

**Impact Evaluation of RTT  
and SIG: OMB Data Collection  
Package**

Part A

November 21, 2011



**MATHEMATICA**  
Policy Research, Inc.



Contract Number:  
ED-IES-10-C-0077

Mathematica Reference Number:  
06844.062

Submitted to:  
National Center for Education  
Evaluation  
555 New Jersey Ave NW, Suite 500h  
Washington, DC 20208  
Project Officer: Thomas E. Wei, PhD

Submitted by:  
Mathematica Policy Research  
P.O. Box 2393  
Princeton, NJ 08543-2393  
Telephone: (609) 799-3535  
Facsimile: (609) 799-0005  
Project Director: Susanne James-  
Burdumy, PhD

## **Impact Evaluation of RTT and SIG: OMB Data Collection Package**

Part A

November 21, 2011

**MATHEMATICA**  
Policy Research, Inc.



## **CONTENTS**

### **PART A. SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT SUBMISSION**

#### **A. Justification**

1. Circumstances Necessitating the Collection of Information
2. Purposes and Uses of the Data
3. Use of Technology to Reduce Burden
4. Efforts to Avoid Duplication of Effort
5. Methods to Minimize Burden on Small Entities
6. Consequences of Not Collecting Data
7. Special Circumstances
8. *Federal Register* Announcement and Consultation
9. Payments or Gifts
10. Assurances of Confidentiality
11. Additional Justification for Sensitive Questions
12. Estimates of Hours Burden
13. Estimates of Cost Burden to Respondents
14. Estimates of Annual Costs to the Federal Government
15. Reasons for Program Changes or Adjustments
16. Plan for Tabulation and Publication of Results
17. Approval Not to Display the OMB Expiration Date
18. Explanation of Exceptions

#### **REFERENCES**

#### **APPENDIX A: PROTOCOL FOR STATE INTERVIEWS**

#### **APPENDIX B: PROTOCOL FOR DISTRICT INTERVIEWS**

#### **APPENDIX C: SCHOOL SURVEY**

#### **APPENDIX D: DATA COLLECTION FORM FOR STATE-LEVEL DATA REQUEST**

#### **APPENDIX E: DATA COLLECTION FORM FOR DISTRICT-LEVEL DATA REQUEST**



## **PART A. SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT SUBMISSION**

This Office of Management and Budget (OMB) package requests clearance for data collection activities to support the Impact Evaluation of Race to the Top (RTT) and School Improvement Grants (SIG). The RTT-SIG evaluation will provide important information on the implementation and impacts of school turnaround efforts and educational reforms funded through these two federal grant programs. The Institute of Education Sciences (IES) at the U.S. Department of Education (ED) has contracted with Mathematica Policy Research and its subcontractors, the American Institutes for Research (AIR) and Social Policy Research Associates (SPR), to conduct this important evaluation.

The RTT-SIG evaluation will include implementation and impact components. For the evaluation of RTT, the implementation component will include semistructured interviews with state officials, and an interrupted time series (ITS) design will be used to examine the relationship between RTT and student outcomes. For the evaluation of RTT- and SIG-funded school turnaround models (STMs), the implementation component will include semistructured interviews with state and district officials and a web survey of school administrators. The plan is for the impact evaluation of STMs to be based on a regression discontinuity design (RDD).

This is the second submission of a two-stage clearance request. The package was submitted in two stages because the study schedule required that recruitment efforts begin before all the study's data collection instruments were developed. The first package (approved June 27, 2011, under OMB # 1850-0884) requested approval for recruitment of states, districts, and schools. This second package requests clearance to collect data that will support the full-scale study.

### **A. Justification**

#### **1. Circumstances Necessitating the Collection of Information**

##### **a. Statement of Need for a Rigorous Evaluation of RTT and SIG**

The investments being made by the U.S. Department of Education in Race to the Top and School Improvement Grants are unprecedented in scope and scale. To advance comprehensive and coherent education reforms across districts for the purpose of improving student outcomes, Congress appropriated \$4 billion in American Recovery and Reinvestment Act of 2009 (ARRA) funding for the main RTT grant competition to encourage and reward states already implementing significant education reforms in four priority areas: (1) standards and assessments; (2) data systems; (3) effective teachers and school leaders; and (4) turning around persistently low-performing schools. RTT grants were awarded competitively in two phases.

Phase I awards were announced in March 2010 to Tennessee (\$500 million) and Delaware (\$100 million). Phase II awards were made in August 2010 to New York (\$700 million); Florida (\$700 million); Georgia (\$400 million); North Carolina (\$400 million); Ohio (\$400 million); Massachusetts (\$250 million); Maryland (\$250 million); Rhode Island (\$75 million); Hawaii (\$75 million); and the District of Columbia (\$75 million).<sup>1</sup>

The SIG program was funded in fiscal year 2009 with \$546.6 million, and received an additional \$3 billion from ARRA (Pub. L. 111-5). SIG funds go to states based on their share of Title I funding; states then distribute the funds to districts with the lowest-achieving Title I schools that demonstrate need and a strong commitment to implement one of four models—turnaround, restart, closure, and transformation—aimed at improving or closing these persistently low-performing schools.

Given the scale and scope of these federal investments, findings from the RTT-SIG evaluation will be highly anticipated and critically scrutinized by a broad audience of policymakers, educators, and other interested parties. These constituents will want to know whether these programs accomplished their goals: Are struggling schools initiating reforms? Are states improving their data systems? Are common standards and assessments being adopted? Are teachers and principals being supported in their attempts to turn around lowest-achieving schools? In addition to these and other questions of program implementation, there is the bottom-line question of whether these reforms affect students' academic achievement and progress beyond high school.

Legislative authorization for the RTT-SIG evaluation is found in the Education Science Reform Act of 2002, Part D, Section 171(b)(2), which authorizes IES to "conduct evaluations of Federal education programs administered by the Secretary (and as time and resources allow, other education programs) to determine the impact of such programs (especially on student academic achievement in the core academic areas of reading, mathematics, and science)."

## **b. Research Questions**

The RTT-SIG evaluation will examine the following research questions:

- How are RTT and SIG implemented at the state, district, and school levels?
- Are RTT reforms related to improvement in student outcomes?
- Does receipt of RTT and/or SIG funding to implement a school turnaround model have an impact on outcomes for lowest-achieving schools?

---

<sup>1</sup> A third round of RTT grants will be awarded using funds appropriated in 2011. The fiscal year 2011 appropriation act also authorizes the Secretary to use RTT funds for grants to States for improving early childhood care and education.

- Is implementation of the four school turnaround models, and strategies within those models, related to improvement in outcomes for lowest-achieving schools?

