

# Impact Evaluation of Teacher Residency Programs

## Supporting Statement for Paperwork Reduction Act Submission

### PART B: Collection of Information Employing Statistical Methods

**September 2021**

**Submitted to:**

U.S. Department of Education  
Institute of Education Sciences  
550 12th Street, SW  
Washington, DC 20202  
ATTN: Meredith Bachman, Project Officer  
Contract: 91990019C0066

**Submitted by:**

Mathematica  
P.O. Box 2393  
Princeton, NJ 08543-2393  
Phone: (609) 799-3535  
Fax: (609) 799-0005  
Project Director: Melissa Clark  
Project Reference: 50911



## Contents

Part B: Collection of Information Employing Statistical Methods.....	1
Introduction.....	1
B.1. Respondent universe and sampling methods.....	2
B.2. Information collection procedures.....	4
B.2.1. Statistical methods for sample selection.....	4
B.2.2. Estimation procedures.....	4
B.2.3. Degree of accuracy needed.....	8
B.2.4. Unusual problems requiring specialized sampling procedures.....	9
B.2.5. Use of periodic (less than annual) data collection to reduce burden.....	9
B.3. Methods to maximize response rates.....	10
B.4. Test of procedures.....	11
B.5. Individuals consulted on statistical aspects of design.....	11
References.....	13

## Exhibits

Exhibit B.1. Key research questions.....	1
Exhibit B.2. Timing of data collection activities.....	2
Exhibit B.3. Respondent universe and sampling methods.....	4
Exhibit B.4. Minimum detectable effects.....	9

## Part B: Collection of Information Employing Statistical Methods

### Introduction

The U.S. Department of Education (ED)'s Institute of Education Sciences (IES) requests clearance for data collection activities to support the first large-scale rigorous study of teacher residency programs. These programs are rapidly increasing in popularity, as a potential way to address persistent inequities in student access to high quality teachers. They combine education coursework with an extensive full-year apprenticeship or “residency,” under the supervision of an experienced mentor teacher, to prepare teachers to fill hard to staff positions in partner districts. They also include financial support to residents in exchange for a commitment to stay in those districts for at least three to five years.

This second of two requests for the study covers data collection activities that will support describing residency programs nationally and estimating the effectiveness of teachers who graduate from these programs. A prior request (1850-0960, approved 4/26/2021) covered the collection of classroom rosters from schools to support random assignment of students to participating teachers.

Collecting information about residency programs is critical given ED's need to provide rigorous evidence on promising approaches to teacher preparation and the increasing popularity of the residency approach. Efforts to better prepare new teachers and retain effective ones are a central goal of Title IIA of the Every Student Succeeds Act (ESSA). Seen as a promising way to achieve these goals, teacher residency programs have grown considerably since the first programs began in the early 2000s, with substantial investment from ED. In addition, ESSA, passed in 2015, allows states and districts to use Title IIA funds to support teacher residency programs, and 15 states and the District of Columbia have promoted these programs in their ESSA plans (National Center for Teacher Residencies 2018).

Despite the growth in residency programs, there is little rigorous evidence on whether they are an effective approach for preparing new teachers and addressing persistent inequities in access to high quality teachers. Some evidence indicates that residency graduates are more effective (Garrison 2019; Papay et al. 2012) and have higher retention in their districts (Papay et al. 2012; Silva et al. 2015) than other teachers. However, findings come from small scale nonexperimental studies and are mixed. Additionally, there is no evidence on the costs or cost-effectiveness of teacher preparation programs. This study will provide evidence on the effectiveness and costs of teacher residency programs and will examine ways to improve teacher preparation more broadly.

IES has contracted with Mathematica and its partners, the National Center for Teacher Residencies, Decision Information Resources, Clowder Consulting, and IRIS Connect (together, the “study team”), to conduct the evaluation, including all data collection. The evaluation will collect data to answer and report on the questions shown in Exhibit B.1.

