

Implementation of Title I/II-A Program Initiatives

Supporting Statement for Paperwork Reduction Act Submission

PART B: Collection of Information Employing Statistical Methods

Contract ED-IES-11-C-0063

January 2018

Prepared for:
Institute of Education Sciences
U.S. Department of Education

Prepared by:
Westat
Mathematica Policy Research

Table of Contents

B. Collection of Information Employing Statistical Methods	1
Overview.....	1
B.1. Respondent Universe and Sampling Methods	2
B.1.1. State Sample	2
B.1.2. School District Sample	2
B.2. Information Collection Procedures	3
B.2.1. Notification of the Sample, Recruitment and Data Collection.....	3
B.2.2. Statistical Methodology for Stratification and Sample Selection.....	5
B.2.3. Estimation Procedures.....	10
B.2.3.1 State and District Level Implementation.....	10
B.2.3.2 Patterns of Cross-Level Implementation.....	11
B.2.3.3. Changes in Implementation over Time.....	11
B.2.3.4. Student Achievement Trends.....	11
B.2.4. Degree of Accuracy Needed.....	12
B.2.5. Unusual Problems Requiring Specialized Sampling Procedures.....	13
B.2.6. Use of Periodic (less than annual) Data Collection to Reduce Burden	13
B.3. Methods to Maximize Response Rates	13
B.3.1. Weighting the District Sample.....	14
B.4. Test of Procedures	14
B.5. Individuals Consulted on Statistical Aspects of Design	15
References.....	15
Appendix A State Survey.....	A-1
Appendix B District Survey.....	B-1
Appendix C Notification Letters and Reminder Emails.....	C-1

Tables

- B-1. Definitions of district size strata for the original district sample
- B-2. Final district sample sizes and relative sampling rates for the original sample by district poverty and size strata
- B-3. Properties of the stratification design for the original district sample

Part B. Collection of Information Employing Statistical Methods

This package is the second of two for the Implementation of Title I/II-A Program Initiatives study. The first data collection was conducted during spring and summer 2014 (OMB # 1850-0902; expiration date: 02/28/17). This package requests approval for the follow-up round of data collection that will include surveys of all states including the District of Columbia, the same nationally representative sample of school districts selected for the 2014 data collection, and a new nationally representative sample of charter school districts.¹ Unlike the first data collection, this follow-up data collection will not include principals and teachers. We anticipate that the state and school district surveys will begin in April 2018.

Overview

The Title I and Title II-A programs are part of the Elementary and Secondary Education Act (ESEA). These programs are intended to promote equal access to education by providing financial assistance to schools and districts which have a high percentage of students from low-income families (Title I); and by improving teacher and principal quality (Title II-A). During the 2014–15 school year, more than 100,000 public schools used Title I funds, and the program served over 25 million children (U.S. Department of Education, 2016a). Ninety-eight percent of districts nationally received Title II-A funding for the 2015–16 school year (U.S. Department of Education, 2016b).

ESEA was most recently reauthorized as the Every Student Succeeds Act (ESSA) in December 2015. Under Title I, ESSA offers states and districts considerable autonomy while requiring them to adopt accountability systems that set state-specific accountability goals and identify and support low-performing schools, challenging academic content standards, and aligned assessments. Under Title II-A, ESSA also provides funding for a broad array of permissible activities to improve the effectiveness of educators and achieve an equitable distribution of effective educators. These four areas—accountability and turning around low-performing schools, improving teacher and leader effectiveness, state content standards, and assessments—were also focuses of Titles I and II under the previous version of ESEA that was in effect at the time of this study's 2014 data collection. Relative to the prior version of the law, however, ESSA gives states substantially more discretion about the details of implementation, particularly with respect to accountability and educator effectiveness.

The purpose of the Implementation of Title I/II-A Program Initiatives study is to describe the implementation of policies and practices funded through Titles I and II-A of ESEA at multiple points in time. The 2014 Title I/II-A data collection examined the implementation of policies and practices at the state, district, school, and classroom levels. That data collection and resulting report² focused on policies and practices funded through Titles I and II-A which were implemented under ESEA Flexibility and other related federal initiatives since the previous National Assessment of Title I concluded in 2006.

The 2018 follow-up data collection will be limited to surveys of states and school districts. Given that the 2017-18 school year is early in the implementation of ESSA, we expect that most new policies and

¹ Charter school districts are districts where all associated schools are charter schools. Based on the 2011-12 Common Core of Data, these districts account for about one-half of all charter schools.