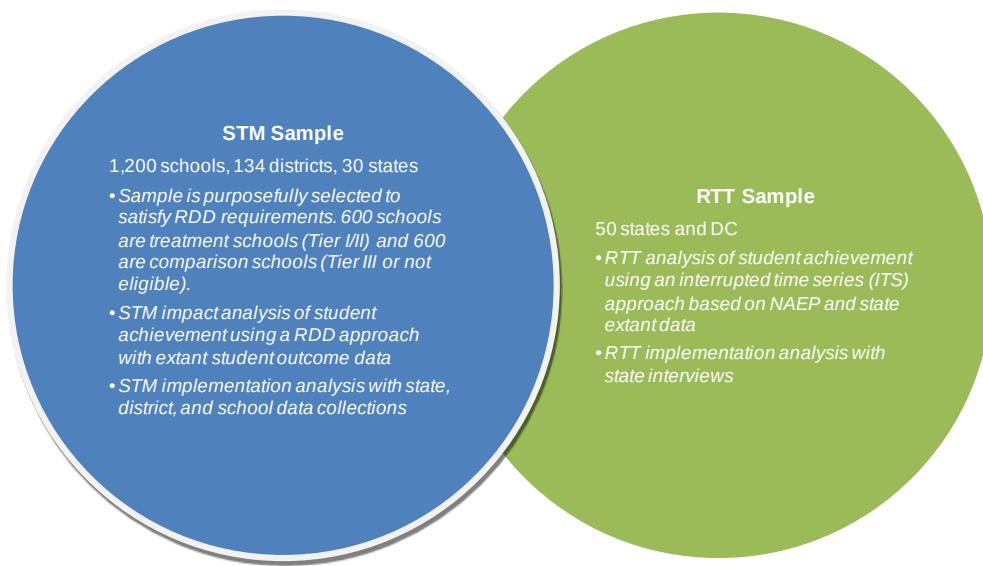
### **c. Study Design**

The RTT-SIG evaluation is designed to provide a descriptive account of the implementation of RTT and SIG; the most rigorous possible estimates of the effects of RTT and SIG; and the contextual information needed to fully understand and interpret those effects. The study will be based on two samples, strategically selected both to provide information on RTT and SIG implementation and to support a rigorous analysis of program impacts (see Figure A.1). To assess the relationship between RTT and student outcomes, the evaluation will use an ITS analysis. All 50 states and the District of Columbia will be included in the sample for the evaluation of RTT (referred to throughout as the RTT sample).

Separately, to estimate the impact of STMs on student achievement, we plan to use a rigorous RDD, exploiting approaches that involve a continuous measure for awarding STM funds (through RTT or SIG) to schools. The evaluation will also assess the correlation between STMs—and the specific turnaround strategies used within such models—and improvements in school outcomes. The sample for the evaluation of STMs (referred to throughout as the STM sample) will consist of about 1,200 schools from an estimated 134 school districts across 30 states (roughly 600 schools will form the treatment group and 600 the comparison group).<sup>2</sup> The districts in the STM sample will be purposefully selected to maximize the statistical precision of the RDD impacts.

---

<sup>2</sup> Because of ED’s interest in the effects of the “restart” school turnaround model, the STM sample will also include the approximately 30 schools implementing that model.

**Figure A.1. Diagram of Study Samples**

- **Research question 1: How are RTT and SIG implemented at the state, district, and school levels?**

The implementation component of the evaluation will gather information to answer this research question, support answering research question 4, and help with the interpretation of research questions 2 and 3.

**RTT Implementation.** From interviews with state representatives from all 50 states and the District of Columbia, we will learn about reforms being implemented in the areas of focus for RTT, such as the steps states are taking to develop standards for college and career preparedness, to improve data systems, to promote an equitable distribution of effective teachers, and to support school turnaround. The interview data collected from states that won RTT grants will provide information on the reforms that were implemented by RTT grantees, while the information from states that did not win RTT grants will aid in the interpretation of impacts by providing a comparison condition.

**SIG Implementation.** From interviews with state representatives from all 50 states and the District of Columbia and with district representatives in the STM sample, we will learn about the school turnaround efforts that have been implemented in districts and the nature and type of supports provided by districts to turnaround schools. Through surveys administered to school administrators in the STM sample, we will learn about the STM models and the specific improvement strategies that are being implemented in these schools.

The evaluation will use several strategies to ensure that the implementation data collected through these activities are comparable and analyzed in a systematic way. A uniform protocol will be used for each data

collection activity. We will also prioritize the use of closed-ended questions in the data collection instruments to ensure we capture quantitative data on the percentage of states, districts, and schools that are implementing particular RTT and SIG reforms. For the open-ended questions, we plan to use Atlas.ti or NVivo software to help organize the qualitative information gathered into themes. This information will be further summarized through the use of indicator or categorical variables amenable to quantitative analysis. This approach will permit the study team to objectively and systematically describe the implementation of RTT and SIG and examine the relationship between patterns in outcomes and key implementation variables.

- **Research question 2: Are RTT reforms related to improvement in student outcomes?**

We will use a quasi-experimental ITS design to assess how student outcomes change following the receipt of RTT grants. The ITS design will take advantage of the timing of RTT grants while also accounting for the application scores used to select RTT winners in the second phase of the competition. The ITS model *projects* the outcomes that would have been expected in the absence of RTT funding and compares the projections with the pattern of outcomes *actually observed* in the post-intervention period. The effect of the intervention is estimated as the difference between the *predicted* pattern of outcomes and the *actual* trend in outcomes in the post-intervention period. This approach estimates the average effect of RTT for states on the cusp of receiving RTT grants.

Importantly, this approach cannot establish causal relationships between the reforms implemented and estimated changes in student outcomes. Thus, appropriate caution must be used when interpreting these results.

The primary data source for this analysis will be state-level National Assessment of Educational Progress (NAEP) scores.

- **Research question 3: Does receipt of RTT and/or SIG funding to implement a school turnaround model have an impact on outcomes for lowest-achieving schools?**

We plan to estimate the impact of STMs using multiple RDD opportunities, each of which we refer to as a “mini-study.” The potential RDD opportunities are all based on the SIG eligibility tier structure. The first two of the three eligibility tiers are defined, in part, by cutoff values on two continuous variables—the fifth percentile of school-level achievement and a graduation rate of 60 percent. Because these first two tiers must be served before the third tier, a substantially higher proportion of schools in the first two tiers receive STM funding than schools in the third tier. This means that there is a discontinuity in the proportion of schools receiving STM funding at those cutoff values, which creates an opportunity to estimate RDD impacts where schools below these cutoffs will form the treatment group (about 600

schools) and schools above the cutoffs will form the comparison group (about 600 schools).

The ongoing recruitment phase of this study will determine whether or not it will be feasible to implement the RDD. Feasibility depends on (1) the “fuzziness” (Trochim 1984; Hahn et al. 2001) of the RDD and (2) the anticipated statistical precision of RDD impacts. The fuzziness of the RDD refers to the difference between the treatment and comparison groups near the RDD cutoff value in the proportion of schools receiving funds to implement an STM (the smaller this difference, the fuzzier the design). RDD fuzziness needs to be considered when assessing the feasibility of the RDD because it leads to finite sample bias when estimating the complier average causal effect (CACE). Using simulations, we have determined that the difference between the treatment and comparison groups near the cutoff in the proportion of schools receiving STM funds needs to be at least 40 percent. If we are unable to find any RDD opportunities that meet this requirement, then the RDD will not be feasible. However, recruiting efforts so far indicate that it is highly likely that we will find RDD opportunities that meet this feasibility requirement.

The anticipated statistical precision of the RDD impacts is primarily affected by sample size and RDD fuzziness. The original goal for this evaluation (as stated in the performance work statement) was to have sufficient statistical precision to detect an impact of at least 0.10 standard deviations with 80 percent probability. If we fall far short of that goal (for example, if the anticipated minimum detectable effect size is 0.20 standard deviations), we will consult members of the technical working group about whether accepting a loss in internal validity in order to increase statistical power is an acceptable tradeoff. We do not yet have enough information to estimate the study’s overall statistical power. If we find acceptable levels of both statistical precision and fuzziness, we will deem the RDD feasible to implement and will plan to move ahead with it.

Assuming that an RDD is feasible, for every RDD mini-study (that is, unique combination of RDD assignment variable and cutoff, outcome, and grade), we will conduct a full RDD analysis aligned with What Works Clearinghouse evidence standards. Specifically, we will estimate impacts within an optimal bandwidth around the assignment variable’s cutoff value and conduct a full set of diagnostic analyses to assess the performance of the RDD. The overall impact of STMs will be a weighted average of the mini-study impacts, where the weight is the inverse of the variance of the mini-study impacts.