**Exhibit B.1. Key research questions**

Key research questions
1. <b>What are the characteristics of residency programs nationally, what strategies do they use to better prepare new teachers, and what are the characteristics of those graduating from these programs?</b>
2. <b>Are residency graduates more effective than other teachers?</b> Does this change as teachers progress in their careers?
3. <b>Do residency graduates remain in teaching longer than other teachers?</b>
4. <b>What explains any differences in effectiveness and retention between residency graduates and other teachers?</b> Are the differences explained by the types of candidates residency programs select? Particular features of residency programs?
5. <b>Are residency programs a cost-effective strategy for improving student achievement?</b>

The study will collect information from multiple data sources, shown in Exhibit B.2.

**Exhibit B.2. Timing of data collection activities**

Data source	Sample	Respondent	Mode and timing
<b>Data collections included in previous clearance request (OMB 1850-0960, approved 4/26/2021)</b>			
Classroom rosters	Students assigned to participating teachers in fall 2021 and fall 2022	School administrative staff	Electronic or paper form collected four times in 2021–2022 and 2022–2023 (before and after random assignment)
<b>Data collections included in current clearance request</b>			
Administrative data on students	Students assigned to participating teachers in fall 2021 and fall 2022	District data office staff	Electronic records collected in fall 2022 and fall 2023
Administrative data on teachers	Participating teachers in 2021–2022 and 2022–2023	District data office staff	Electronic records collected in fall 2022 and fall 2023
Teacher surveys	Participating teachers in 2021–2022 and 2022–2023	Teachers	Web survey in spring 2022 and 2023
Residency program interviews	All residency programs nationwide	Residency programs directors	Telephone interviews in summer 2022
District cost interviews	Participating districts	District residency coordinator	Telephone interviews in spring 2022

**B.1. Respondent universe and sampling methods**

The study will include two different groups of respondents to learn about (1) the characteristics of residency programs nationwide and (2) the effectiveness and retention of residency graduates.

To describe all residency programs nationally, the study team will interview nearly the full universe of residency programs nationwide. As of January 2021, we had identified approximately 140 such programs

through a comprehensive web search. We will update this search in spring 2022 to identify the full set of residency programs in operation at that time. We will then conduct interviews with most of these programs. For the residency program interviews, the study team intends to interview all programs outside of Louisiana and a subset of approximately 10 programs within Louisiana. Under Louisiana’s Believe and Prepare program, all traditional undergraduate teacher preparation programs were required to have their graduates complete a full-year residency under an expert mentor by the 2017–18 school year. As a result, Louisiana has a much larger number of residency programs than in other states, and it will be more efficient to interview a sample of these programs. The study team will use publicly available information to select a sample from Louisiana that is geographically diverse within the state and varies on other characteristics that could be associated with teacher preparation approaches. These include program size; program length; subjects of focus (general education versus specific areas, such as English learner or special education programs); whether they serve public schools, private, or charter schools; and degree type (graduate versus undergraduate degree).

To learn about the effectiveness and retention of residency graduates relative to other teachers, the study will use a random assignment design in a purposive sample of residency programs, districts, and schools. We describe the selection of the sample below.

- **Selection of residency programs.** The study will include a sample of approximately 25 residency programs. The study team will first identify programs that (1) offer at least a full-year residency with a mentor teacher and (2) prepare at least some teachers to teach English language arts or math in grades 3 through 8 (the grades and subjects that will be the focus of the study). An estimated 81 programs nationwide meet these initial eligibility requirements out of an estimated 140 total nationwide. During the study’s recruitment effort, the study team will reach out to these programs to verify that they meet the study’s eligibility requirements and determine whether they are willing and able to participate in the study.
- **Selection of districts.** After identifying residency programs that are eligible and willing to participate, the study team will reach out to these programs’ partner districts to identify a sample of approximately 15 districts. The study team will focus outreach on districts that employ a sufficient number of residency program graduates (at least seven), to ensure the study can efficiently meet its sample size targets.
- **Selection of schools.** Within districts that are willing to participate in the study, the study team will reach out to schools that employ residency graduates (based on lists provided by the programs). In this outreach, the team will seek to identify schools that will have at least one eligible “classroom match” in the 2021–2022 or 2022–2023 school year and are willing and able to participate in the study. Classroom matches are groups of classes within the same grade and subject, with at least one taught by a residency graduate and at least one taught by some other type of teacher (or a “comparison teacher”). Eligible classroom matches will be made up of either English language arts, math, or general education classes in grades 3 through 8. We anticipate that the sample will include approximately 70 schools.
- **Selection of teachers.** Within schools that have at least one eligible classroom match, the sample will include all teachers (residency graduates and comparison teachers) in the eligible matches. We expect that each match will include one residency graduate and 2.5 comparison teachers, on average. The study will aim to include a sample of 100 residency graduates and 250 comparison teachers for a total sample of 350 teachers over the two study school years. To the extent possible, we will aim to include matches in which residency graduates and comparison teachers have similar levels of experience.