² Troppe, P., Milanowski, A., Heid, C., Gill, B., and Ross, C. (2017). *Implementation of Title I and Title II-A Program Initiatives: Results from 2013-14*. (NCEE 2017-4014). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

activities will not reach the school and classroom levels by the time the surveys are administered in Spring 2018. This follow-up data collection will provide information on state and district activities in the same four core areas as in the 2014 data collection. This approach will allow the study to not only describe the implementation of Title I and Title II-A provisions in 2018 but also compare it to 2014, in order to see if states are, for example, continuing to implement specific policies or are leaving more decisions up to districts.

The 2018 data collection will also include a new nationally representative sample of charter school districts and questions on school choice, an area of increasing federal and state policy attention. With the addition of charter school districts, the study can investigate whether the implementation of policies varies in charter school districts compared to traditional school districts. A charter school district often consists of only a single charter school, which may operate quite differently from traditional school districts, not only due to small size but also due to the absence of an elected school board and freedom from state regulations.

B.1. Respondent Universe and Sampling Methods

The study sample will include the universe of states and the District of Columbia, and a nationally representative sample of school districts. The district sample for the follow-up data collection will use the same sample of districts as was used for the 2014 surveys but add a new nationally representative sample of charter school districts to support comparisons of policies and practices with non-charter school districts.

B.1.1. State Sample

We will survey all 50 states and the District of Columbia.

B.1.2. School District Sample

The follow-up survey will collect data from the same sample of school districts used for the 2014 survey as well as a new sample of charter school districts.

For the 2014 survey, we selected a nationally representative sample of districts. This provides unbiased estimators of district characteristics, and provided the first stage of selection for samples of schools and teachers for the 2014 surveys. A nationally representative sample is necessary as Title I/II-A covers most of the U.S. public school educational system.

We also are interested in statistically comparing the implementation of initiatives funded by Title I and Title II-A by **district level of poverty** and **size of districts** based on student enrollment. Poverty is included because Title I is specifically intended to ameliorate the effects of poverty on local funding constraints and education opportunity. District capacity to implement initiatives also will be of interest for this study. Success in implementation of initiatives, particularly those that build upon each other such as having longitudinal data systems and identifying effective teachers, might be tied to district organizational capacity. Therefore, we also will examine implementation by district size.

For the 2018 data collection, we also are interested in examining the implementation of Title I/II-A initiatives by **charter school district status**. As noted above, charter school districts are likely to differ from traditional school districts in several ways: they are not governed by local school boards, they are exempt from some state rules and regulations, and they often consist of only a single school. The original district sample did not include enough of these districts to support national estimates for charter districts. In order to produce national estimates for charter districts for 2018, we will add a new sample of these districts.

For the original district sample, we implemented a “minimax” sample design that struck a balance between producing efficient student-weighted estimates and efficient district-weighted estimates. This was appropriate because the policy levers of Title I and Title II-A are directed at districts (and schools and states). For some purposes, it will be useful to understand the average experience of students across the country, but for other purposes, we will want to understand the experience and behavior of the average districts (and schools)—the units that the U.S. Department of Education (ED) expects its policies to immediately influence. The minimax design is a compromise between a design that is relatively efficient (i.e., allows estimates with narrow confidence intervals) for answering questions about the number or proportion of U.S. public school students in districts implementing initiatives of interest, and one that is relatively efficient for answering questions about the number or proportion of U.S. school districts implementing such initiatives.³

To construct the original district sampling frame, we used data primarily from the National Center for Education Statistics’ (NCES) Common Core of Data (CCD), with supplementary data from sources such as the U.S. Bureau of the Census’s district-level SAIPE (Small Area Income and Poverty Estimates) program for school-district percentages of children in families in poverty.

We drew an original district sample of 570 districts out of 15,762 school districts. See Table B-2 in Section B.2 for the universe counts by stratification classifications for the original sample. Three of these districts have closed since the 2014 data collection, leaving 567 of the original districts for the 2018 data collection. We will draw a new sample of 152 charter school districts for the 2018 data collection (see section B.2.2 for additional information on the design for this sample).

B.2. Information Collection Procedures

B.2.1. Notification of the Sample, Recruitment and Data Collection

Re-introduce the Study to State Education Agencies. We will begin by sending the chief state school officer and the state Title I administrator a notification letter (see Appendix C) explaining the study, thanking the state for its previous participation in the 2014 survey, reminding states of the importance of the state’s involvement, and the mandatory nature of the state’s response. We will then follow up with a phone call to the state Title I administrator to answer questions about the study and identify

³ The minimax design differs from the one used for the Title I study that concluded in 2006. The previous study selected districts probability proportional to size (PPS), with size measured by student enrollment. The PPS design is quite efficient for estimating the proportion of students enrolled in districts implementing policies of interest. However, when estimating the percent of districts implementing a policy, the PPS design is relatively inefficient, compared to a simple random sample. This is because relatively few small and medium-sized districts are included in a PPS design. This, in turn, requires the small and medium-sized districts in the sample to be given greater weight to better represent the population of districts nationwide and can lead to relatively wide confidence intervals around estimates of proportions of districts.

additional state-level respondents based on areas of expertise. The survey will be sent to the chief state school officer in each of the 50 states and the District of Columbia beginning in April 2018, with the expectation that different sections of the survey may be filled out by different staff.

