Mathematica
Progress Together

Impact Evaluation to Inform the Teacher and School Leader Incentive Program

Part B: Collection of Information Employing Statistical Methods

April 28, 2020

CONTENTS

TABLES

PART B. SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT SUBMISSION

This package requests clearance for the remaining data collection activities to support an evaluation of the Teacher and School Leader Incentive Program (TSL). The Institute of Education Sciences (IES), U.S. Department of Education (ED) has contracted with Mathematica and its partners Public Impact, Applied Engineering Management Corporation, Decision Information Resources, and Dr. Jason Margolis to conduct this evaluation.

This study will include two evaluation components:

1. ***Descriptive study of TSL grantees' programs.*** Data collection includes interviewing all districts included in a fiscal year 2017 TSL grant to obtain information on TSL grantees' programs and experiences.

2. ***Implementation, impact, and cost-effectiveness study of designating one or more "teacher leaders" as coaches in schools.*** A random assignment study of this common TSL strategy will be conducted in non-TSL schools. Data will also be collected from TSL grantee schools on their implementation of the teacher leader role, in order to connect the impact findings from non-TSL schools to the TSL implementation experience.

This package provides a detailed discussion of both evaluation components. However, the package only requests clearance for the items associated with the second component that have yet to receive clearance. An initial package requesting clearance for early data collection activities was approved on January 16, 2020 under OMB Control No: 1850-0950.[1]

---

[1] https://www.reginfo.gov/public/do/PRAViewICR?ref_nbr=201908-1850-005

COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

B1.   Respondent universe and sampling methods

The study will examine two groups of districts and schools:

1. The universe of 2017 TSL grantee districts along with a random sample of schools within the TSL districts funding teacher leaders. The TSL grantee sample will include all 14 TSL grantees and the 25 districts covered by a 2017 TSL grant. The sample of schools will be drawn from among the 12 TSL grantees funding teacher leaders through their TSL grant, which includes 23 districts.

2. A purposive sample of 100 schools across 10 school districts for the random assignment evaluation. Within each of these districts, we will match schools into pairs prior to random assignment. The schools will be matched based on the similarity of their grade span, grades and subjects that they propose for teacher leader roles, and student characteristics. In spring 2020, we will randomly assign one school in each pair to the treatment group and one to the control group. The treatment group will receive the teacher leader intervention, including funding and training for two teacher leaders for two years (2020-2021 and 2021-2022 school years). The control group will proceed with business-as-usual.

B2.   Procedures for the collection of information

1.   Statistical methods for sample selection

Below we explain the samples for the 2017 TSL grantees and for the random assignment evaluation.

a.   2017 TSL grantees

TSL grantee districts, and principals and teachers from schools covered by TSL grants will be asked to complete interviews or surveys.

School districts. We will not be sampling 2017 TSL school districts. We will ask all 25 districts covered by a 2017 TSL grant to participate in the district interviews.

Schools. We will randomly select 100 schools from the 23 TSL districts that received and are using TSL funds to support a teacher leader role. The principal of each of these 100 schools will be asked to complete a principal survey.

Teachers. We will randomly select 2 eligible teachers from each of the 100 randomly selected TSL-supported schools. Eligible teachers will be those in the same grades as those implementing teacher leader roles for the random assignment evaluation (for example, 4th and 5th grades) and who teach math or ELA.

b. Random assignment evaluation

The random assignment evaluation will include a purposive sample of approximately 10 districts that together include about 100 schools.

Selection of school districts. We will recruit a purposive sample of districts and schools that are willing to participate and meet the study's three main eligibility criteria:

1. District includes elementary schools that are high-need

2. District uses a teacher evaluation system based in part on student learning

3. School does not have teacher leader or teacher support systems in place that is similar to the study's teacher leader model, which relies on frequent and systematic observation and coaching to improve instructional practices

We will recruit, on average, 10 schools from each district. Based on the Common Core of Data and National Council on Teacher Quality (2018), 358 districts in 29 states meet these criteria. Although we will not be able to generalize to all schools, we will obtain valid estimates of the impact of the intervention for a policy-relevant sample of schools that meet our eligibility requirements and are willing to participate.