If an RDD is not feasible, we will use an ITS design to assess how student outcomes change following the implementation of an STM in every school in our sample that implements one. The ITS design will take advantage of the timing of STM implementation. The ITS model projects the outcomes that would have been expected in the absence of the STM and compares the

projections with the pattern of outcomes actually observed in the post-intervention period. The effect of STM implementation is estimated as the difference between the predicted pattern of outcomes and the actual outcomes observed in the post-intervention period. To strengthen the validity of our estimates and to increase statistical power, our ITS design will also incorporate a comparison group of schools that did not receive STM funding. This will be accomplished by using the difference in outcomes between schools that do and do not receive STM funding as the outcome in the ITS analysis. The sample of schools used to estimate impacts with this ITS design would be based on the sample identified for estimating RDD impacts, but augmented to reduce any observed differences between our analysis sample and the national population of SIG grantees.

Student-level data used to answer research question 3 will come from two extant sources: (1) state-level data systems whenever possible; and (2) when we cannot obtain the necessary data from the states in which districts are located, from school district data systems. The outcomes of interest for this study are student standardized test scores (state assessments) from the 2011-2012, 2012-2013, and 2013-2014 school years; high school graduation and attendance rates; and (to the extent that data are available) college enrollment rates and completion of at least a year of college credit.

- **Research question 4: Is implementation of the four school turnaround models, and strategies within those models, related to improvement in outcomes for lowest-achieving schools?**

To examine the correlation between improvements in school outcomes and specific turnaround strategies, we will draw on the implementation data collected from schools implementing an STM. We will conduct an ITS analysis within each school and then correlate the ITS outcomes with school-level turnaround models or strategies.

As with the ITS design described for research question 3, this correlational analysis cannot establish a causal relationship between turnaround models/strategies and estimated changes in school outcomes. Thus, caution must be used when interpreting these results, because specific turnaround models or strategies may not have caused the observed changes in outcomes.

Outcomes to be examined include student achievement on state assessments, high school graduation rates, attendance, and (to the extent data are available) college enrollment rates and rates of completion of at least a year's worth of college credit. Data sources for turnaround models, strategies, and practices include state and district interviews and school surveys.

## d. Data Collection Needs

This study includes several data collection efforts that are described below and summarized in Table A.1.

**Table A.1. Data Collection Needs**

Data source	Sample	Respondent	Mode	Schedule
State Interview	Both RTT and STM Samples	Representatives from 50 states and the District of Columbia	Telephone interview	Spring 2012 Spring 2013 Spring 2014
District Interview	STM Sample	134 district administrators	Telephone interview	Spring 2012 Spring 2013 Spring 2014
School Survey	STM Sample	1,200 school administrators	Web with email, hard-copy, and telephone follow-up	Spring 2012 Spring 2013 Spring 2014
Administrative Data on Student Outcomes	Both RTT and STM Samples	State or district staff	Electronic or hard copy	Fall 2012 Fall 2013 Fall 2014
NAEP Scores	RTT Sample	N/A	Publicly available	Spring 2012 Spring 2014

N/A = not applicable; NAEP = National Assessment of Educational Progress.

**State Interview.** In the spring of 2012, 2013, and 2014, we will conduct semistructured telephone interviews with representatives from the state education agency in every state and the District of Columbia (Appendix A). The interviews will consist of topic-specific modules that may be administered to different state-level respondents. States in the RTT sample that did not receive RTT grants will be asked about their implementation of RTT-related reforms. All states will be asked detailed questions about their policies and supports for school turnaround.

**District Interview.** In the spring of 2012, 2013, and 2014, we will conduct semistructured telephone interviews with district-level administrators from each district included in the STM impact study (about 134 districts) (Appendix B). These interviews will document school turnaround efforts and supports provided by the district to turnaround schools. Like the state interview, the district interview will consist of topic-specific modules that may be administered to different district-level respondents.

**School Survey.** In the spring of 2012, 2013, and 2014, we will conduct a web survey of school administrators (principals, assistant principals, or other staff knowledgeable about school turnaround activities) at the approximately 1,200 schools that are part of the STM sample (Appendix C). We will contact administrators by email or (if email is unavailable or invalid) cover letter. 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 need for 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 either providing answers on the telephone or completing a hard-copy version. To ease burden on respondents, we will limit the length of the survey to 45 minutes. Because the information we need to obtain from schools is considerable, items on the instrument capture specific areas of interest through closed-ended questions and offer specific and mutually exclusive response options.

**Administrative Data on Student Outcomes.** In the fall of 2012, 2013, and 2014, we will request standardized test scores on state proficiency assessments; high school graduation rates; attendance; and (to the extent data are available) college enrollment rates and completion of at least a year of college credit. In addition to test scores, we will request that the state (or district if necessary) provide data on student characteristics such as sex, race/ethnicity, birth year, grade, eligibility for free or reduced-price lunch, and English language learner status. Student-level data will be collected for the STM impact analysis only; the RTT outcome analysis will rely on administrative data aggregated to the state, district, or school levels. We are assuming that we will be able to obtain the necessary student-level outcome data directly from state data systems for some states, but that we will need to obtain student-level data from districts in other states. We will develop two forms to collect outcomes data—one for states and one for districts—to show states (and districts, when necessary) the data we need (appendices D and E).

**NAEP Scores.** In the spring of 2012 and 2014, we will obtain public-use aggregate state-level NAEP scores from ED. The NAEP scores are available for grades 4 and 8, for both math and reading, every other year. In 2001, participation in state NAEP tests was made mandatory for states receiving Title I funds. Thus, we plan to use at least 4 years of data prior to RTT grants (2003, 2005, 2007, 2009) and two years of post-RTT data (2011 and 2013, which we anticipate will become available in spring 2012 and spring 2014).

#### **e. Study Activities and Data Collection Timeline**

This clearance request pertains to the collection of data through interviews with state representatives (Appendix A), interviews with district representatives (Appendix B), a survey of school administrators (Appendix C), and collection of administrative data from states and districts (appendices D and E). The RTT-SIG evaluation is expected to be completed in five years, with three years of data collection. Table A.2 shows the schedule of data collection activities.

**Table A.2. Study Timetable**

Activity	Date
	<b>2011</b>
Solidify Participation of Study Sample	6/2011 through 1/2012
	<b>2012</b>
Collect Interview Data	3/2012 through 6/2012
Collect Survey Data	3/2012 through 6/2012
Collect Administrative and NAEP Data	7/2012 through 10/2012
	<b>2013</b>
Collect Interview Data	3/2013 through 6/2013
Collect Survey Data	3/2013 through 6/2013
Collect Administrative Data	7/2013 through 10/2013
	<b>2014</b>
Collect Interview Data	3/2014 through 6/2014
Collect Survey Data	3/2014 through 6/2014
Collect Administrative and NAEP Data	7/2014 through 10/2014

NAEP = National Assessment of Educational Progress.

## 2. Purposes and Uses of the Data

To address the study's research questions, the evaluation will collect and analyze data from several sources. Table A.3 lists the questions and the data sources that will be used to answer them. We describe the study's use of each data source in more detail below. Information will be collected by Mathematica Policy Research and its partners AIR and SPR, under contract with ED [contract number ED-IES-10-C-0077].

**Table A.3. Research Questions and Data Sources**

Research Question	Data Source(s)
1. How are RTT and SIG implemented at the state, district, and school levels?	School surveys State and district interviews
2. Are RTT reforms related to improvement in student outcomes?	NAEP data Aggregated state extant data
3. Does receipt of RTT and/or SIG funding to implement a school turnaround model have an impact on outcomes for lowest-achieving schools?	State and district extant data
4. Is implementation of the four school turnaround models, and strategies within those models, related to improvement in outcomes for lowest-achieving schools?	State and district extant data School surveys State and district interviews

NAEP = National Assessment of Educational Progress.

**State Interviews.** These interviews will focus on RTT policies and practices at the state level, as well as state policies and practices designed to support school turnaround through RTT and SIG. We will use this information primarily to examine research question 1, but perhaps also to examine whether impacts of STM vary with respect to these implementation details (research question 4). States in the RTT sample that did not receive RTT grants will be asked about their implementation of RTT-related reforms.