The mailing will include a password and secure web address to access an electronic version of the questionnaire and instructions on how to submit the questionnaire via a secure SharePoint site. After the initial mailing, we will also send a follow-up email that also includes the web address and password. Project staff will monitor completion rates, review the instruments for completeness throughout the field period, and follow up by email and telephone as needed to answer questions and encourage completion. During these calls, state representatives will be given the option of completing the module by telephone with the researcher. Each of the four topic areas (accountability and low-performing schools, improving teacher and leader effectiveness, state content standards, and assessments) will on average require between 15 minutes and 1 hour to complete.

Researchers knowledgeable about the four content areas will review the completed questionnaire and documentation downloaded from state and other publicly available web sites for completeness. We will then conduct follow up calls with states to clarify any ambiguous answers in the survey.

Re-introduce the Study to District Leaders. We will send notification letters (see Appendix C) by email and mail to the superintendents of the sampled districts. Notification letters will thank the district in the original sample for its participation in the 2014 survey and remind them of the study's importance and benefits. For districts in the new charter school district sample, the notification letter will introduce the study to them and underscore the study's importance and benefits. Sending the notification letters both by email and mail will increase the likelihood that addressees will receive our communications in a timely manner. States and districts receiving Title I and Title II-A funds have an obligation to participate in Department evaluations (Education Department General Administrative Regulations (EDGAR) (34 C.F.R. § 76.591)), and virtually all states and districts receive Title I and/or Title II-A funds.

As was found in the 2014 data collection, the district surveys will require input from several key individuals. We will ask the district superintendent to provide contact information (including email) for a person designated by the superintendent as the study liaison, who will coordinate the completion of the survey by the appropriate staff. Once the district liaison is identified, we will conduct all follow-up directly with the district liaison.

For those districts that do not respond and identify a study liaison online within five days, we will make follow-up telephone calls to district superintendents to confirm receipt of the letter, answer any questions, confirm the identity of the study liaison, and obtain the designated liaison's contact information. We will follow all required procedures, and as necessary, obtain approval of the district for its participation in the study through submissions of the required research application.

Administer Surveys. By email and mail, we will send a notification letter to district liaisons. The letter and email will underscore the purpose of the study and the importance of participation and informing district respondents that completing the web-based survey is mandatory and required by law. Mailed letters will include the survey URL and login information for responding to the survey as a web-based instrument. For security purposes, the first email will provide the URL and the User ID with the respondent's password provided in a second email.

All communications will include a toll-free study number and a study email address for respondents' questions and technical support. Based on Westat's experience on large-scale data collections, we will assign several trained research staff to answer the study hotline and reply to emails in the study mailbox. We will train them on the purpose of the study, the obligations of district respondents to participate in the evaluation, and the details for completing the web-based survey. Content questions will be referred to the study leadership. An internal FAQ document will be developed and updated as needed throughout the course of data collection to ensure that the research staff has the most current information on the study.

The web will be our primary method of data collection for the district surveys. We will offer respondents the option of emailing them an electronic version of the survey (e.g., PDF or Word document) to complete and return by email or completing a paper-and-pencil instrument. However, we have found that the vast majority of respondents in districts prefer the web-based approach. A phone survey option will be offered to all respondents as part of the nonresponse follow-up effort. Since the web-based surveys will include data checks, we will use the web-based surveys to enter any surveys received on hard copy or by phone.

Westat will develop a web-based data monitoring system (DMS) to track the sample for each instrument, record the status of district approvals, generate materials for mailings, and monitor survey response rates.

The state survey will use electronic fillable PDFs. The survey will be divided into modules, each a separate fillable PDF, which will allow the appropriate staff with expertise in each area to respond separately. This approach will reduce burden for respondents as (a) each individual will have fewer questions to answer and (b) respondents will be asked questions concerning topics in which they are well versed and answers should be readily available. State Education Agency staff may prefer to complete individual sections of the survey with the help of colleagues who can help ensure that responses to questions about state policies and programs are accurate. The fillable PDF documents facilitate this collaborative approach to completing the survey by making it easier to view and print an entire survey section or just parts of it, both before and after it is completed. Each state will be given access to a secure SharePoint site to facilitate access to the survey for multiple respondents and to allow respondents to forward any documents they feel would help us better understand their state policies. An electronic fillable PDF is also a more cost-effective electronic version for a survey of 51 entities.