Selection of schools. Within the participating districts we will invite eligible schools to participate in the study. We will include 100 schools in the 2020–2021 and 2021-2022 school years. Schools eligible and willing to participate in the evaluation will be randomly assigned to the treatment or control group as described in section B1 above.

Selection of teachers. Within the treatment schools, we will include all teachers on the two teams of teachers being supported by teacher leaders. This includes the two teacher leaders and the 2 to 5 other teachers in their same grade or grade/subject. Within each control school, we will include all teachers in the same grades and subjects as their matched treatment school. Teachers will participate in the study in the 2020–2021 and 2021-2022 school years.

Selection of students. To estimate impacts on student outcomes after one year of the intervention, we will define the student analysis sample as the students who are on study teachers' rosters at the beginning of the first intervention school year (fall 2020). To estimate the second year impacts, we will define the student analysis sample as the students enrolled in study schools in fall 2020 who are projected to be in the grades covered by the study teams in fall 2021, along with the new students assigned to study teachers in fall 2021.

2. Data collection

We have already received clearance for the district interview protocol, teacher leader applicant background form, and the school information questionnaire. At this time, we are requesting clearance for the teacher leader activity form, principal survey, teacher survey, and administrative data collection request memo.

a. Data collection approved under previous clearance request

District interviews. To describe the key strategies used by TSL grantees to improve educator effectiveness as well as understand the specific activities they implemented, we will conduct interviews in spring 2020 and spring 2021 with all 25 districts covered by a 2017 TSL grant. The spring 2020 interview will be conducted in three stages. The stages are:

1. *Initial email.* We will inform TSL grantee districts about the interviews and ask one question about how they use the TSL grant funds to improve educator effectiveness.

2. *Introductory interview.* We will conduct a 15-minute phone call to learn more about the activities the district identified in the email.

3. *In-depth interview.* We will hold a 45-minute follow-up call to get additional information about the three highest-priority activities identified in the introductory interview.

The spring 2021 interview will feature many of the same topics as the first, but will focus on experiences since the first interview.

Teacher leader applicant background form. Ideally, teachers selected for the teacher leader role will be the applicants with strong teaching and coaching skills. To examine the extent that this occurred, we will have principals complete a teacher leader applicant background form.

School information questionnaire. We will use the school information questionnaire for several purposes.

- To form pairs of similar schools for random assignment.

- To identify high-priority teachers that might benefit from coaching.

- To estimate the impact on teacher leaders and their students.

b.   Data collection under current clearance request

Teacher leader activity form. To monitor and describe teacher leader activities, teacher leaders will indicate on a weekly basis what they did in their role, whom they supported (specific teachers, the full team, a subset of the team), and the focus of their activities. We will also use these data to examine whether impacts may have been associated with the amount and content of the coaching they provided. It is expected that teachers will spend 10 minutes per week completing the web-based form.

Principal surveys. We will administer a 30-minute web-based survey in spring 2021 to all principals of treatment and control schools and 100 randomly selected schools that receive TSL funds to support teacher leader roles for three main purposes:

1.   To describe the implementation of the teacher leader role in treatment schools

2.   To estimate the impact of teacher leaders on principals' satisfaction and recruitment strategies

3.   To compare the teacher leader role implemented by TSL schools and non-TSL treatment schools

The principal survey will collect information related to the types and frequency of coaching, mentoring, common lesson planning, and professional development in the principal's school, strategies for recruiting and retaining teachers, activities of teacher leaders, and principal satisfaction.

We will also survey all principals of treatment and control schools again in spring 2022 to contribute to the analysis of implementation and impacts after two years.