This will help ensure a sufficient sample of early career residency graduates and comparison teachers (those in their first five years of teaching) within the same matches, so we can examine residency graduates’ relative effectiveness among early career teachers. We anticipate that the sample will include roughly 50 early career residency graduates and 50 early career comparison teachers in the same matches.

- Selection of students.** Within each participating classroom match, the study team will randomly assign students in that grade to a class taught by a residency graduate or a class taught by a comparison teacher, in the summer before the school year begins. Random assignment will ensure that teachers begin the year with similar students. The study can then attribute any differences in student achievement at the end of the year to the teachers rather than differences in their students or schools. The sample will include all students randomly assigned to a matched classroom. To encourage compliance, before random assignment we will allow schools to exempt a limited number of students (no more than 10 percent per class) from random assignment and to instead assign those students at the school’s discretion. We will exclude students who are not randomly assigned from the sample. We anticipate that each class will include an average of 24 randomly assigned students, for a total sample of 8,400 students.

Exhibit B.3 describes the respondent universe for each data collection activity within the sample described above.

**Exhibit B.3. Respondent universe and sampling methods**

Data source	Respondent	Respondent universe (estimated)	Sampling approach
Classroom rosters <sup>a</sup>	School administrative staff (one per school)	70	Census of participating schools
Administrative data on students	District data office staff (one per district)	15	Census of participating districts
Administrative data on teachers	District data office staff (one per district)	15	Census of participating districts
Teacher surveys	Teachers	350	Census of participating teachers
Residency program interviews	Residency programs directors (one per program)	115	Census of all residency programs nationwide outside of Louisiana; purposive sample of 10 programs within Louisiana
District cost interviews	District residency coordinator (one per district)	15	Census of participating districts

<sup>a</sup>This data collection was covered in a prior clearance request.

**B.2. Information collection procedures**

**B.2.1. Statistical methods for sample selection**

The study team will not use sampling methods for data collection, beyond those to select the purposive study sample described in Section B.1. A purposive sample rather than a probability sample is needed to support the study’s random assignment design. The study will include all participating residency programs, districts, schools, teachers, and students in applicable data collections.

### B.2.2. Estimation procedures

The analysis will provide comprehensive information about residency programs and their graduates. Below we discuss planned estimation procedures for each of the study's key research questions.

***What are the characteristics of residency programs nationally, what strategies do they use to better prepare new teachers, and what are the characteristics of those graduating from these programs?***

Understanding the landscape of teacher residency programs nationally will provide valuable information for districts and teacher preparation programs on the range of residency program approaches. We will calculate summary statistics such as means and percentages to describe key features of residency programs nationwide. These will include the length and intensity of the residency; residents' responsibilities in the classroom; selection, training, and compensation of mentors; coursework requirements; integration of coursework and the residency; the types of financial support that residents receive (such as stipends or loan forgiveness); and the number of years that residents commit to teach in the district. We will focus in particular on how residency programs prepare their graduates to teach special populations, including English learner students and students with disabilities. We will also describe the characteristics (such as grade and subject areas) of program graduates.

We will also compare features of the residency programs in the random assignment sample and other residency programs nationally. If the features of the two sets of programs are similar, this may suggest that the findings on residency graduates' effectiveness and retention based on the random assignment sample are relevant for the full set of residency graduates nationwide.