B.2.2. Statistical Methodology for Stratification and Sample Selection

The follow-up district survey will use the same districts used for the 2014 survey and a new sample of charter school districts. The district sample for the 2014 survey was stratified by poverty status and district size. The poverty strata were defined based on the percentage of children in families in poverty. The high-poverty stratum consisted of the roughly 25 percent of districts with percentages greater than the national 75th percentile.⁴ The low/medium-poverty stratum consisted of the complement set (roughly 75 percent of the districts).

⁴ This percentile is weighted by enrollment, and is found using the U.S. Bureau of the Census school district SAIPE (Small Area Income and Poverty Estimates). Linking the most recent SAIPE District data file to the 2011-12 Common Core of Data District Universe Frame, we found that this 75th percentile is 27.7% percent of children in families in poverty.

The district size strata for the original sample are given in Table B-1. It should be noted that for comparing adjacent classes, each class has an enrollment range roughly three times greater than the preceding class (in terms of minimums, mean value, or maximums). In addition, a separate stratum was created for small states (according to the number of districts) to guarantee that every state has at least one selected district.

Table B-1. Definitions of district size strata for the original district sample

District oversampling class	Lower bound district enrollment	Upper bound district enrollment
G	1	500
F	501	1,500
E	1,501	5,000
D	5,001	15,000
C	15,001	50,000
B/A	50,001	no limit

Note: District classes A and B were merged only for presentation purposes in this document to avoid sample disclosure.

A total of 570 districts were sampled, with oversampling of high-poverty districts by a factor of 3. This oversampling brought the expected sample size for high-poverty districts in line with the expected sample size for low/medium-poverty districts.

Table B-2 presents the final district sample sizes and relative sampling rates (as compared to the stratum with the lowest sampling rate) for the original sample by district poverty and size strata. The counts are based on the 2011-12 school-year CCD frame. Note that under a probability proportionate to size by enrollment design, the relative sampling rates between neighboring district size classes would be 3, as that is roughly the enrollment ratio. By using powers of 1.80 rather than powers of 3 as relative sampling factors, we are oversampling the strata with the higher enrollments, but not to the full extent justified by the ratios of enrollment means.

Table B-2. Final district sample sizes and relative sampling rates for the original sample by district poverty and size strata

Poverty Stratum	District size class	District count	Student enrollment (in 1000s)	Relative sampling rate	District sample size
Low/medium poverty	G	3,961	937.4	1.0	24
Low/medium poverty	F	3,430	3,127.0	1.8	55
Low/medium poverty	E	3,060	8,426.0	3.2	97
Low/medium poverty	D	1,112	9,139.5	5.8	65
Low/medium poverty	C	346	8,728.7	10.5	36
Low/medium poverty	B/A	67	6,172.2	18.9+	19
Low/medium poverty	Total	11,976	36,530.8		296
High poverty	G	1,687	384.7	3.0	25
High poverty	F	948	838.6	5.4	49
High poverty	E	763	2,095.3	9.7	89
High poverty	D	265	2,172.1	17.5	56
High poverty	C	98	2,592.6	31.5	37
High poverty	B/A	25	4,101.1	56.7+	18
High poverty	Total	3,786	12,184.4		274

Notes: District size class was defined in terms of student enrollment intervals: G: 500 or less; F: 501 to 1,500; E: 1,501 to 5,000; D: 5,001 to 15,000; C: 15,001 to 50,000; B/A: 50,001 and over. District classes A and B were merged only for presentation purposes in this document to avoid sample disclosure.

We call this sample design a ‘minimax’ design, as it is designed to equalize the efficiency for two types of estimates. The first type of estimate counts each district as one in the population, so that the base weight is the inverse of the district probability of selection. This type of ‘count-based’ estimate answers questions such as “What percentage of districts have characteristic X?” The second type of estimate includes enrollment of the district, so that the sampling base weight is the enrollment divided by the probability of selection. This type of ‘enrollment-based’ estimate answers questions such as “What percentage of students are enrolled in districts which have characteristic X?” A probability proportionate to enrollment design will lead to optimal efficiency for the second type of estimate, but will have poor efficiency for the first type of estimate (as the district weights will be close to equality for the enrollment-based estimate, but will vary considerably for the count-based estimate). On the other hand, a simple stratified design with no oversampling of larger district-size strata will have high efficiency for count-based estimates, but poor efficiency for enrollment-based estimates. This ‘middle-ground’ design oversampled the higher enrollment district-size strata, but proportional to the 0.535 root⁵ of the enrollment mean in the stratum, rather than to enrollment directly,⁶ and has reasonable efficiency for both count-based estimates and enrollment-based estimates (the design is set up to equalize the efficiency for both types of estimates, at the cost of not being as good for each type of estimate as the optimal design for that type of estimate). Table B-3 summarizes the properties of this design for the original sample.⁷

⁵ 1.8 is the 0.535 root of 3.