Teacher surveys. To learn about teachers' perspectives about the type and amount of support they receive and to estimate the impact of teacher leader roles on teacher satisfaction, we will administer a 35-minute web survey in spring 2021 to teachers in the same 100 randomly-selected TSL-supported schools from the principal survey and to teachers in treatment and control schools. We will administer a second round of the teacher survey in spring 2022 to teachers in the treatment and control schools only to estimate impacts after two years of implementation. For consistency with the survey sample in the impact evaluation component, all teachers in the TSL sample will teach English language arts or math in grades 3 through 6. In particular, we will select teachers in TSL grantee schools from the grades and subjects in those schools that are served by teacher leaders (within that grade range). As with the principal survey, we will use the information from teachers to describe implementation of the teacher leader role in treatment schools, estimate impacts on teacher-reported outcomes, and compare the teacher leader role implemented by TSL schools and non-TSL treatment schools.

The teacher survey will ask about demographic and background characteristics, the types and frequency of support, coaching, mentoring, and professional development teachers receive or provide.

District administrative records on teachers. We will collect the following district administrative data on teachers in treatment and control schools in fall 2020, fall 2021, or fall 2022:

- *Student rosters of study teachers*. In order to estimate the impact of teacher leaders on the student achievement of two key subgroups (students of high-priority teachers and those of the teacher leaders) we will collect student rosters in the fall of each implementation year (fall 2020 and fall 2021).

- *Teachers' district evaluation scores from the year before the intervention (the 2019– 2020 school year)*. We will use these data to compare the effectiveness of teacher leaders with other teachers in their schools and compare the evaluation ratings of the teacher leader with the ratings of the teachers they support. We will also use this information to examine the correlation between the effectiveness of a school's teacher leaders and their impacts on student achievement.

- *Teachers' school, grade, and subject assignments*. In order to examine the impact on the retention and recruitment of effective teachers, in fall 2022 we will collect information on teachers' school, grade, and subject assignments for teachers in treatment and control schools in spring 2020 (before random assignment), fall 2020, fall 2021, and fall 2022.

- *Teachers' demographic characteristics*. To describe the study sample, we will collect information on teachers' demographic characteristics (for example, age, gender and race), educational background (e.g. certifications, degrees, and scores on licensure or certification exams), and years of teaching experience. We will collect these data at the same time we collect data on teachers' school, grade and subject assignments.

District administrative records on students. To estimate the impact of teacher leaders on student achievement, we will collect administrative data on students from each district in the random assignment evaluation. We will collect state test score data in reading and math from the baseline year (2019-2020) and each implementation year (2020-2021 and 2021-2022). We will use demographic, socioeconomic, and baseline test score data to describe the students in the study and compare the characteristics of students in treatment and control schools.

3.   Estimation procedures

This study will conduct analyses for an implementation brief and a final report.

Analyses for implementation brief. The implementation brief will describe TSL grantee districts and their programs. A key goal of this part of the study is to enhance our understanding

of the strategies that TSL grantee districts use to improve educator effectiveness and how they carry out those strategies. We will also describe several aspects of grantee districts and their programs. For example, we will describe the location and size of grantee districts, the types of activities supported by TSL grants, and districts' challenges in implementing and sustaining TSL activities.

Analyses for final report. The final report will cover four broad sets of analyses, including impact, subgroup, non-experimental, and implementation analyses.

Impact analyses. We will use regression models to estimate the impact of teacher leader roles on principal, teacher, and student outcomes after each year of implementation. Key outcomes include, principals' strategies for recruiting teachers, the retention and recruitment of effective teachers, principals' and teachers' satisfaction, and students' math and ELA achievement.