**District Interviews.** Interviews with districts in the STM sample will focus on how state and district STM policies play out in districts and schools, including documenting the STM supports and information received by the districts from the states. We will use the information from the interviews with districts in the STM sample to examine the implementation of SIG (research question 1) and whether impacts of STM vary with respect to these implementation details (research question 4).

**School Surveys.** These surveys will focus on implementation of STM in schools and the STM-related supports, information, and policies rolled out by the state and district. We will use this information to examine SIG implementation (research question 1) and whether impacts of STM vary with respect to these implementation details (research question 4).

**State and District Extant Data.** The data for the outcomes analyses will come from student-level administrative data maintained by states and districts, as well as from the NAEP data the study team obtains from ED. (Student-level data will be collected for the STM impact analysis only; the RTT outcomes analysis will rely on administrative data aggregated to the state, district, or school levels.) The outcomes of interest for this study are student standardized test scores (state assessments and NAEP); high school graduation and attendance rates; and (to the extent data are available) college enrollment rates and completion of at least a year of college credit.

### **3. Use of Technology to Reduce Burden**

The data collection plan is designed to obtain information in an efficient way that minimizes respondent burden. When feasible, we will gather information from existing data sources, using the most efficient methods available. Existing data sources will be computer files provided by states (or from districts when not available from the state) that include test scores for school-administered tests and student demographic information. Mathematica will work with state and district personnel to determine the most efficient and least burdensome procedures for their staff, and will capitalize on any electronic systems already in place. Whenever possible, states and districts will be able simply to upload or enter data into any electronic spreadsheet, or to an equivalent file, and transfer it to Mathematica through a secure FTP site. Regardless of the form in which it is received, these data will be converted into a consistent format so that they can be combined with data submitted by other states or districts and are

suitable for analysis. If it is too burdensome or not possible for a state or district to provide data in electronic format, we will provide clear instructions on how to submit the information in hard-copy form, to be coded by the study team.

We will interview state and district representatives by telephone, which will allow us flexibility to schedule the interviews at their convenience and to separate topical modules from a particular instrument for a given respondent. This mode of data collection is also appropriate for the conversational exchange necessary to obtain answers to any open-ended questions, and to allow probing for more detail than in a self-administered survey.

A web-based survey will be the primary mode of data collection for school administrators 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 level of response errors.

#### **4. Efforts to Avoid Duplication of Effort**

Because no other national study has been conducted or is under way to address the same research questions as this one, ED determined that an in-depth national study examining the implementation and impacts of RTT and SIG is needed. To minimize burden on participants and avoid duplication of effort, ED will coordinate efforts between this evaluation and several other ongoing studies of ARRA, including the Integrated Evaluation of ARRA Funding, Implementation, and Outcomes (IEA) and the Study of School Turnaround (SST).

With regard to the IEA, some overlap among respondents is inevitable, since the ARRA evaluation is collecting data from state officials from all 50 states and from a nationally representative sample of district administrators and school principals. However, the topics of data collection and the data collection strategies are notably different. For example, the IEA will administer a closed-ended survey to state officials. While the state-level interview for the RTT-SIG impact study addresses some of the same broad topics covered in the IEA survey, the RTT-SIG interviews will probe deeper than a survey can. In addition, the study team for the RTT-SIG evaluation has compared data collection instruments to those from IEA and has deleted any duplicative questions from the study instruments.

With respect to the SST, we anticipate little overlap among respondents for these two studies for two reasons:

- The SST began data collection in the spring of 2011 and will continue through 2013, while the RTT-SIG evaluation seeks to begin data collection in spring of 2012, thus reducing simultaneous data collection efforts by one year.
- The SST has a small sample of 60 schools in six states. Given this sample size, the likelihood of overlap is relatively small.
- There is limited overlap in respondent groups, since much of the SST focuses on collecting data from teachers, students, parents, union representatives, school improvement teams, instructional coaches, and external support providers. None of these respondent groups are part of the impact evaluation of RTT-SIG data collection plans. In the few cases where there is overlap, the study teams will investigate the feasibility of conducting joint interviews. That is, a researcher from one study team could conduct the interview while a representative from the other study team listens, adding questions only as necessary to address study requirements. The teams will explore this option once the extent of sample overlap is known, keeping in mind that the focus of the two studies differs in important ways (with the SST focusing on the change process and how reforms are implemented over time, while the RTT-SIG evaluation focuses on documenting the reforms that were implemented, including RTT reforms, which are not a focus of the SST).

Whenever possible, the evaluation contractor will use existing data from EDFacts; state SIG and RTT grant and subgrant applications; Consolidated State Performance Reports; Office of Elementary and Secondary Education monitoring reports; and federal, state, and local administrative files. This will further reduce respondent burden and minimize duplication of data collection efforts.

## 5. Methods to Minimize Burden on Small Entities

No small businesses or entities will be involved as respondents.

## 6. Consequences of Not Collecting Data

The data collection plan described in this submission is necessary for ED to conduct an evaluation of RTT and SIG and determine whether its investment has improved student outcomes. Moreover, RTT and SIG together represent one of the largest investments in education reform in American history (about \$7 billion); failing to conduct this evaluation would mean missing the opportunity to learn lessons relevant to future education reform efforts.

The consequences of not collecting specific data are outlined below:

- Without the information from the **state interviews**, we will be unable to examine the implementation of RTT policies and practices at the state level and practices designed to support school turnaround through RTT and SIG.
- Without the information from the **district interviews**, we will be unable to understand how state and district STM policies play out in districts and schools or to document the STM supports and information districts receive from states.
- Without the **school surveys**, we will not capture school-level implementation of STM policies. The surveys also serve as a key data source for examining whether impacts of STM vary with respect to implementation details.
- Without **state and district administrative records**, we will not be able to analyze the impact of receipt of RTT and/or SIG funding on student outcomes.

## 7. Special Circumstances

There are no special circumstances involved with this data collection.

## 8. **Federal Register Announcement and Consultation**

### a. **Federal Register Announcement**

The 60-day notice to solicit public comments was published in Vol. 76, No. 182, Pg. 58251 of the *Federal Register* on September 20, 2011. No public comments were received

### b. **Consultations Outside the Agency**

In formulating the evaluation design, the study team sought input from the technical working group (TWG), which includes practitioners and experts in evaluation methods and data analysis, state assessment programs, and education reform. We will continue to consult with the TWG throughout the study on other issues that would benefit from their input. Table A.4 lists the TWG members.

**Table A.4. Technical Working Group Members**

Name	Title and Affiliation	Expertise
Thomas Fisher	Fisher Education Consulting	State assessment programs
Brian Jacob	Walter H. Annenberg Professor of Education Policy and Director of the Center on Local, State and Urban Policy at the Gerald R. Ford School of Public Policy, University of Michigan	Education reform, evaluation methods and data analysis
Elizabeth Stuart	Assistant Professor in the Department of Mental Health and the Department of Biostatistics, Johns Hopkins University	Evaluation methods and data analysis
Guido Imbens	Professor of Economics, Harvard University	Evaluation methods and data analysis
Thomas Cook	Joan and Sarepta Harrison Chair in Ethics and Justice Professor of Sociology, Psychology, Education and Social Policy; Faculty Fellow, Institute for Policy Research, Northwestern University	Evaluation methods and data analysis
James Spillane	Spencer T. and Ann W. Olin Chair in Learning and Organizational Change and Professor, School of Education and Social Policy, Northwestern University	Education reform
Jonathan Supovitz	Associate Professor and Director, Consortium for Policy Research in Education, University of Pennsylvania	Education reform
Sean Reardon	Associate Professor, Stanford University	Education reform, evaluation methods and data analysis
Thomas Kane	Professor of Education and Economics, Harvard University	Education reform, evaluation methods and data analysis
Eric Smith	Former Commissioner of Education, State of Florida	Practitioner

### c. **Unresolved Issues**

It is currently unknown whether an RDD will be feasible to address research question 3 (impact of STMs on student outcomes). Ongoing recruitment efforts with states and districts will be used to assess feasibility,

with an expected determination by January 2012. If an RDD is not feasible, an ITS design will be used instead to address research question 3.