⁶ This design is close to a ‘square root’ design, except that it is stratified design rather than a fully PPS design (sampling rates are equal within strata), and the root used is slightly larger than $\frac{1}{2}$.

⁷ As processed to drop ineligible schools and entities, schools with no enrollment, etc.

Table B-3. Properties of the stratification design for the original district sample

Power Property	Enrollment-based weight estimates	Count-based weight estimates
Effective sample size: All districts	294.6	292.4
Effective sample size: High-poverty districts	237.7	174.8
Effective sample size: Low/medium-poverty districts	179.6	186.9
Minimum detectable effect size (MDES) comparing Poverty District Strata	27.7%	29.5%

The effective sample sizes are the sample sizes for a simple random sample which would provide the same precision as the design.⁸ Note that the effective sample size for all-district estimates is about half of the district sample size of 570. This large ratio is caused partially by the oversampling of high-poverty districts. Note also an equalization of effective sample sizes for the two types of estimates. This is the ‘minimax’ aspect. The MDES (minimum detectable effect size) is computed for evaluating the null hypothesis of no difference between the high-poverty and the low/medium-poverty districts for a range of district-level characteristics.⁹ This is a very important comparison at the district level for this study, and the sample design is carefully tailored with this comparison in mind (we effectively equate the sample designs for the two strata, though the high-poverty districts are only one quarter of enrollment). The sample design does achieve an MDES lower than 30 percent for both types of estimates.¹⁰

The original district frame is based on the 2011–12 CCD frame, as processed through National Assessment of Educational Progress (NAEP) macros to purge entities that are not in scope (e.g., administrative districts, district consortiums, entities devoted to auxiliary educational services, etc.), as well as a canvassing of new districts from the preliminary CCD frame from the following year (only districts with positive enrollments on the CCD frame, with other out-of-scope entities purged out). All school districts and independent charter districts with at least one eligible school and at least one enrolled student were included in the frame.¹¹

Further Details of the Original District Sampling. The primary strata were by district-size class and poverty status as described in the previous section. Within the poverty and district size class strata, we implicitly stratified districts by drawing a systematic sample from a list of districts ordered by the implicit stratification characteristics. Districts in the small state stratum (all states with expected district sample sizes less than or equal to 5) were implicitly stratified by Census region, state, poverty status, urbanicity, and district enrollment. Districts in large states (all states with expected district sample sizes greater than 5) were implicitly stratified by poverty status, Census region, urbanicity, and district enrollment.

⁸ The effective sample size is equal to the population variance divided by the sampling variance under the design.

⁹ We assume a null hypothesis of no difference with a two-sided critical region with a 5 percent alpha level. We find the smallest population difference that would be detectable with this test with 80 percent power. The MDES is this population difference divided by the (assumed) common population standard deviation for each subgroup.

¹⁰ This assumes 100 percent district response. District nonresponse will degrade these power results. However, we achieved a district response rate of 99 percent on the 2014 survey and expect minimal district nonresponse for the follow-up survey as this study is by law mandatory for the districts.

¹¹ In defining district eligibility, we follow the criteria from the NAEP. The NAEP macros for excluding districts are applied also in the generation of the district frame here.

The implicit stratification by Census region and urbanicity promotes the nationally representative nature of the sample.

The districts¹² were selected within the district size and poverty strata using a “minimax” design. The minimax design oversampled the size strata corresponding to larger enrollment (but not as heavily as a probability proportionate to size design would).¹³ In addition, districts in the high-poverty stratum were oversampled by a factor of three to improve analytic precision. High-poverty districts are roughly one-quarter of the districts, but with oversampling were roughly one-half of the sample. The realized sample sizes were 296 and 274 for low-/medium- and high-poverty districts, respectively.

The high-poverty stratum was defined by SAIPE estimates of percentages of 5 to 17 year old children in poverty for the school district. We computed the weighted 75th percentile percentage over all districts in the U.S., and the mean value became the cutoff. Districts with percentages lower than the cutoff were designated ‘low/medium-poverty’ and districts with percentages higher than the cutoff were designated ‘high poverty.’ Independent districts such as charter school districts were associated with the public school district that they are associated with geographically, as only the primary geographically-based public school districts have poverty estimates from SAIPE.