To estimate impacts on student, teacher, and principal outcomes, we will use the following regression model:

$$(1) \quad y_{isb} = \beta T_{sb} + \gamma X_{isb} + \delta Z_b + \varepsilon_{isb},$$

where

- y_{isb} is the outcome for individual *i* (either student, teacher, or principal) in school *s* and random assignment block *b*

- T_{sb} indicates whether the school was assigned to the treatment group to receive funding for up to two teacher leaders

- X_{isb} is a set of student-level covariates (included in the student analyses only)

- Z_b is a set of indicators for the study's random assignment blocks (matched pairs of schools)

- ε_{isb} is an individual-level error term

- γ and δ are vectors of coefficients

- β represents the average impact of the teacher leader roles

We will estimate robust standard errors that account for the clustering of outcomes within schools in the teacher and student analyses.

For the student-level regressions, the covariates will include baseline math and reading achievement, student demographic characteristics (such as gender, race, ethnicity, free or reduced-price lunch eligibility, special education status, and Limited English proficiency). We do

not plan to include covariates in the principal- or teacher-level regressions because these analyses will include smaller numbers of individuals, and we expect that once we include the block fixed effects, these covariates will explain little of the variation in outcomes.

In our main analysis, we will give equal weight to each school (regardless of the number of students or teachers in the analysis sample in each school). Estimates that give equal weight to each school will reflect the impact of teacher leader roles on the average school in the sample. However, as a sensitivity test, we will also estimate the impacts of the teacher leader role for the average teacher or student by giving equal weight to each student or teacher in the analysis.

We will use the standard (frequentist) statistical approach to estimate the impact of teacher leader roles (β) along with the standard error, statistical significance, and p-value of this estimate. However, to avoid common misinterpretation of p-values (Wasserstein and Lazar 2016; Greenland et al. 2016), we will complement these estimates with a Bayesian measure of the probability that the intervention had a positive effect.

Subgroup analyses. To help districts and schools identify which teachers and students might benefit the most from teacher leader roles, we will estimate impacts for subgroups defined by teacher and student characteristics, including:

- Teachers identified by their principals (prior to random assignment) to have the greatest potential to benefit from the coaching

- Novice teachers (less than 3 years of teaching experience) and those with low baseline evaluation scores (as measured by their district evaluation system)

- Teacher leaders (those chosen to be a teacher leader in a treatment school and those identified prior to random assignment by principals of control schools as likely to be chosen to be a teacher leader)

- Students with high-needs (based on their eligibility for free or reduced-price lunch) or low baseline achievement (below median achievement in the district's study schools at baseline)

Non-experimental analyses. We will conduct non-experimental analyses to better understand the effectiveness of teacher leader roles and their potential application for policy and practice.

- **Correlational analyses.** The correlational analyses will help districts and schools decide how to adopt or refine their teacher leader roles. We will conduct several analyses to explore how various characteristics of the teacher leader roles might be related to the impacts of the teacher leader on student achievement. In particular, we will examine whether estimated impacts tend to be larger in matched pairs of schools in which the treatment school teacher leader has particular characteristics (such as having better

baseline evaluation scores) or engages in particular activities (such as doing more modeling of teaching practices).

- **Assess cost-effectiveness of teacher leader roles.** Districts choosing how to spend limited resources will benefit from knowing not only what impacts the teacher leader role has on principal, teacher, and student outcomes but also whether this strategy is cost-effective relative to other policies. The cost of implementing teacher leader roles includes teacher leader stipends, compensation for teachers who cover teacher leader's release time, support to help principals select teacher leaders, and initial and follow-up training for teacher leaders and their principals. If the teacher leader role has positive impacts on student achievement, we will calculate a cost-effectiveness ratio—total per-student cost of this strategy divided by its total impact. We will compare this ratio to those calculated for other policies that have been rigorously evaluated and that districts could implement instead of teacher leader roles, such as pay-for-performance or incentives to work in low-performing schools. Thus, these findings will yield lessons about the most promising components to emphasize in human capital management systems.

Implementation analyses for the random assignment evaluation. Two important goals of the implementation analyses for the random assignment evaluation are to highlight overarching lessons for the TSL program and to provide context for the impact results. We will conduct several implementation analyses.