***Are residency graduates more effective than other teachers? Does this change as teachers progress in their careers?***

Understanding residency graduates' effectiveness relative to other teachers in the same school will provide important evidence on whether residency programs are a promising approach for improving student achievement. To examine residency graduates' relative effectiveness, we will estimate the following regression model, using data for the full sample of students who were randomly assigned:

$$(1) Y_{ijk} = \alpha_k + \beta_1 T_{jk} + EXP_{jk} \beta_2 + X_{ijk} \beta_3 + \varepsilon_{ijk}$$

where  $Y_{ijk}$  is end-of-year math or English language arts achievement score of student  $i$  who was randomly assigned to teacher  $j$  in classroom match  $k$ ;  $\alpha_k$  is a classroom match fixed effect;  $T_{jk}$  is a binary indicator for being assigned to a residency graduate (rather than a comparison teacher);  $EXP_{jk}$  is a set of controls for teachers' experience level (such as an indicator for each year of experience up to the seventh year, an indicator for having from 7 to 10 years of experience, and an indicator for more than 10 years)<sup>1</sup>;  $X_{ijk}$  is a set of baseline student characteristics (such as prior achievement, gender, race/ethnicity, eligibility for free or reduced-price lunch, English learner status, disability status);  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  are parameters to be estimated; and  $\varepsilon_{ijk}$  is a random error term. The parameter of interest,  $\beta_1$ , is the average difference in effectiveness between residency graduates and the other teachers with whom they were compared. We will cluster standard errors at the teacher level.<sup>2</sup>

<sup>1</sup> The controls for teachers' years of experience will depend on the distribution of experience among teachers in the sample.

<sup>2</sup> Clustering would not be required in a purely experimental design in which students were randomly assigned to teachers of the same experience level. However, because we are not able to do this and instead plan to control for

Residency graduates' effectiveness relative to other teachers could vary with teachers' experience. For example, because of the intensive clinical experience provided by residency programs, early-career residency graduates may be more effective than other early-career teachers. However, any differences could fade over time as both groups of teachers gain classroom experience. Understanding residency graduates' relative effectiveness at different stages of their careers will help districts consider the short- and long-term benefits of hiring a residency graduate.

To understand whether residency graduates are more effective than other teachers early in their careers and examine whether any differences in effectiveness remain or fade over time, we will conduct subgroup analyses that compare the effectiveness of (1) early-career residency graduates (in their first five years of teaching) and other early-career teachers and (2) more experienced residency graduates (more than five years of teaching) and other more experienced teachers. To do so, we will estimate a version of model (1) that includes an interaction between an indicator for whether the teacher is an early-career teacher and an indicator for whether the teacher is a residency graduate. The coefficient on this interaction term will provide information about whether the relative effectiveness of more experienced residency graduates is lower than that of early-career residency graduates, which would suggest that the relative effectiveness of residency graduates fades over time.

***Do residency graduates remain in teaching longer than other teachers?***

Residency programs could improve the retention of residency graduates by improving their preparedness for and satisfaction with teaching and by requiring them to commit to teach in the district for a minimum number of years. To examine whether residency graduates remain in their districts or schools longer than other teachers, we will use a statistical approach known as survival analysis (Miller 1997). This approach accounts for the fact that we will only observe teachers for at most three school years, not for the entire time that they work in their districts or schools.

***What explains any differences in effectiveness and retention between residency graduates and other teachers? Are the differences explained by the types of candidates residency programs select? Particular features of residency programs?***

Residency programs could lead to differences in effectiveness and retention by selecting better candidates than other preparation programs or better preparing these candidates to become teachers. These two possibilities could have different implications for teacher preparation policy. In particular, strategies to improve training could be more broadly applicable to other teacher preparation programs that are less selective than residency programs.

To shed light on how residency programs' selection and training of candidates relate to their graduates' effectiveness and retention, we will conduct analyses to explore these relationships. For example, we will examine how effectiveness and retention vary with the stipends programs provide residents or the training they provide mentor teachers. We will also examine how effectiveness and retention vary with residency graduates' and comparison teachers' characteristics and training, such as their racial and ethnic diversity and the amount of coaching and feedback they received during their preparation to become a teacher. We will use two main approaches for estimating these relationships.