New Sample of Charter School Districts for 2018 Data Collection. There is an interest in examining districts with only charter schools as compared to other districts (districts with no charter schools or ‘mixed districts’: districts with charter and non-charter schools). There were 24 charter-only districts selected in the original district sample. Three of these districts have since closed. For the follow-up data collection, we want to increase this sample to 173 to provide a nationally representative sample of charter school districts and to achieve an MDES of 30 percent for comparing charter-only to all other districts. This will be accomplished by drawing a new sample from the most recent CCD frame (2016-17) of eligible charter school districts. This new sample will draw 152 additional charter school districts.¹⁴ The charter school district sample will be sampled in such a way as to not have oversampling for the high-poverty stratum (as was the case for the original sample), to maximize the precision of this sample as representative of all charter school districts, while retaining all other aspects of the original sample design. As noted above, the original district sample will be retained for the 2018 data collection (including all 21 charter school districts from the original sample still in operation as well as all 546 non-charter-only districts). As a result, all of the precision properties of the original district sample will be retained. In addition, the new sample will have a precision at least as high as the original sample of 570 districts for national estimates, and for high poverty/medium-low poverty comparisons as given in Table B-3.

¹² Districts with only one school had a sampling rate set to one-quarter of other districts in the same poverty/district size stratum (with correspondingly higher weights to ensure unbiased estimates). They were still represented in the study, but we had fewer of these districts. This method of under sampling is similar to that done in the NAEP for schools with very small numbers of students.

¹³ Note that this relative oversampling factor is somewhat larger than the square root of the relative mean enrollment size, and that within each district size stratum the districts are selected with equal probability.

¹⁴ The necessary sample size is 125 completed charter school interviews: the 152 allows for charter school district nonresponse as this was experienced among the 24 selected in the original sample. The longitudinal respondents from the additional 21 from the original sample still in operation should also contribute to the precision of the comparison, which should further add to the power.

B.2.3. Estimation Procedures

The follow-up report will include analyses with four main objectives: (1) describing the extent to which policy and program initiatives related to the objectives of Title I and Title II-A are being implemented at the state and district levels, including how implementation varies by selected state and district characteristics; (2) describing patterns of cross-level implementation; (3) describing changes in implementation since 2014; and (4) describing trends in student achievement. Each set of planned analyses is described below.

B.2.3.1. State and District Level Implementation

The primary goal of the follow-up data collection is to describe the implementation of policy and program initiatives related to the objectives of Title I and Title II-A. To achieve this goal, extensive descriptive analyses will be conducted using survey data. We anticipate that relatively straightforward descriptive statistics (e.g., means, frequencies, and percentages) and simple statistical tests (e.g., tests for differences of proportions) will typically be used to answer the research questions.

While simple descriptive statistics such as means and percentages will provide answers to many of our questions, cross-tabulations will be important to answering questions about variation across state and district characteristics. The primary characteristics of interest for the cross-tabulations are:

- District poverty level: Poverty is included because Title I is specifically intended to ameliorate the effects of poverty on local funding constraints and educational opportunity.
- District size: Size is included because it may be related to district capacity to develop and implement programs.
- District urbanicity: Urbanicity is included because of the relationships between educational opportunity and rural isolation and the concentration of poverty in urban schools.
- Concentration of English learners: Concentration of English learners is included because of the increased emphasis on ensuring that this group of students reaches English proficiency and meets state standards, and the recognition that modifications in testing as well as instruction will be needed to facilitate progress of these students.
- District charter school status: There is interest in examining whether district policies and practices vary in districts consisting entirely of charter schools.

The use of stratification (and oversampling when necessary) in our sample design was introduced to ensure reasonable power for specific subgroup comparisons, and our sample includes units that vary on other characteristics (e.g., urbanicity) to allow the first four comparisons described above. The new charter school district sample will allow for comparisons of charter school districts to other districts.

Because of the use of a statistical sample, survey data presented for districts will be weighted to national totals (tabulations will provide standard errors for the reported estimated statistics). In addition, the descriptive tables will indicate where differences between subgroups are statistically significant. We will use Chi-Square tests to test for significant differences among distributions and *t*-tests for differences in means.

B.2.3.2. Patterns of Cross-Level Implementation

Planned analyses of cross-level implementation involve examining responses of districts by categories of state responses. These analyses will examine the relationship between policies and programs originating at the state level and implementation “on the ground” in districts. Though the planned analyses cannot support causal conclusions about the effects of state actions on district implementation, they can provide evidence on the extent to which district practices are consistent with state policies.

Conceptually, these analyses posit that certain state policy choices influence what happens in districts. Examples of potential cross-level analyses include:

- The actions planned by districts to turn around lowest-performing schools by the interventions required by states and/or the guidance provided by states.
- District reports of major challenges implementing state content standards by whether the state reported making major changes to content standards since April 2014.

These cross-level analyses will be based on cross-tabulations. For example, a straightforward set of descriptive tables will show the relationships between survey responses at the state level and the mean responses or percentages at the district level. Breakdowns by the previously discussed state and district characteristics will further sharpen the interpretation of these relationships.