1. We will compare TSL grantee and evaluation districts, including district, principal and teacher characteristics and aspects of their teacher leader roles. This comparison will help policymakers understand the extent to which findings based on the schools in the random assignment evaluation may be relevant for TSL grantee districts.

2. We will describe implementation of teacher leader roles in the treatment schools, such as teacher leader qualifications and characteristics, the composition of teacher teams, and how teacher leaders spend their time. A detailed understanding of the teacher leader roles implemented will provide important information for districts considering implementing similar teacher leader roles, support replication of these roles in other districts, and provide context for impact findings.

3. We will describe and compare the amount and types of supports teachers in treatment and control school receive, such as teachers' exposure to coaching and mentoring. A clear description of what the control group schools are doing compared to the treatment schools can help identify possible reasons for the presence or absence of impacts.

4.   Degree of accuracy needed

In the analysis for the implementation report, we will conduct interviews with all 2017 TSL grantee districts. Thus, the resulting descriptive statistics based on data from these interviews will cover the relevant universe and not be subject to sampling error.

In the analysis contributing to the impact report, an analysis of the statistical power of the design led to the following conclusions:

- The target sample size of 100 schools will enable the study to detect impacts on student test scores as small as 0.11 standard deviations in the full sample—students of teacher teams in treatment schools and those from the identical grades and subjects in the paired control schools (Table B.1). This effect is smaller than the average effect of coaching (0.18) documented in a meta-analysis by Kraft et al. (2018), so the evaluation will be able to detect effects equal to or even a bit smaller than the typical coaching intervention.

- Among students of high-priority teachers—about half of the sample in each school—the minimum detectable effect will still be 0.11.

- For teacher satisfaction and retention, the minimum detectable impacts of 11 and 10 percentage points are similar to the average effects of other teacher-focused interventions (Chiang et al. 2017 and Clotfelter et al. 2008).

Table B.1 displays the minimum detectable effect sizes for the full sample of teachers as well as a 50 percent subsample. The full sample will include 100 schools, 50 in the treatment group and 50 in the control group. The study design will maximize power to detect impacts by matching schools with similar characteristics into blocks within each district. The characteristics used to match schools will include grade span, grades and subjects that they propose for teacher leader roles, and average school baseline performance. We will also use students' prior test scores as covariates in the impact analysis to increase statistical power.

The calculations in Table B.1 assume the following:

- 80 percent power and 5 percent significance level for a two-tailed test

- An average of 8 teachers per school and 22 students per teacher

- A response rate of 95 percent for district records, based on all districts responding, but missing records for 5 percent of cases

- A response rate of 85 percent for the teacher survey

- 87 percent of teachers receive the top two evaluation scores

- A sample of 100 schools, with schools split evenly between the treatment and control groups

- For student achievement, 16 percent of the variation in outcomes occurs across schools, and 80 percent of the cross-school variation and 40 percent of the within-school variation can be explained by covariates (Deke et al. 2010)

- For teacher satisfaction and retention outcomes, assumptions on school-level clustering of outcomes and explanatory power of covariates come from Wellington et al. (2016)

- For teacher retention, assumptions on the percentage of teachers that receive the top two evaluation scores come from Garet et al. (2017).

Table B.1. Minimum detectable effect sizes with 100 schools

Outcome	Data source	Units	Minimum detectable effect		
			Full sample	50 percent subsample of schools	50 percent subsample of teachers in all schools
Students' reading and math test scores in spring 2021	District records	Standard deviations	0.11	0.15	0.11
Percentage of teachers who are satisfied with their jobs in spring 2021	Teacher survey	Percentage points	11	16	14
Retention of effective teachers (based on baseline evaluation scores) from spring 2020 to fall 2021	District records	Percentage points	10	15	14

5.   Unusual problems requiring specialized sampling procedures

We do not anticipate any unusual problems that require specialized sampling procedures.

