

# **Volume I:**

Request for Clearance for a Cognitive Study to  
Test the Feasibility of Tracking Principals'  
Employment Status

1850-0803

July 13, 2008

## **Volume I**

### **Justification**

NCES has as its legislative mission the collection and publication of data on the condition of education in the Nation (the National Education Statistics Act of 1994, P.L. 103-382, October 20, 1994 (20 USC 9001)). This study is specifically designed to support this mission by providing a means to study retention of key personnel in the education process, namely, elementary and secondary school principals. To date, nationally representative studies of principal retention focusing on the event of leaving and the reasons for doing so have not been conducted. This study will provide a better understanding of why principals leave schools and the profession more generally, and allow for the analysis of changes in these reasons when it starts to be repeated as part of the regular Schools and Staffing Surveys. In addition, this study is designed to study the feasibility of using a third party proprietary annual data set to conduct analysis that will yield annual retention rates for principals.

The Schools and Staffing Survey (SASS) surveys teachers and principals in a nationally representative sample of schools every 4 years. SASS is designed to collect information on the characteristics and qualifications of teachers and principals, teacher hiring practices, professional development, class size and other conditions of American schools. In year 5, SASS is followed by a survey known as the Teacher Follow-up Survey (TFS). The TFS is a follow-up of a sample of the elementary and secondary school teachers who participated in the previous year's Schools and Staffing Survey (SASS). The TFS sample includes teachers who leave teaching in the year after the SASS data collection and those who continue to teach either in the same school as last year or in a different school. The purpose of the TFS is to determine how many teachers remained at the same school, moved to another school, or left the profession in the year following the SASS administration.

While NCES is currently collecting the sixth data point on teacher movements, there are no comparable data for the principals who lead American schools. Yet, like teachers, principals are expected to experience large rates of turnover over the next decade as baby boomers reach retirement age. Additionally, there are challenges in retaining and attracting the most highly qualified principals at individual schools. This feasibility study has two components. The first is to test whether the annual proprietary data held by "Quality Education Data" (QED) are sufficiently reliable to be used to develop year-to-year estimates of the rates at which American elementary and secondary school principals change schools and leave the profession. The second component, a survey of principals who leave the profession (leavers) and a survey of principals who change positions (movers), will serve two functions. First, it will confirm or validate the principal status assigned based on the analysis of QED data. Second, it will serve as a field test for a potential new survey component to add to the SASS suite of surveys.

### **Design**

## Volume I

The NCES Common Core of Data (CCD) serves as the sample frame for NCES public school data collections. It is an aggregate school, district, and state universe data collection for public schools in the United States that allows NCES to draw samples of schools or school districts; typically, NCES uses the school directory from this frame to contact schools or districts to request lists of students or teachers to sample for specific surveys. In order to study job transitions within the education labor force, historically, NCES has drawn a sample in one year (e.g., teachers in SASS) and then reinterviewed a subsample of those respondents in the next year to learn about those who stayed in the same job, moved to a different education job, or left the profession (stayers, movers, and leavers) and their reasons for their job decisions. This approach is both resource and labor intensive, and it involves data loss associated with three levels of nonresponse—school or district to get the list, respondent nonresponse in the first year of data collection, and respondent nonresponse in the subsequent round of data collection. While these procedures are necessary to get in-depth information on the educational background, work experience, and reasons for the decision to stay, move, or leave; they may not be necessary to develop annual estimates of movers, leavers, and stayers among elementary and secondary principals. A database that includes annual universe data on the individual staff associated with each school could be used to identify stayers, movers, and leavers by comparing school and principal names across two consecutive years.

Principal names are not collected as part of CCD. However, they are part of QED data. The QED describes its National Education Database™ as being primarily for use by market researchers. The QED reports that its database “...contains comprehensive contact and demographic information for early childhood centers, K-12 schools and higher education institutions.” The QED also reports that its database contains demographic and contact information on every educator in the U.S. Thus, it should be possible to compare two consecutive years of QED data to identify principals who remained at the same school (stayers), who changed schools (movers), and those who left the profession (leavers). Ideally, that information along with the demographic data on the principals would provide the basis for a demographic profile of the principals in the stayers, movers, and leavers subpopulations. However, an attempt at just such an exercise revealed that a number of the desired demographic data fields had missing data for large proportions of the principals. This missing data raises questions about the validity of the QED data on individual educators.

As a first step in examining the data validity, the QED can be linked to the NCES Common Core of Data through the CCD school ID number to compare school, district, and state level aggregates across the two data collections (years  $y$  and  $y+1$ ) and to compare school contact information (and fill in missing contact information in the QED). However, this does not allow for a validation of the QED data entries for principals and the use of those data to categorize each principal as a stayer, mover, or leaver. This can be done through drawing samples of one or more of the three identified principal subpopulations and surveying the sample members to confirm their job status in year  $y+1$ . The survey can also be used to validate the demographic information on principals that is available in QED and to fill in the demographic information for those sample cases where it is missing.