## **9. Payments or Gifts**

Burden payments have been proposed for the school administrator survey to partially offset respondents' time and effort in completing the survey. During each round of data collection, we propose a \$30 payment to administrators from comparison-group schools who complete the questionnaire, in acknowledgment of the 45 minutes required. This amount is within the incentive guidelines outlined in the March 22, 2005, memo, "Guidelines for Incentives for NCEE Evaluation Studies," prepared for OMB. The payment is proposed because high response rates are needed to make the survey findings reliable, and we are aware that school administrators are the target of numerous requests to complete surveys on a wide variety of topics from state and district offices, independent researchers, and ED.

## **10. Assurances of Confidentiality**

The study team has established procedures to ensure the confidentiality and security of its data. This approach will be in accordance with all relevant regulations and requirements, in particular the Education Sciences Institute Reform Act of 2002, Title I, Subsection (c) of Section 183, which requires the director of IES to "develop and enforce standards designed to protect the confidentiality of persons in the collection, reporting, and publication of data." The study will also adhere to requirements of Subsection (d) of Section 183 prohibiting disclosure of individually identifiable information as well as making the publishing or inappropriate communication of individually identifiable information by employees or staff a felony.

The study team will protect the full privacy and confidentiality of all individuals who provide data. The study will not have data associated with personally identifiable information (PII), as study staff will be assigning random ID numbers to all data records and then stripping any PII from the data records. In addition to the data safeguards described here, the study team will ensure that no respondent names, schools, or districts are identified in publicly available reports or findings, and if necessary, the study team will mask distinguishing characteristics. A statement to this effect will be included with all requests for data:

"Mathematica Policy Research and its subcontractors AIR and SPR follow the confidentiality and data protection requirements of IES (The Education Sciences Reform Act of 2002, Title I, Part E, Section 183). **Responses to this data collection will be used only for research purposes.** The reports prepared for the study will summarize findings across the sample and will not associate responses with a specific district, school, or individual. We will not provide information that identifies respondents to anyone outside the study team, except as required by law."

Mathematica employs the following safeguards to ensure confidentiality:

- All Mathematica employees sign a pledge that emphasizes the importance of confidentiality and describes their obligation.
- Secure FTP services allow encrypted transfer of large data files with clients, if necessary. Internal networks are all protected from unauthorized access utilizing defense-in-depth best practices, which incorporate firewalls and intrusion detection and prevention systems. The networks are configured so that each user has a tailored set of rights, granted by the network administrator, to files approved for access and stored on the LAN. Access to hard-copy documents is strictly limited. Documents are stored in locked files and cabinets. Discarded materials are shredded.
- Computer data files are protected with passwords, and access is limited to specific users, who must change their passwords on a regular basis and conform to strong password policies.
- Especially sensitive data are maintained on removable storage devices that are kept physically secure when not in use.

After the study concludes, the study data will be transmitted to the National Center for Education Statistics (NCES) for safekeeping as a restricted-use file. All other versions of the data will be destroyed. Prior to transmittal, the data will undergo careful screening, and modification if necessary, to ensure that there is no unacceptably high level of disclosure risk for protected respondents. Researchers wishing to access the data for secondary analysis must apply for an NCES license and agree to the applicable rules and procedures guiding the use of restricted-use files.

## **11. Additional Justification for Sensitive Questions**

No questions of a sensitive nature will be included in this study.

## **12. Estimates of Hours Burden**

Representatives from all 50 states and the District of Columbia, officials from 134 school districts, and administrators from 1,200 schools will be asked to participate in data collection activities over three years. These include student record collection, interviews with state and district representatives, and surveys of school administrators.

We estimate that the total number of state, district, and school respondents involved in the three years of data collection activities will be 1,526 and the total number of burden hours will be 7,376, for an average of 4.83 hours per respondent (Table A.5). The data collection instruments are contained in appendices A, B, C, D, and E, and will include the following appropriately tailored Public Burden Statement:

According to the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless such collection displays a valid OMB control number. Public reporting burden for this collection of information is estimated to average XX minutes/hours per response, including time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. The obligation to respond to this collection is required to obtain or retain a benefit (The Education Department General Administrative Regulations, 34 C.F.R. § 76.591). Send comments regarding the burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to the U.S. Department of Education, Washington, DC 20202-4651 and reference the OMB Control Number 1850-0884. If you have comments or concerns regarding the status of your individual submission of this form, write directly to: Institute of Education Sciences, U.S. Department of Education, 555 New Jersey Ave. NW, Washington, DC 20208.

### **13. Estimates of Cost Burden to Respondents**

There are no additional respondent costs associated with this data collection beyond the burden estimated in item A12.

### **14. Estimates of Annual Costs to the Federal Government**

The estimated cost for this five-year study, including development of a detailed study design, data collection instruments, justification package, recruitment, data collection, data analysis, and report preparation, is \$18,606,416 or approximately \$3,721,283 per year.

### **15. Reasons for Program Changes or Adjustments**

There is an overall annual decrease in burden hours of 2,303 hours. This program change is a result of the burden hours from this second phase (the data collection phase) being added and the burden hours from the first phase (the recruitment phase) being eliminated since they will be completed by the time this second phase is approved.

### **16. Plan for Tabulation and Publication of Results**

#### **a. Tabulation Plans**

Our tabulation plans include four sets of analyses aligned to the research questions.

**Table A.5. Burden in Hours to Respondents**

Respondents/ Activity	Total Number of Respondent s	Number of Responses per Responden t	Number of Administratio ns	Total Number of Response s	Average Burden Hours per Response	Total Burden (Hours)
<b>States</b>						
State Interview (Appendix A)	51	1	3	153	4.5	689
Student and School Records Request (Appendix D)	51	1	3	153	8.0	1,224
<b>Total State Staff</b>	<b>102</b>			<b>306</b>		<b>1,913</b>
<b>Districts</b>						
District Interview (Appendix B)	134	1	3	402	1.5	603
Student Records Request (Appendix E)	90	1	3	270	8.0	2,160
<b>Total District Staff</b>	<b>224</b>			<b>672</b>		<b>2,763</b>
<b>Schools</b>						
School Survey (Appendix C)	1,200	1	3	3,600	0.75	2,700
<b>Total School Staff</b>	<b>1,200</b>			<b>3,600</b>		<b>2,700</b>
<b>Overall Total</b>	<b>1,526</b>			<b>4,578</b>		<b>7,376</b>