6.   Use of periodic (less frequent than annual) data collection cycles to reduce burden

In order to limit respondent burden as much as possible, we have carefully considered what the minimum amount of data is needed to answer the research questions and how to structure the data collection. For example, we will request administrative data no more than once a year, and whenever possible, we will request multiple years of data within a single request to reduce the number of separate requests.

B3. Methods to maximize response rates and deal with nonresponse

The study will employ multiple strategies to maximize response rates while minimizing burden on respondents. These strategies include establishing positive relationships with respondents and school and district staff; sending letters to respondents to alert them to upcoming requests to complete the surveys; providing data collection instruments that are accessible in both web and mobile formats and accepting administrative data files in formats that are most convenient for districts. To reassure respondents on the confidentiality of the data they provide, we will include a statement on confidentiality and data collection requirements (Education Sciences Reform Act of 2002, Title I, Part E, Section 183) in all letters and data collection instruments. Finally, we will include a reminder for TSL grantee districts that their participation in the district interview for this ED-sponsored study is expected as a condition of their grant. For other respondents on other data collection instruments, we will include a statement indicating that participation is voluntary, yet we will also emphasize the importance of each response for the study findings.

District interviews. Flexibility in scheduling and conducting district interviews will help us obtain high response rates. Members of the study team will identify the contact at each district best suited to respond to the interview protocol. Initial emails will be sent to each contact to identify a time that would be most convenient for their schedule. Reminder emails will be sent to non-responders, highlighting the importance of their participation and our flexibility to meet their scheduling needs. The use of a semi-structured interview protocol will allow flexibility in focusing on key concepts and themes, while staying within the confines of the one-hour interview. We anticipate that these qualities, in addition to ED support, will facilitate a response rate of 100 percent.

Teacher leader applicant background form. Teacher leader application forms will be distributed to principals in all treatment schools as an attachment sent via email. Principals will complete the form for each teacher leader applicant and return to the study team through an upload to the study's secure file transfer site (due to teacher PII collected on the form). The study team will coordinate with principals throughout the fielding period to confirm that all forms are being returned in a timely manner via periodic email and phone check-ins. Utilizing the communication channels and relationships necessary for implementation of the study intervention, as well as the development of a succinct form, we anticipate a response rate of 100 percent. The study team will identify nonresponse and reporting errors by checking for complete and reasonable answers as soon as forms are received and will follow-up with respondents as needed for correction.

School information questionnaire. School information questionnaires will be distributed to principals in all study schools prior to random assignment, as an attachment sent via email. Principals will complete the form and return to the study team through an upload to the study's secure file transfer site (due to teacher PII collected on the form). The study team will coordinate with principals throughout the fielding period to confirm that questionnaires are being returned in a timely manner via periodic email and phone check-ins. Utilizing the communication

channels and relationships necessary for implementation of the study intervention, as well as the development of a succinct form, we anticipate a response rate of 100 percent.

Teacher leader activity forms. Teacher leaders in treatment schools will receive invitations and weekly email reminders to complete each weekly web-based checklist. The study team will monitor checklist completion and follow-up with individual teachers to ensure timely completion of each log. Through district and school support of the study, as well as a succinct, easily accessible data collection instrument, we anticipate a 90 percent response rate.

To ensure high quality data, we pretested the instrument for clarity, accuracy, length, flow, and wording. Based on the pretest, the instrument is estimated to take 10 minutes to complete each week. Additionally, data from each completed web checklist will be reviewed throughout the fielding period for accuracy and consistency. We will follow up with respondents for clarification, as needed. The use of the web mode allows for sophisticated skip logic and fills within the instrument, further improving the overall reliability of the data collected.