B.2.3.3. Changes in Implementation over Time

Analyses will address how implementation of policies and programs may have changed since the 2014 data collection. These analyses will document changes over time and will not attempt to derive any causal inferences or attribute any changes directly to the passage of ESSA.

These analyses will examine key policies that may have changed since 2014, such as:

- The number of states requiring teacher evaluation systems with certain components compared with the number of states requiring systems with those same components in 2014.
- The percentage of school districts with lowest-performing schools that reported receiving higher levels of professional development for teachers and principals in these schools compared with the percentage in 2014.

These analyses will be based on a comparison of responses to the same questions posed in 2014 and 2018. The analyses of state responses will report differences over time in the number of states reporting particular policy choices. The analyses of district responses will be weighted to national totals and the tables will indicate where differences between 2014 and 2018 are statistically significant.

B.2.3.4. Student Achievement Trends

We will conduct a descriptive analysis of state-level and national trends in student proficiency levels, using NAEP results and results on state assessments. These analyses will address the research question of whether students are making progress on meeting state academic achievement standards within states and how this progress varies across states. The analyses will be entirely descriptive and will not attempt to derive any causal inferences. While these data do not provide evidence on the effects of any particular ESSA policy or of ESSA as a whole, they provide context for the implementation analyses in the remainder of the report.

The state-level analysis will use two data sources:

- State-by-state NAEP proficiency rates since 2005, math, reading, and science, grades 4, 8, and 12; and
- State-by-state proficiency rates on each state's own assessments since 2007, math and reading, grades 4, 8, and 9-12 (from *EDFacts*).

For the analyses using state-by-state NAEP proficiency rates, we use 2005 as the baseline year because it was the final year included in the last national assessment of Title I. For the analyses using state-by-state proficiency rates on each state's own assessments, we use 2007 as the baseline year because *EDFacts* does not consistently have proficiency data for all states prior to 2007. The results of this analysis will consist of a comparison of changes in student proficiency levels across states.

We will also examine national trends in student proficiency rates for the nation as a whole and for certain subgroups of students. For the nationwide average, racial/ethnic subgroups, students with disabilities, and English learners, we will present the following information:

- Percent proficient on NAEP, in 2005 and most recent year, and the difference between the two years; and
- Percent proficient on state assessments, in 2007 and the most recent year, and the difference between the two years.

For each of these items we will report results separately for math and reading, and for grades 4, 8, and high school. NAEP assessments are given in grade 12, but the tested high-school grade(s) on state assessments vary across states, and we will report results for the tested high-school grade(s) for each state. Additionally, we will report the percentage of states that are improving, declining, or remaining the same in their proficiency rates on NAEP.

Finally, we will use existing data to show the trend in the nationwide graduation rate and the latest graduation rates for each state using the four-year adjusted-cohort graduation rate data from *EDFacts*. These data will be examined by student subgroups (e.g., major racial and ethnic groups, economically disadvantaged).

Please see Section B.3.1 for a discussion of our weighting procedures.

B.2.4. Degree of Accuracy Needed

For the follow-up survey, we require statistical precision at the district level. The original sample size of 570 districts roughly evenly split between high-poverty districts and the complement stratum will provide an MDES of 30 percent for comparing these two important district subgroups. A sample size of 152 charter districts (with 125 expected responses) plus the original 21 charter districts still in operation will provide an MDES of 30 percent for comparing charter school districts with other districts.

B.2.5. Unusual Problems Requiring Specialized Sampling Procedures

There are no unusual problems requiring specialized sampling procedures.

B.2.6. Use of Periodic (less than annual) Data Collection to Reduce Burden

The follow-up survey will be collected during the 2017–18 school year.

B.3. Methods to Maximize Response Rates

This OMB package addresses only the 2017–18 school year data collection. We achieved very high response rates for the state survey (100 percent) and district survey (99 percent) during the 2014 data collections. As a result, we expect to achieve very high response rates again for the follow-up survey to the same states and school districts. We plan to work with states and school districts to explain the importance of this follow-up data collection effort and to make it as easy as possible to comply. For all respondents, a clear description of the study design, the nature and importance of the study, and the OMB clearance information will be provided.

For the states, we will be courteous but persistent in follow-up with participants who do not respond in a timely manner to our attempts. As noted earlier, we will also be very flexible gathering our data, allowing different people to respond to the different content areas and in whichever mode is easiest -- electronic, hard copy or telephone format.

We will initiate several forms of follow-up contacts with respondents who have not responded to our communication. We will use a combination of reminder postcards, emails and follow-up letters to encourage respondents to complete the surveys. The project management system developed for this study will be the primary tool for monitoring whether surveys have been initiated. After 10 days, we will send an email message (or postcard for those without email) to all non-respondents indicating that we have not received a completed survey and encouraging them to submit one soon. Within seven business days of this first follow-up, we will mail non-respondents a hard copy package including all materials in the initial mailing. Ten days after the second follow-up, we will telephone the remaining non-respondents to ask that they complete the survey and offer them the option to answer the survey by phone, either at that time or at a time to be scheduled during the call.