To examine relationships between effectiveness and program features, we will conduct subgroup analyses using the following regression model:

---

teacher experience in the impact estimation model, clustering at the teacher level is required to account for the correlation of outcomes within teachers.



$$(2) Y_{jk} = \alpha_k + \beta_1 T_{jk} + EXP_{jk} \beta_2 + SUB_{jk} T_{jk} \beta_3 + X_{jk} \beta_4 + \varepsilon_{jk}$$

which includes an interaction term ( $SUB_{jk} * T_{jk}$ ) that measures whether the residency graduate attended the subgroup of programs that has a particular feature (such as offering high stipends to residents) or not. We will estimate similar subgroup models to examine relationships between retention and program features.

To examine relationships between effectiveness and teachers' characteristics or training, we will conduct correlational analyses using the following regression model:

$$(3) Y_{jk} = \alpha_k + \beta_1 T_{jk} + EXP_{jk} \beta_2 + CHAR_{jk} \beta_3 + X_{jk} \beta_4 + \varepsilon_{jk}$$

where  $CHAR_{jk}$  is a control variable for teachers' characteristics or training experiences—such as whether teachers share the same race or ethnicity as their students or perceptions of the support that teachers received from their mentor teacher. These control variables will be defined for both residency graduates and comparison teachers. The parameter  $\beta_3$  will provide information on whether that characteristic or training experience is associated with effectiveness. We can then calculate the average difference in that characteristic or training experience between residency graduates and comparison teachers  $\overline{CHAR}$  and determine the extent to which that difference explain differences in overall effectiveness:

$$(4) \overline{CHAR} \times \hat{\beta}_3$$

We will use a similar model to examine relationships between retention and teachers' characteristics or preparation experiences.

### ***Are residency programs a cost-effective strategy for improving student achievement?***

Districts considering partnering with teacher residency programs thus need reliable information on the cost-effectiveness of this approach—or how the costs of preparing and hiring a teacher from a residency program compare with the benefits. Although hiring a residency graduate may require more initial up-front investment from districts than hiring a teacher from another type of program, it may ultimately yield important benefits in terms of reduced teacher attrition and improved student achievement.

To measure cost-effectiveness to districts, we will examine all costs that may differ when the district hires a residency graduate instead of a teacher from another type of program. These include districts' contributions to the training that residents receive before becoming the teacher of record, as well as districts' contributions to any follow-up support provided by the residency programs. They also include any differential costs associated with hiring residency graduates or other teachers, providing them with professional development, and replacing them if they leave.<sup>3</sup>

To determine whether teacher residency programs are a cost-effective approach for districts, we will calculate a cost-effectiveness ratio. To calculate the ratio, we will first calculate the total costs to districts of preparing and hiring a residency graduate, above and beyond the cost of hiring another teacher. We

<sup>3</sup> In many cases we will measure costs in terms district resources (such as time spent recruiting and training mentor teachers or facilities required to provide professional development). To provide cost estimates of hiring residency graduates that are relevant to school districts nationally, we will determine the value of staff time and facilities based on national average prices (for example, average teacher salaries), where those data are available. We will identify these prices using a database of national average prices for educational resources (Hollands et al. 2015).

will then divide this cost by the relative effectiveness of residency graduates in improving student achievement, compared to that of a teacher from another type of program (estimated as discussed above). Because many of the costs of preparing and hiring residency graduates are incurred up front, but the potential cost savings of reduced professional development costs and reduced attrition are realized later, we will discount future costs in our calculations. We will also discount relative effectiveness in our calculations because districts do not realize benefits of increased effectiveness until residents become the teacher of record.

### **B.2.3. Degree of accuracy needed**

For the analysis of the characteristics of residency programs nationally, we will analyze data from interviews with all programs outside of Louisiana and a subsample of 10 programs from Louisiana. Thus, the resulting descriptive statistics based on data from these interviews will cover nearly all of the relevant universe and thus sampling error will be limited.