Principal and teacher surveys. To ensure completion of surveys, we will send an invitation letter both by mail and email with a link to the web-based survey. In previous studies in similar settings, we have found that some sample members do not check school email accounts frequently. Therefore, we will also give the option of completing a hard-copy survey, which will be mailed to them at their schools. Over the data collection period, we will send email and mail reminders. We will also conduct reminder phone calls and allow sample members to complete the survey over the phone with a trained supervisor, if they prefer. We propose to offer $30 to respondents who complete a survey. We will also coordinate in-person school visits with our field staff during the last four weeks of data collection to provide teachers with a hard-copy version of the teacher survey. This in-person connection has helped motivate sample members to participate in past surveys. Based on Mathematica's experience in conducting surveys with principals and teachers, we expect at least an 85 percent response rate for the principal and teacher surveys. Because sample members will receive full information on study commitments, we anticipate high levels of cooperation. We will be courteous but persistent in our follow up with participants who do not respond quickly to our attempts to reach them.

To ensure high quality data, we have used many items that have been successfully used in other federal studies. We pretested each survey instrument for clarity, accuracy, length, flow, and wording. Pretest administrations indicated that the principal survey should take approximately 30 minutes to complete, while the teacher survey pretests ran longer than the planned 30 minutes. As a result, we revised the teacher survey to remove superfluous text and eliminated questions that were less central to the planned analysis. The final instruments included with this request are reflective of these revisions. Based on the pretest and subsequent revisions, the principal survey is estimated to take 30 minutes to complete and the teacher survey is estimated to take 35 minutes to complete. Additionally, data from completed web surveys will be reviewed throughout the fielding period for accuracy and consistency. We will follow up with respondents for clarification, as needed. The use of the web mode allows for sophisticated skip logic and fills

within the instrument, further improving the overall reliability of the data collected. Finally, for surveys that are completed on paper, trained quality-control staff will identify item nonresponse and reporting errors by checking for complete and reasonable answers as soon as a hard-copy questionnaire is received and will follow-up with respondents if problems are identified.

District administrative records. Because we will develop an MOU with each district specifying in detail all data requirements, we anticipate full district participation for administrative records. To further solidify administrators' cooperation, we will adhere to additional data collection requirements that districts may have such as preparing research applications and providing documentation of institutional review board (IRB) approvals. Reducing districts' burden in the submission of study data will facilitate attaining a response rate of 100 percent on student records and educator administrative data. Federal rules permit ED and its designated agents to collect student demographic and existing achievement data from schools and districts without prior parental or student consent (Family Educational and Rights and Privacy Act (FERPA) (20 U.S.C. 1232g; 34 CFR Part 99)). To maximize the response rate and minimize burden on schools and parents, we will follow these federal rules.

B4. Tests of procedures or methods to be undertaken

As much as possible, data collection instruments for the study draw on surveys, forms, and protocols that have been used successfully in previous federal studies. For example, the principal and teacher surveys include items used on the Evaluation of Data-Driven Instruction, the Impact Evaluation of the Teacher Incentive Fund, and the Impact Evaluation of Support for Principals.

We did not pretest the school records data request as it was closely modeled on forms that have been effectively used for other studies, such as the Impact of Teacher Feedback using Classroom Videos and the Impact Evaluation of the Teacher Incentive Fund.

The teacher leader activity form, principal survey and teacher survey pretests assessed the content and wording of individual questions, organization and format of the questionnaire, respondent burden time, and potential sources of response error. Pretest participants were recruited across six schools in five school districts. In order to resemble our study sample, districts recruited for the pretest were either TIF program grantees (which included similar teacher leader structures) or other districts that indicated having teachers in similar roles. We conducted the pretest of the teacher leader activity form with five teachers in teacher leader positions similar to our proposed intervention. Principal and teacher survey pretests were conducted with four principals and six elementary school teachers (3rd through 6th grade teachers). The purpose of the pretest was to identify problems that study respondents might have providing the requested information and to confirm the level of burden. We sent a full survey packet to pretest respondents and asked them to complete the survey. Respondents returned completed forms by email. The study's instrument design team conducted a debriefing telephone interview with each respondent reviewing problems respondents may have encountered and following a protocol to probe on a number of items to be sure the survey questions were

communicated clearly and collected accurate information. The results of the pretest were used to revise and improve the teacher leader activity form and survey instruments. Pretest administrations indicated the teacher survey ran longer than the planned 30 minutes. As a result, we revised the survey to remove superfluous text and eliminated questions that were less central to the planned analysis. The final instruments included with this request are reflective of these revisions. Respondent burden is estimated at 10 minutes for the teacher leader activity form (each weekly response), 30 minutes for the principal survey and 35 minutes for the teacher survey.