To maximize response rates, we also will (1) provide clear instructions and user-friendly materials, (2) offer technical assistance for survey respondents using a toll-free telephone number or email, and (3) monitor progress regularly. In recognition of the fact that district administrators have many demands on their time, and typically, these administrators receive numerous requests to participate in studies and

complete surveys for federal and state governments, district offices, and independent researchers, we plan to identify a district liaison. For most districts, completion of the district survey will require input from multiple respondents, and the district liaison's role will be pivotal in positively impacting participation, collecting high quality data, and achieving a minimum 85 percent response rate.

B.3.1. Weighting the District Sample

After completion of the field data collection in each year, we plan to weight the district follow-up data to provide a nationally representative estimator at the district level. The district weighting process will involve developing unit-base sampling and replicate weights, then adjusting these weights as necessary to account for survey nonresponse.

Replicate weights will be generated to provide consistent jackknife replicate variance estimators (statistical packages such as STATA and SAS Version 9.2 allow for easy computation of replicate variance estimates). The development of replicate weights will facilitate the computation of standard errors for the complex analyses necessary for this survey. For districts selected with certainty into the sample, the replicate weights will equal the base sampling weights. For the noncertainty districts, the replicate weights will be generated using the jackknife replication method.

We anticipate limited nonresponse at the district level, which we will adjust for by utilizing information about the non-responding districts from the frame. This information will be used to generate nonresponse cells with differential response propensities. The nonresponse adjustments will be equal to the ratio of the frame weighted count to the sum of weights for respondents. This will adjust for bias from nonresponse, and also adjust for differences from the frame (accomplishing poststratification). If the cell structure is too rich, we may utilize raking (multi-dimensional adjustment). The cell structure for districts will include district poverty status, Census region, urbanicity, and district size. With the large charter-school district supplement, charter status will also be included in the cell structure. Depending on differences in survey response between the 2014 and 2018 data collections (i.e., a district responded in 2014 but not in 2018), we also will consider incorporating information from the 2014 survey data into the nonresponse analysis to leverage additional information to calibrate the weights of 2018 respondents to 2014 respondent percentages for selected survey items.

There are two types of district weights that will be generated: enrollment-based weights and unit-based weights. The first set of weights will add to total enrollment and the second to the simple count of districts on the frame. A separate set of nonresponse adjustments will be calculated for each set of weights. We also will generate longitudinal unit-based and enrollment-based weights for districts that participate in both the 2014 and 2018 data collections. Districts in the new charter school district sample will not have a longitudinal weight.

B.4. Test of Procedures

The study team pretested the state and district surveys to ensure that questions added to or substantially revised from the 2014 survey were clear and that the average survey completion time is within expectations. The state survey was divided into its five sections to reduce the burden of the pretest for any one respondent, and each section was pretested with three or four states. The pretest states represented a mix in terms of state policies and practices relevant to the survey such as previous

ESEA flexibility status and use of the Common Core State Standards and the aligned assessments. States completed a Word or paper version of the survey section and participated in a debriefing interview.

The study team recruited seven pretest districts to complete a paper or Word version of the survey and participate in a debriefing interview. Five districts were traditional public school districts, and two were charter-only districts. The five traditional public school districts varied in terms of size, student population, and state policy context.

As a result of the pretests, the study team reduced the number of questions on both surveys and reorganized other questions to improve clarity and reduce respondent burden. We also revised the state survey time to complete to 180 minutes to more accurately reflect the expected average completion time. The average time to complete for the district survey remains at 60 minutes, which we believe is accurate given the reduction in the number of questions and other efforts to streamline and clarify the survey questions.

B.5. Individuals Consulted on Statistical Aspects of Design

The individuals consulted on the statistical aspects of the school district sample design include:

Patty Troppe, Westat, Project Director, 301-294-3924

Anthony Milanowski, Westat, Co-Principal Investigator, 240-453-2718

Brian Gill, Mathematica, Co-Principal Investigator, 617-301-8962

Lou Rizzo, Westat, Senior Statistician, 301-294-4486

David Morganstein, Westat, Vice-President Statistical Group, 301-251-8215

References

U.S. Department of Education. (2016a). Retrieved from <http://eddataexpress.ed.gov/>

U.S. Department of Education. (2016b). Findings from the 2015–2016 Survey on the Use of Funds Under Title II, Part A. Retrieved from: <http://www2.ed.gov/programs/teacherqual/leasurveyfunds rpt82016.pdf>