For the analysis of the relative effectiveness and retention of residency graduates, the targeted sample of 100 residency graduates and 250 comparison teachers will be sufficient to estimate meaningful differences in student achievement and teacher retention. For example, the targeted sample will yield a minimum detectable effect size of 0.08 standard deviations on student achievement (Exhibit B.4). Thus, the study will have the statistical power to detect effects similar to effects of teacher preparation programs found in other studies. For example, for teachers in grades 4 through 8, Papay et al. (2012) found positive impacts of 0.07 standard deviations on student math achievement for residency graduates in their fourth and fifth years of teaching compared with other teachers with similar years of experience. The minimum detectable effects on teacher retention are within the ranges of the average effects of other teacher-focused interventions (Chiang et al. 2017; Clotfelter et al. 2008; Silva et al. 2015).

Analyses focused on differences in effectiveness and retention among particular subgroups of teachers will also have sufficient sample sizes. For example, a subgroup of half the study teachers (such as those from residency programs that offered residents above-average stipends) would yield a minimum detectable effect of 0.11 standard deviations. However, subgroups of 25 residency graduates and their matched comparison teachers would have limited statistical power to detect effects on teacher retention. The targeted subsample of 50 early-career residency graduates matched with 50 early-career comparison teachers will yield a minimum detectable effect of 0.12 standard deviations for student achievement.

**Exhibit B.4. Minimum detectable effects**

Number of residency graduates	Number of comparison teachers	Minimum detectable effects	
		Students' achievement (standard deviations)	Percentage of teachers retained in their schools (percentage points)
<b>All experience levels</b>			
100	250	0.08	12
50	125	0.11	15
25	63	0.16	19
<b>Early-career teachers</b>			
50	50	0.12	17
25	25	0.18	20

Note: Calculations assume (1) 80 percent power and 5 percent significance level for a two-tailed test; (2) an average of 24 students per teacher; (3) a response rate of 95 percent for district records on student test scores to measure student achievement and 85 percent for the teacher survey; (4) 2.5 comparison teachers for each residency graduate in a sample of teachers from all experience levels (authors' calculations based on Schools and Staffing Survey 2011–2012 and Common Core of Data 2011); (5) one comparison teacher for every residency graduate in a sample of early-career teachers; and (6) a correlation of -0.45 between being a residency graduate and teachers' years of experience (authors' calculations from data collected from web searches and program interviews and National Teacher and Principal Survey 2017–2018). For student achievement, the calculations assume that 54 percent of the variation can be explained by covariates and the teacher-level interclass correlation is 0.15 (Schochet 2008). The calculations assume that standard errors are clustered at the teacher level. Assumptions for the teacher retention analysis come from Silva et al. (2015).

**B.2.4. Unusual problems requiring specialized sampling procedures**

There are no unusual problems requiring specialized sampling procedures.

**B.2.5. Use of periodic (less than annual) data collection to reduce burden**

To minimize burden, the study team has carefully considered the necessary frequency of data collection.

- The residency program interviews and district cost interviews will each be collected only once over the two years of the study.
- The teacher surveys will only be conducted once per study year. Annual data collection is needed because the sample of teachers will change from year to year. To minimize burden, teachers who are in the sample in both years will complete a shortened version of the survey that only includes questions about information that could have changed since the previous year.
- Administrative data collection on students and teachers will be collected once for each study school year, to ensure we receive complete information on all students and teachers in the sample in both years of the study. Annual data collection is necessary because some districts may not remain in the sample for the full two years. Districts also often prefer annual data collection, since data from previous years may be archived and more difficult for them to access later.

- Classroom rosters will be requested at four times during each of the 2021–2022 and 2022–2023 school years (Appendix A). This frequency is required to conduct random assignment at the beginning of each school year and to monitor student movement throughout each school year.

### B.3. Methods to maximize response rates

To maximize response rates, we will use strategies that have proven successful in past studies the study team has conducted with similar populations of school and district administrators, teachers, and residency programs (for example, the Evaluation of Secondary Math Teachers from Two Highly Selective Routes to Alternative Certification, the Impact Study of Feedback for Teachers Based on Classroom Videos, the Impact Evaluation to Inform the Teacher and School Leader Incentive Program, and the Implementation Evaluation of Teacher Residency Programs). Below we discuss the specific strategies we plan to use for each data collection activity.