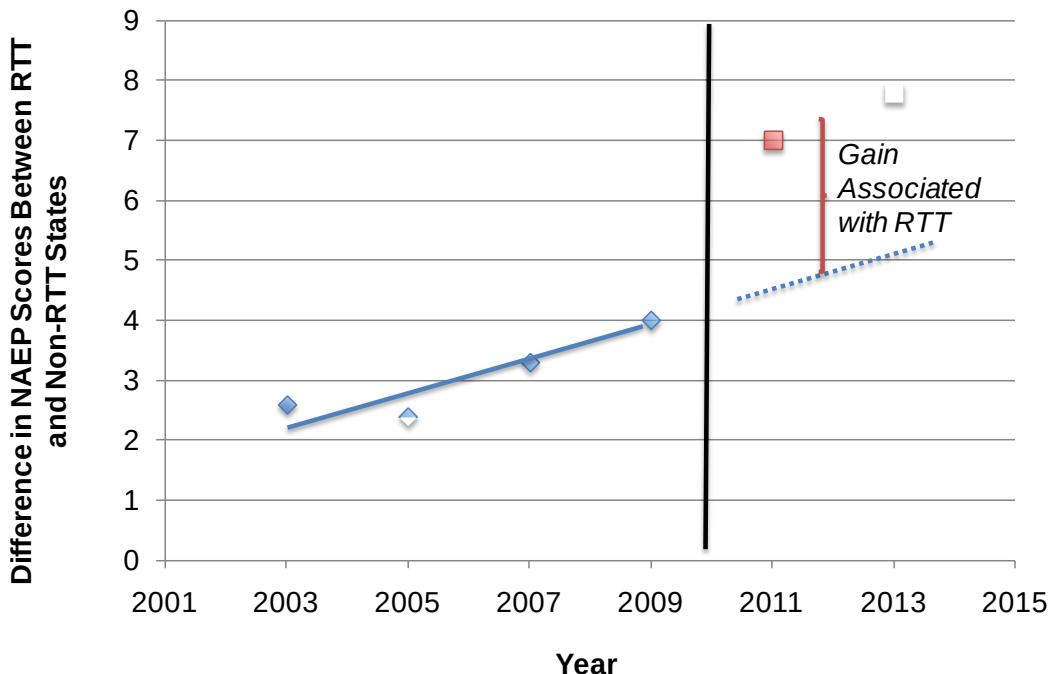
**Implementation Analysis.** To thoroughly document the extent to which states, districts, and schools have implemented RTT and SIG systems and requirements, we will use data collected through interviews with state and district representatives and surveys of school administrators. We will use descriptive analyses to report observed patterns in the data. We will also describe implementation by key groups at different levels. For the RTT sample, we will report findings separately for RTT states and non-RTT states. For the STM sample at the district level, we will report what district representatives recount in response to questions about the treatment and comparison schools in their districts. For the STM sample at the school level, we will report findings separately for schools that received STM funding and for schools that did not. We will use the data to compare responses to questions about implementation from year to year, and implementation analyses to help interpret the impacts of STM by describing the implementation of STMs in the treatment group and the reform experiences of schools in the comparison group.

Because we plan to estimate the impacts of STMs using an RDD (which generates impact estimates that apply primarily to schools at the cutoff value of an assignment variable), it will also be important to describe the difference in reform experiences between the treatment and comparison groups near the cutoff value of the RDD assignment variable. In calculating the difference in treatment and comparison group reform experiences at the cutoff value of the assignment variable, we will use the same analytic techniques for calculating outcome differences. The data source for these analyses will be the school administrator survey. In comparing the average experiences of the full treatment and comparison groups, we will draw on the school administrator survey and interviews with district representatives.

**RTT Outcomes Analysis.** We will use an ITS design to assess how student outcomes change following the receipt of RTT grants. The ITS design will take advantage of the timing of RTT grants. The ITS model *projects* the outcomes that would have been expected in the absence of RTT funding and compares the projections with the pattern of outcomes *actually observed* in the post-intervention period. The effect of the intervention is estimated as the difference between the *predicted* pattern of outcomes and the *actual* trend in outcomes in the post-intervention period.

To strengthen the validity of our estimates and to increase statistical power, our ITS design will also incorporate a comparison group of states that did not win RTT funding. This will be accomplished by using the estimated *difference* in outcomes between RTT winners and losers who were on the cusp of winning or losing RTT as the outcome in the ITS analysis. In forming this comparison group, we will take advantage of the application scores used to select RTT winners in the second phase of the competition. The cutoff value on this score is the lowest application score received by one of the 10 winning Phase II applicants.

The ITS approach is illustrated graphically in Figure A.2. Figure A.2 shows the estimated difference in NAEP scores between states that just won RTT in Phase II and states that just lost RTT in 2003, 2005, 2007, and 2009, with the cutoff year between receiving and not receiving RTT funds taken to be 2010. We focus on Phase II winners because Phase I winners lack a Phase II application score, which is needed to adjust for the application score when calculating the difference between RTT and non-RTT outcomes in each year. The solid line shows the (linear) trend in outcomes estimated on the basis of the period before the intervention, which is then extended to the post-intervention period (dashed line). The gains associated with RTT in the post-RTT years are estimated as the average deviation from this projected trend. Actual outcomes in the post-RTT years are shown by the squares in the figure.

**Figure A.2. Illustration of the ITS Design**

For our benchmark ITS analysis, we will use an approach consistent with Figure A.2. Specifically, we will estimate a linear trend using the difference in NAEP scores<sup>3</sup> from 2003, 2005, 2007, and 2009 between the 10 Phase II RTT grantees and 12 non-RTT states in our sample. We will then analyze outcomes by examining the difference between that trend and the observed difference between those two groups of states in 2011 and 2013.

Estimation involves a two-step procedure. In the first step, we estimate the average difference in outcomes between states that just won RTT and states that just lost RTT in each year. We will estimate this difference using a simplified RDD approach<sup>4</sup> in each year using equations 1 and 2.

$$(1) \quad Y_i^R = \beta_0^R + \beta_1^R X_i^R + \beta_2^R Z_i^R + \varepsilon_i^R$$

$$(2) \quad Y_i^L = \beta_0^L + \beta_1^L X_i^L + \beta_2^L Z_i^L + \varepsilon_i^L$$

<sup>3</sup> We frame this discussion in terms of NAEP scores, but the same methods will be used for other outcomes.

<sup>4</sup> Due to the relatively small number of states involved in this estimation and the potential for subjectivity in scoring RTT applications, this RDD approach will not have the same level of rigor as that used to estimate impacts of STMs. For example, we do not anticipate that findings based on this RDD analysis would meet WWC standards. However, even a simplified RDD approach will provide a more valid comparison than an approach that compares the average RTT state to the average non-RTT state, since the RDD approach does at least adjust for the RTT application score.

Equations 1 and 2 will be estimated separately in each year. The superscripts  $R$  and  $L$  denote the right (treatment) and left (comparison) sides of the RDD cutoff value,  $Y_i$  is the outcome for state  $i$ ,  $X_i$  is the RTT application score centered at the cutoff value,  $Z_i$  is a set of mean-centered pre-RTT covariates, and  $\varepsilon_i$  is the error term. The interpretation of the constant term in a regression is the expected mean outcome when all covariates equal zero. Thus, the assignment variable is centered at the RDD cutoff value so that the intercept terms in equations 1 and 2 represent the predicted value of the outcome variable at the cutoff value. Thus, the RDD-adjusted difference in outcomes between RTT and non-RTT states is estimated by the difference in intercept terms:  $\Delta Y_t = \beta_0^R - \beta_0^L$ .

In the second step,  $\Delta Y_t$  becomes the outcome in an ITS estimation equation:

$$3) \quad \Delta Y_t = \alpha + \beta[t(1 - RTT_t)] + \delta \cdot RTT_t + \varepsilon_t$$

(

where  $\Delta Y_t$  is the difference between RTT and non-RTT states in NAEP scores (in a particular subject and grade) in year  $t$ . The term  $\alpha$  is a constant,  $\beta$  is a linear trend, and  $RTT_t$  is an indicator of whether year  $t$  is after RTT funding begins. The time variable,  $t$ , will be centered at 2010 (the year RTT grantees began receiving their funding). This regression does not include additional covariates due to a lack of available degrees of freedom. Because  $\Delta Y_t$  is an estimate with potentially varying precision across years, we will estimate equation (3) using inverse variance weights.

The outcome gains associated with RTT are  $\delta - 2 \cdot \beta$ . As discussed, the gains associated with RTT are the deviation relative to the pre-existing trend. The reason we subtract  $2 \cdot \beta$  is to account for the rise in the projected trend line over the two years between 2010 and 2012 (2012 is the average of 2011 and 2013, the two post-RTT data points whose average is estimated by  $\delta$ ). We will also calculate outcome gains separately for 2011 and 2013 by replacing the single RTT indicator with two, one for each post year (for the study's first report with outcome findings, only 2011 data will be available).

