

Institute of Educational Sciences  
National Center for Education Statistics

## COORDINATING DESIGNS FOR MULTIPLE SURVEYS

May 2022

### EXECUTIVE SUMMARY

The National Center for Education Statistics (NCES) and other Centers within the Institute of Education Sciences (IES) collect data on a national scale from districts, schools, and individual administrators, teachers, students and parents. The surveys, assessments and other studies vary widely in purpose and scope, but all contribute to the information available about individual schools, districts and states. In addition to creating databases for individual data collections, this compiled information is used in turn for designing future samples.

In 2020, NCES commissioned the National Institute of Statistical Sciences (NISS) to assemble a pair of panels of technical experts, *Post-COVID Surveys* and *Setting Priorities for Federal Data Access to Expand the Context for Education Data* to consider opportunities for changing the sampling paradigm and process. From the NCES point of view dual goals were to address the rising nonresponse and lack of participation and to enrich the information base by linking data collections. From the points of view of school districts, schools and participants, the need was to understand and alleviate the burden of participation. These two panels became the first in a series to examine the principal issues in greater detail. The next pair of panels, *Connecting the Dots, I & II* examined the technical design and the implementation issues of coordinating the sampling across multiple data collections.

This *FORUM* responds to the findings and recommendations of those four panels. The broad recommendations of the first two panels were to develop a combined approach for data collections during each academic year and to expand the information base while gaining efficiency by linking data and eliminating redundant requests. The second pair of panels focused on identifying specific steps required to achieve the original goals for NCES and for the education community.

The goal of NCES-sponsored research (through NISS) to further advance innovations required by a coordinated sampling approach was to define the statistical challenges, identify the critical points for technical solution, and examine the several posited strategies for solution. This *FORUM* brought together statistical experts and NCES staff for presentations of the preliminary work on fleshing out the most promising sampling paradigms from *Connecting the Dots, I & II*. Following those presentations, a technical working session engaged additional statistical experts and NCES staff in formal and informal discussion.

- The first *FORUM* presentation set the context with a brief outline of the progression of panels leading to this point. The two presentations that followed outlined different strategies for

## CONNECTING THE DOTS

developing a new class of design approaches for coordinating multiple surveys with varied objectives.

### *Statistics: The Problem Space*

A sequence of expert panels on design innovation, technical issues and the challenges to implementation led to the research into new design strategies that are the focus of this FORUM. These panels set the context, *i.e.*, motivation, rationale and constraints, that define statistical challenges for coordinated sampling designs. Several specific strategies were proposed for further development and consideration.

### *Some Topics for Coordinated Sampling of Schools over Time*

Sampling design for longitudinal surveys is tricky as we wish to balance the trade-off between the response burden and the statistical efficiency. To reduce the response burden, it is desirable to develop a negatively coordinated sampling design. We propose a modification of the Swiss method to allow for differential coordination for each sampling unit. In addition, statistical methods for imposing balancing conditions in planned missingness and adaptive sample size allocation for stratified random sampling are discussed.

### *Coordinating Sample Design for Multiple NCES Surveys*

We have developed four strategies for a coordinated sampling process to potentially reduce response burdens: 1) Independently select schools for each survey, compute the burden for each selected school, and randomly substitute schools from the same stratum. 2) Independently select schools for each survey, compute the burden for each selected school, and reject samples that exceed the burden, 3) Sequentially sample schools based on a random survey order and decrease the selection probability for schools selected in previous surveys, and 4) Use matrix sampling of to assign surveys to schools using a probabilistic mechanism, *i.e.*, create replicates.

Discussion of technical points opened with formal comments from invited discussants. This working session continued engaging NCES staff, presenters and discussants in open discussion of technical points, specific applicability to NCES studies, and requirements for moving forward. The session culminated in five recommendations for immediate next steps.

### **Recommendations**

- Pursue multiple promising strategies simultaneously, developing each selected methodology sufficiently to permit validation and feasibility testing of implementation, and to allow (preliminary) estimation of precision and bounds on bias. Strategies deserving of consideration for development include strategies derived from block design of experiment from a common sampling frame, also Bayesian design incorporating prior or non-probability information, a modified Swiss method with balance, and matrix sampling that creates separate sampling frames.
- Create both simulation and realistic test files and resources for evaluation of strengths, vulnerabilities and capacity for scaling up with extension to additional IES/Department of Education studies and surveys.

## CONNECTING THE DOTS

- Initiate consolidation of stratum definitions and items to enable data sharing for sampling design, reduce redundant data requests, and facilitate efficient, balanced sampling designs.
- Expand the detail available for CCD schools; for sampled schools collect data on participation history (e.g., agreement/refusal with rationale at each level of decision-making) as a basis for designing samples for future studies, and for research purposes.
- Remain open to continuing innovation of sampling strategy, especially to take advantage of alternative data sources (e.g., non-probability online data sources), to adjust selection probability (e.g., burden metrics) and to use model-based prediction (e.g., propensity for nonresponse, magnitude of potential contribution to bias).