- **Classroom rosters (Appendix A, approved 4/26/2021).** To ensure classroom rosters are returned in a timely manner, the study team will appoint a liaison for each participating school. The liaison will serve as the single point of contact for the school throughout the study. The liaison will coordinate with school administrative staff throughout the roster collection period, following up by email and telephone as needed to answer questions and encourage submissions. To make the process as easy as possible for the school administrative staff, we will allow them to submit their rosters on either the study-provided form or in the format of their choosing, uploaded to the study’s secure file transfer site. We expect a 100 percent response rate for the roster collection.
- **Administrative data on students and teachers (Appendix D).** To help ensure district cooperation with administrative data collection, the study team will develop a memorandum of understanding with each district specifying all administrative data requirements. We will also comply with any other district requirements, such as preparing research applications and providing documentation of institutional review board approvals. To encourage timely data submission, we will provide user-friendly instructions for encrypting and password-protecting files before uploading them to our secure file transfer site. We expect a 100 percent response rate for the administrative data collection.
- **Teacher survey (Appendix E).** To maximize response rates on these surveys, we will provide clear instructions and user-friendly materials, offer technical assistance for survey respondents using a toll-free telephone number or email, and monitor progress regularly. We will use a combination of reminder emails and follow-up letters to encourage respondents to complete the surveys. We will also offer \$30 incentive for survey completion. Based on Mathematica’s experience conducting surveys with teachers, we expect at least an 85 percent response rate for the teacher survey.
- **Residency program interviews (Appendix F).** To maximize response rates on the residency program interviews, the study team will send an advance email to all program directors to clearly describe the nature and importance of the study. A study team member will reach out to programs via telephone and email to set up and conduct the interviews at a time convenient for the program directors. A spreadsheet will be shared with the program director prior to the interview that they can complete on their own time. The study team will follow up with telephone calls and emails as needed to encourage participation. We will also offer a \$50 incentive for completion of this interview. We anticipate an 85 percent response rate for the residency program interviews.
- **District cost interviews (Appendix G).** The study team will work with the residency program coordinators to schedule a cost interview at a convenient time. The study team will use follow-up

telephone calls and emails to encourage participation. We anticipate a 100 percent response rate for the district cost interviews.

#### B.4. Test of procedures

We pre-tested the teacher survey, residency program interview, and district cost interviews to identify problems that study respondents might have providing the requested information and to confirm the level of burden. For example, the pretests assessed the content and wording of individual questions, organization and format of the instruments, amount of time it took to respond, and potential sources of response error.

We pretested the teacher survey with eight teachers in grades 1 through 12. Five of the teachers had participated in a residency program as part of their initial certification program, representing three different residency programs, and had between 1 and 3 years of teaching experience. The other three had participated in a traditional or alternative certification program as part of their initial teacher certification and had between 1 and 13 years of teaching experience. We sent the survey to pretest respondents via email and asked them to complete and return it by email. We then reviewed the completed surveys and conducted debriefing interviews by phone with each respondent to review any problems they may have encountered. Interviewers followed a protocol to probe on more complex or nuanced items to be sure the survey questions were clear and collected accurate information. Respondent burden to complete the survey averaged 30 minutes and ranged from 15 to 45 minutes as reported by pretest respondents. We used the results of the pretest to revise and improve the survey instrument.

We pretested the residency program interview protocol by phone with the directors of four residency programs. Before the phone call, we sent the program directors the spreadsheet instrument to be completed before the interview. Program directors gathered records and answered these questions on their own time. We then conducted a phone call to ask them the interview questions. During the interviews, we probed on the clarity and relevance of the interview questions, following a structured protocol. Respondent burden to complete the spreadsheet and interview averaged approximately 75 minutes, including the time to gather records. We used the pretest findings to revise and streamline the spreadsheet instrument and interview protocol.