In order to address issues that may arise during the fielding of data collection instruments, the study will provide a help desk, and study staff will be available to answer questions throughout the data collection period. Staff will be trained to respond to frequently asked questions about the study and individual forms, so they can provide technical assistance and report any issues that come up in the field.

B5. Individuals consulted on statistical aspects of the design and on collecting and analyzing data

The following individuals were consulted on the statistical aspects of the study:

Table B.2. Individuals consulted on statistical design

Name	Title	Telephone Number
Hanley Chiang	Senior Researcher, Mathematica	617-674-8374
Phil Gleason	Associate Director, Human Services Research and Senior Fellow, Mathematica	202-264-3443
Kristin Hallgren	Senior Researcher, Mathematica	609- 275-2397
Jason Margolis	Professor of Education, Duquesne University	412-396-6106
Alison Wellington	Senior Researcher, Mathematica	202-484-4696

The following individuals will be responsible for data collection and analysis:

Table B.3. Individuals responsible for data collection and analysis

Name	Title	Telephone Number
Julie Bruch	Senior Researcher, Mathematica	617-301-8964
Megan Davis-Christianson	Lead Program Analyst, Mathematica	609-275-2361

Name	Title	Telephone Number
Phil Gleason	Associate Director, Human Services Division and Senior Fellow, Mathematica	202-264-3443
Kristin Hallgren	Senior Researcher, Mathematica	609- 275-2397
Alicia Harrington	Survey Researcher, Mathematica	609-750-3193
Mariesa Herrmann	Senior Researcher, Mathematica	609-716-4544
Jason Margolis	Professor of Education, Duquesne University	412-396-6106
Alison Wellington	Senior Researcher, Mathematica	202-484-4696
Eric Zeidman	Associate Director, Human Services Division and Senior Survey Researcher, Mathematica	609-936-2784

REFERENCES

Chiang, Hanley, Cecilia Speroni, Mariesa Herrmann, Kristin Hallgren, Paul Burkander, and Alison Wellington. "Evaluation of the Teacher Incentive Fund: Final Report on Implementation and Impacts of Pay-for-Performance Across Four Years." NCEE 2017-4004. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2017.

Clotfelter, Charles T., Elizabeth Glennie, Helen F. Ladd, and Jacob L. Vigdor. "Would Higher Salaries Keep Teachers in High-Poverty Schools? Evidence from a Policy Intervention in North Carolina." *Journal of Public Economics*, vol. 92, no. 5-6, 2008, pp. 1352-1370.

Deke, John, Lisa Dragoset, and Ravaris Moore. "Precision Gains from Publically Available School Proficiency Measures Compared to Study-Collected Test Scores in Education Cluster-Randomized Trials." NCEE 2010-4003. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2010.

Garet, Michael S., Andrew J. Wayne, Seth Brown, Jordan Rickles, Mengli Song, and David Manzeske. "The Impact of Providing Performance Feedback to Teachers and Principals." NCEE 2018-4001. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2017.

Kraft, Matthew A., David Blazar, and Dylan Hogan. "The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence." *Review of Educational Research*, vol. 88, no. 4, 2018, pp. 547–588.

Wellington, Alison, Hanley Chiang, Kristin Hallgren, Cecilia Speroni, Mariesa Herrmann, and Paul Burkander. "Evaluation of the Teacher Incentive Fund: Implementation and Impacts of Pay-for-Performance After Three Years." NCEE 2016-4004. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, August 2016.

This page has been left blank for double sided copying.

Mathematica
Progress Together

mathematica.org