*, vol. 28, no. 6, 2004, pp. 564-577.
- Garet, Michael S., Stephanie Cronen, Marian Eaton, Anja Kurki, Meredith Ludwig, Wehmah Jones, Kazuaki Uekawa, Audrey Falk, Howard Bloom, Fred Doolittle, Pei Zhu, and Laura Sztejnberg. "The Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement." Washington, DC: U.S. Department of Education, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, September 2008.
- Glazerman, Steven, Paul Decker, and Daniel Mayer. "Alternative Routes to Teaching: The Impacts of Teach For America on Student Achievement and Other Outcomes." *Journal of Policy Analysis and Management*, vol. 25, no. 1, 2006, pp. 75-96.
- Hanushek, Eric A. "Efficient Estimators for Regressing Regression Coefficients." *American Statistician*, vol. 28, no. 2, 1974, pp. 66-67.
- Hanushek, Eric A., and S. Rivkin. "Generalizations About Using Value-Added Measures of Teacher Quality." *American Economics Review*, vol. 100, no. 2, 2010, pp. 267-71.
- Rivkin, S., E. Hanushek, and J. Kain. "Teachers, Schools, and Academic Achievement." *Econometrica*, vol. 73, no. 2, 2005, pp. 417-458.
- Rockoff, J. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review* (AEA Papers and Proceedings), vol. 94, no. 2, 2004, pp. 247-252.

Tuttle, Christina, Steven Glazerman, and Tara Anderson. "ABCTE Teachers in Florida and Their Effect on Student Performance." Washington, DC: Mathematica Policy Research, April 27, 2009.

Yoon, K. S., T. Duncan, S.W.Y. Lee, B. Scarloss, and K. Shapley. "Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement." Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2007.