We pretested the district cost interview protocol by phone, with the residency program coordinators of four districts in two states. Each district partnered with one residency program. Before the interview, we held a 20-minute orientation call with each program coordinator. During the interviews, we probed on the clarity and relevance of the interview questions, following a structured protocol. Respondent burden to gather records and complete the cost workbook averaged 60 minutes, ranging from 37 to 90 minutes. Respondent burden to complete the interview averaged 40 minutes, ranging from 26 to 50. Total respondent burden to complete the data collection process averaged 120 minutes. We used the pretest findings to revise and streamline the cost workbook and interview protocol.

### B.5. Individuals consulted on statistical aspects of design

We consulted the following people on the statistical aspects of the study:

Name	Title/affiliation
Melissa Clark	Senior Researcher, Mathematica
Mariesa Herrmann	Senior Researcher, Mathematica
Jill Constantine	Senior Vice President, Mathematica
Phil Gleason	Associate Director and Senior Fellow, Mathematica

The following people will be responsible for the data collection and analysis:

Name	Title/affiliation
Melissa Clark	Senior Researcher, Mathematica
Mariesa Herrmann	Senior Researcher, Mathematica
Eric Zeidman	Deputy Director of Human Services, Mathematica
Megan Davis	Lead Program Analyst, Mathematica
Ryan Callahan	Senior Survey Researcher, Mathematica

## References

- American Association of Colleges for Teacher Education. "Teacher Quality Partnership Grants 2019 Fact Sheet." 2019. Available at [https://secure.aacte.org/apps/rl/res\\_get.php?fid=1329&ref=rl](https://secure.aacte.org/apps/rl/res_get.php?fid=1329&ref=rl). Accessed November 11, 2020.
- 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." Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, December 2017.
- Clotfelter, Charles, Elizabeth Glennie, Helen Ladd, and Jacob 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. 55, June 2008, pp. 1352–1370.
- Garrison, Anne Walton. "Memphis Teacher Residency: Teacher Effectiveness in 2017–18." Memphis, TN: Memphis City Schools, Department of Research and Performance Management, February 2019.
- Hollands, F.M., Hanisch-Cerda, B., Levin, H. M., Belfield, C.R., Menon, A., Shand, R., Pan, Y., Bakir, I., & Cheng, H. *CostOut - the CBCSE Cost Tool Kit*. Center for Benefit-Cost Studies of Education, Teachers College, Columbia University, 2015. Retrieved from: [www.cbcssecosttoolkit.org](http://www.cbcssecosttoolkit.org), October 19, 2020.
- National Center for Teacher Residencies. "2017–18 Network Partner Report." Chicago, IL: National Center for Teacher Residencies, 2018.
- Miller, Rupert G. *Survival Analysis*. New York: John Wiley & Sons, 1997.
- Papay, John P., Martin R. West, Jon B. Fullerton, and Thomas J. Kane. "Does an Urban Teacher Residency Increase Student Achievement? Early Evidence from Boston." *Educational Evaluation and Policy Analysis*, vol. 34, no. 4, 2012, pp. 413–434.
- Raudenbush, Stephen W., Andres Martinez, Howard Bloom, Pei Zhu, and Fen Lin. "Studying the Reliability of Group-Level Measures with Implications for Statistical Power: A Six-Step Paradigm." Unpublished manuscript. Chicago, IL: University of Chicago, 2011.
- Schochet, Peter Z. "Statistical Power for Random Assignment Evaluations of Education Programs." *Journal of Educational and Behavioral Statistics*, vol. 33, no. 1, March 2008, pp. 62–87.
- Silva, Tim, Allison McKie, and Philip Gleason. "New Findings on the Retention of Novice Teachers from Teaching Residency Programs." NCEE Evaluation Brief 2015-4015. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, August 2015.

---

Mathematica  
Princeton, NJ • Ann Arbor, MI • Cambridge, MA Chicago,  
IL • Oakland, CA • Seattle, WA Tucson, AZ • Woodlawn,  
MD • Washington, DC  
EDI Global, a Mathematica Company  
Bukoba, Tanzania • High Wycombe, United Kingdom



[mathematica.org](https://mathematica.org)