**STM Impact Analysis.** Unless it is not feasible, we plan to use an RDD to estimate the impact of STM funding on student outcomes. The rules from ED about the prioritization of state STM funds to the persistently lowest-achieving schools create the opportunity for an RDD, generally considered one of the strongest quasi-experimental designs (see, for example, Shadish et al. 2002). Student-level data used for this analysis will come from two sources: (1) the school districts recruited for the study, and (2) extant data

from the data systems of states. The recruited school districts will include about 600 treatment group and 600 comparison group schools.

The RDD component of this study can be characterized as a set of many mini-studies, each corresponding to a specific combination of state, outcome, and grade level. For some mini-studies, schools could be assigned using two assignment variables (average achievement and graduation rate), and we will estimate separate impacts for each assignment variable. We will conduct a separate RDD analysis for each mini-study, because the relationship between the outcome and the assignment variable could vary across mini-studies, and estimating that relationship accurately is essential for obtaining unbiased impacts.

For each mini-study, we plan to estimate intent to treat (ITT) impacts and the CACE. The ITT impact is the impact of being below the cutoff value on the assignment variable (the impact of being eligible for STM funding). Because not all schools below the cutoff value will actually receive STM funding (preliminary calculations reveal that about 70 percent of schools that would be in our treatment group received funding), the ITT impact does not correspond to the impact of being offered, or of receiving, STM funding. Both the impact of being offered STM funding (measured using SIG award information from state websites and reported in Hurlburt et al. 2011) and the impact of actually receiving STM funding (measured using the STM school survey) will be estimated using CACE analysis. Therefore, the CACE impacts are likely to be of greatest policy interest.

The ITT impact estimation equations for the mini-studies, in which the unit of assignment is the school, are:

$$(4) \quad Y_{ij}^R = \beta_0^R + \beta_1^R X_{ij}^R + \beta_2^R Z_{ij}^R + u_j^R + \varepsilon_i^R,$$

$$(5) \quad Y_{ij}^L = \beta_0^L + \beta_1^L X_{ij}^L + \beta_2^L Z_{ij}^L + u_j^L + \varepsilon_i^L,$$

where the superscripts *R* and *L* denote the right and left sides of the RDD cutoff value,  $Y_{ij}$  is the outcome (for example, scores on the state assessment or postsecondary matriculation) for student *i* in school *j*,  $X_{ij}$  is the assignment variable centered at the cutoff value (either school-level achievement or graduation rate, depending on which assignment variable is used in a particular mini-study),  $Z_{ij}$  is a set of mean-centered baseline covariates,  $u_j$  is a school-level error term, and  $\varepsilon_i$  is a student-level error term. The interpretation of the constant term in a regression is the expected mean outcome when all covariates equal zero. Thus, the assignment variable is centered at the RDD cutoff value so that the intercept terms in equations 4 and 5 represent the predicted value of the outcome variable at the cutoff

value. Similarly, the covariates  $(Z_{ij})$  are mean-centered. The RDD impact of STM funding receipt on the outcome is estimated by the difference in intercept terms:  $\delta^{RD} = \beta_0^L - \beta_0^R$ . The baseline covariates  $(Z_{ij})$  are included in this model to increase precision and will vary by state and district depending on data availability.

An RDD in which the difference in the intervention participation rate between the treatment and comparison groups is less than 100 percent is known as a “fuzzy” RDD (Trochim 1984; Hahn et al. 2001). In the context of a fuzzy RDD, it is possible to estimate the impact either of receiving an *offer* of STM funding or of *actually receiving* STM funding by calculating the CACE. To calculate CACE impacts, we will add two estimating equations:

$$(6) \quad P_{ij}^R = \gamma_0^R + \gamma_1^R X_j^R + \gamma_2^R Z_{ij}^R + \nu_j^R + e_i^R,$$

$$(7) \quad P_{ij}^L = \gamma_0^L + \gamma_1^L X_j^L + \gamma_2^L Z_{ij}^L + \nu_j^L + e_i^L,$$

where  $P$  is an indicator of whether a school is offered (or receives) STM funding, and other variables are defined similarly to equations 4 and 5. The impact on being offered (or receiving) STM funding is  $\delta^P = \gamma_0^L - \gamma_0^R$ , and the

$$\text{CACE impact is } \delta^{CACE} = \frac{\delta^{ITT}}{\delta^P}.$$

If it is not feasible to implement an RDD, we will use an ITS design to assess how student outcomes change following the implementation of an STM. The ITS design will take advantage of the timing of STM implementation. The ITS model projects the outcomes that would have been expected in the absence of the STM and compares the projections with the pattern of outcomes actually observed in the post-intervention period. The effect of STM implementation is estimated as the difference between the predicted pattern of outcomes and the actual outcomes observed in the post-intervention period. To strengthen the validity of our estimates and to increase statistical power, our ITS design will also incorporate a comparison group of schools that did not receive STM funding. This will be accomplished by using the difference in outcomes between schools that do and do not receive STM funding as the outcome in the ITS analysis. The sample of schools used to estimate impacts with this ITS design would be based on the sample identified for estimating RDD impacts, but augmented to reduce any observed differences between our analysis sample and the national population of SIG grantees.

**Relating Student Outcome Gains to STMs and Practices.** We will use an ITS design to assess how student outcomes change following the implementation of an STM. The ITS design will take advantage of the timing of STM implementation. The ITS model *projects* the outcomes that would

have been expected in the absence of the STM or practice and compares the projections with the pattern of outcomes *actually observed* in the post-intervention period. The effect of the STM or practice is estimated as the difference between the *predicted* pattern of outcomes and the *actual* outcomes observed in the post-intervention period.

After outcome gains have been estimated for every school in our sample that implemented an STM, we will examine the relationship between those gains and the specific STM and individual practices that each school implemented.

When interpreting findings we will clarify that variation in outcome gains across STMs and practices could be due to unobserved characteristics of schools and cannot necessarily be attributed to the models or practices themselves. This is because the mechanism used to assign STMs and associated practices to schools is unknown, meaning that we cannot adjust for it.

Our approach to estimating the relationship between improvements in student outcomes and specific models or practices involves four steps. First, we will assess which STMs and practices can be analyzed with the available data, creating “bundles” of practices when practices cannot be analyzed individually. Second, we will estimate outcome gains for every grade in each school in our sample that implemented an STM. Third, we will examine the relationship between the estimated school-specific gains and the specific STM and practices implemented in schools. Fourth, we will aggregate these relationships across grades.

The ITS model is shown in equation (8), where:  $Y$  is the outcome (in the case of test scores,  $Y$  is transformed into a z-score<sup>5</sup>);  $t$  is the year (centered at the 2010-2011 school year);  $STM^-$  is a binary variable that equals 1 for years prior to 2010-2011 and 0 otherwise;  $STM^T$  ( $T$  corresponds to outcome year 1, 2, or 3) is a binary variable that equals 1 when  $t = T$  and 0 otherwise;  $\varepsilon$  is an error term; and  $\alpha$ ,  $\delta^-$ ,  $\delta^1$ ,  $\delta^2$ , and  $\delta^3$  are parameters to be estimated.

$$(8) \quad Y_t = \alpha + \delta^- (t \cdot STM_t^-) + \delta^1 \cdot STM_t^1 + \delta^2 \cdot STM_t^2 + \delta^3 \cdot STM_t^3 + \varepsilon_t$$

For an outcome year  $T$ , the outcome gain associated with STM implementation for a given grade in a given school is  $\Delta Y^T = \delta^T - T \cdot \delta^-$ , which is the distance between the outcome in year  $T$  and the trendline projected by the ITS model from the preintervention time period.

---

<sup>5</sup> An individual student's test score is converted into a z-score by subtracting from the student's score the statewide mean score and then dividing by the statewide standard deviation.

To assess the relationship between outcome gains and whether schools implement specific models, we will estimate equation (9) separately for each grade and outcome year ( $T$ ) of interest, where:  $i$  indexes schools;  $\Delta Y_i^T$  is the outcome gain for a given grade of interest;  $STM$  is a set of binary variables indicating which STM a school implemented;  $X$  is a set of school characteristics that includes demographic characteristics of the student body and the RDD assignment variables;  $STATE$  is a set of binary variables indicating the state where the school is located;  $\varepsilon$  is an error term; and  $\beta_0$ ,  $\beta_{STM}$ ,  $\beta_X$ , and  $\beta_{STATE}$  are parameters to be estimated.

$$(9) \quad \Delta Y_i^T = \beta_0 + \beta_{STM} STM_i + \beta_X X_i + \beta_{STATE} STATE_i + \varepsilon_i$$

The differences in outcome gains between STMs are given by the  $\beta_{STM}$  estimates. Because we include state indicator variables, these estimates are based only on within-state variation, meaning that the effects of STMs are not confounded with cross-state differences. Also, because we are focusing on differences among STMs in outcome gains, our estimates will be unaffected by changes over time that affect all schools within a state (for example, changing state assessments).

To estimate the relationship between individual practices (or bundles of practices) and outcome gains, we will estimate equation (10), which adds to equation (9) the term  $\beta_p P_i$ , where  $P$  is a set of binary variables indicating which practices (or bundles of practices) were implemented in each school and  $\beta_p$  is a set of parameters representing the differences in outcome gains associated with those practices. Schools that implemented the closure model will not be included in this analysis, since the closure model cannot involve any other practices.

$$(10) \quad \Delta Y_i^T = \beta_0 + \beta_p P_i + \beta_{STM} STM_i + \beta_X X_i + \beta_{STATE} STATE_i + \varepsilon_i$$

## b. Publication Plans

Table A.6 displays the anticipated timetable for project publications, which include three reports and six evaluation briefs.

The topics for the evaluation briefs will be determined at a later date. These briefs will be released separately and prior to the release of each report.

**Table A.6. Timetable for Project Publications**

Activity	Date
<b>Reports</b>	
Draft Report 1	June 2013
Revised Draft Report 1	August 2013
Final Report 1	February 2014
Draft Report 2	January 2014
Revised Draft Report 2	March 2014
Final Report 2	September 2014
Draft Report 3	December 2014
Revised Draft Report 3	January 2015
Final Report 3	July 2015
<b>Evaluation Briefs</b>	
Draft Evaluation Briefs 1 and 2	March 2013
Revised Draft Evaluation Briefs 1 and 2	April 2013
Final Evaluation Briefs 1 and 2	July 2013
Draft Evaluation Briefs 3 and 4	December 2013
Revised Draft Evaluation Briefs 3 and 4	January 2014
Final Evaluation Briefs 3 and 4	April 2014
Draft Evaluation Briefs 5 and 6	November 2014
Revised Draft Evaluation Briefs 5 and 6	December 2014
Final Evaluation Briefs 5 and 6	March 2015

## **17. Approval Not to Display the OMB Expiration Date**

All data collection instruments will include the OMB expiration date.

## **18. Explanation of Exceptions**

No exceptions are requested.

**REFERENCES**

Hahn, J., P. Todd, and W. Van Der Klaauw. "Regression Discontinuity." *Econometrica*, vol. 69, no. 1, 2001, pp. 201-209.

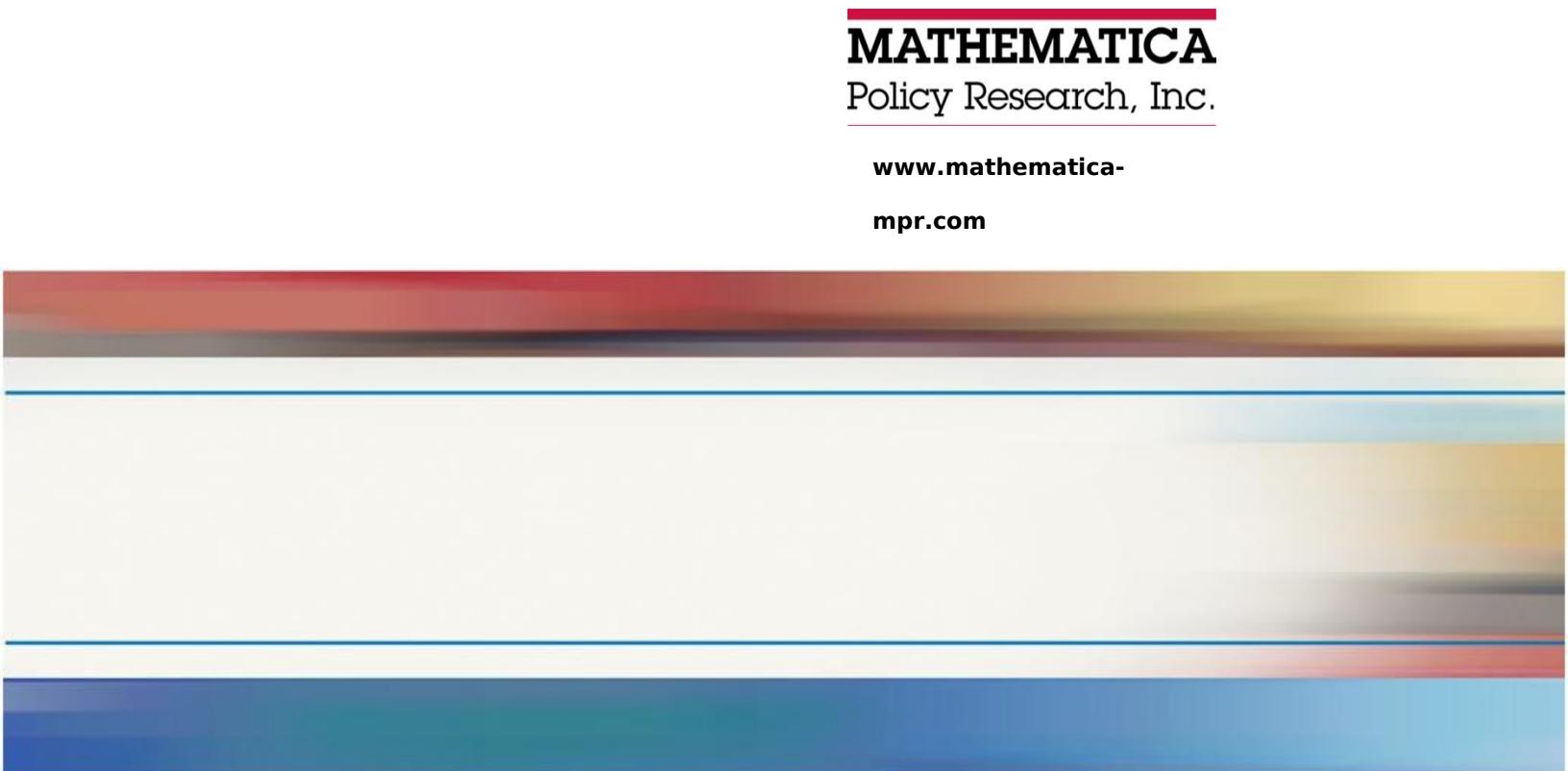
Hurlburt, S., K. C. Le Floch, S. B. Therriault, and S. Cole. "Baseline Analyses of SIG Applications and SIG-Eligible and SIG-Awarded Schools." (NCEE 2011-4019.) Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, 2011.

Shadish, W.R., T.D. Cook, and D.T. Campbell. *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin, 2002.

Trochim, W.M.K. *Research Design for Program Evaluation*. Beverly Hills, CA: Sage Publications, 1984.







**MATHEMATICA**  
Policy Research, Inc.

[www.mathematica-  
mpr.com](http://www.mathematica-mpr.com)

Improving public well-being by conducting high-quality, objective research and surveys

Princeton, NJ ■ Ann Arbor, MI ■ Cambridge, MA ■ Chicago, IL ■ Oakland, CA ■ Washington, DC

Mathematica® is a registered trademark of Mathematica Policy Research