## **Volume I**

Furthermore, given NCES interest in adding a principal follow-up component to SASS, data comparable to that currently collected on teachers' reasons for moving from one school to another or leaving the profession altogether can be collected for principals using only slight modifications of items already developed and tested in the teacher follow-up component of SASS.

Conducting this proposed study in the fall of 2008 provides a unique opportunity to maximize its utility, given the proximity to the most recent administration of SASS. Specifically, principal data derived from a comparison of QED data for school years 2006-07 and 2007-08 yields the stayers, movers, and leavers as of fall 2007. Given that SASS was conducted in the 2007-08 school year, the SASS sample will include a sample of the QED identified 2006-07 to 2007-08 stayers and movers (the possibility of maximizing overlap between the QED samples and the 2007-08 SASS sample will be considered during sample selection). The QED sample of leavers will provide information on the complement of leavers for the year prior to the 2007-08 SASS study year. If the feasibility study validates the use of QED data from consecutive years to identify stayers, movers and leavers, the comparison will be repeated for 2007-08 and 2008-09. Combining this information with data from the 2007-08 SASS will yield background demographic and work experience data and 2007-08 working environment and conditions for the 2007-08 to 2008-09 stayers, movers, and leavers among principals. (SASS did not collect principal names or contact information for principals so it is not possible to do a direct follow-up study of principals using the 2007-08 SASS as the base year.)

Again, assuming the feasibility study validates the use of QED data from consecutive years to identify stayers, movers, and leavers among principals, this process can be repeated annually to identify annual universe based rates of job status transitions among principals. In addition, if the principal surveys of movers and leavers are deemed successful, they can be repeated as a full-scale component of the 2011-2012 SASS (assuming that the collection of names for principal respondents is included in the SASS principal data collection).

Volume II of this submission includes draft questionnaires for principal movers and principal leavers.

The individual survey items for the surveys of principal movers and leavers are derived from existing items from the SASS TFS data collection instruments. Thus, the survey items are modifications of items that have been successfully used in full-scale data collections for teachers or have undergone cognitive testing for use in the 2008-09 SASS TFS. The draft questionnaires include a cross walk to the TFS source items. The screening items and educational attainment specific to principals are the only new data items on the mover questionnaire.

Although the survey items for principals could be tested with a smaller sample size, the validation of the quality of the QED data on school and principal names and the use of these data to adding individual principals to the job status categories of stayers, movers, and leavers requires a sample size sufficient to yield national estimates of the number of principals who are movers and the number who are leavers cross-tabulated by basic

## **Volume I**

school characteristics. These crosstabulations will allow for an analysis of whether the quality of the QED data varies by important school and principal demographic characteristics. In addition, the stayer classification will be evaluated on the basis on the extent of incorrect classifications in the mover and leaver subpopulations. To meet these goals, samples of 1,200 movers and 1,200 leavers are proposed. This sample size is based on the sample size typically used in the NCES Fast Response Survey System (FRSS) to produce national estimates with basic cross-tabulations by school characteristics. A response rate of 85 percent is anticipated.

The two proposed samples will each be random samples of the principals in the identified subpopulation (movers or leavers). Using school level data for the principals' 2006-07 schools, the principals will be stratified by the instructional level (elementary, secondary/combined schools), enrollment size (less than 300 students, 300 to 499, 500 to 999, 1,000 to 1,499, 1,500 or more), and percentage of students eligible for free or reduced-price lunch or percent of students who are minorities (less than 35 percent, 35 to 49 percent, 50 to 74 percent, 75 percent or more) (the variable with the least amount of missing data will be selected from these two). The allocation of the total sample to the primary strata will initially be made in proportion to the aggregate square root of enrollment. Then within the primary strata, principals will be selected with probabilities proportionate to the square root of the school size (i.e., enrollment). The probabilities of selection will be used in developing the weights, with nonresponse bias analyses and adjustments for nonresponse if necessary. Such an allocation is efficient for estimating the proportion of principals in schools with specific characteristics, and will also support estimates of the proportion of principals with specific characteristics (this assumes one head principal per school, as is done in CCD). For a sample of this size, the standard error of an estimated proportion for the total sample is expected to be in the range of .015 to .020. For basic cross-tabulations, the standard errors of estimated proportions may be expected to range from .018 to .040.

The survey of principals who move will be conducted by sending a letter introducing the study and a survey to each principal at the school the principal moved to in 2007-08. The option to complete the survey online using a web link will also be provided. After a three-week period telephone follow-up will begin. The survey of principals who left the profession will be conducted by sending an advance letter to the principal's residence describing the study, and then by calling each respondent for a computer assisted telephone interview. Additional mailings will be used for nonresponse follow-up. (Principal's current mailing addresses and phone numbers will be obtained through a separate research project conducted for NCES by NISS. That project will study the feasibility of using noncredit third-party tracing companies to locate respondents based solely on name, field of employment, and place of employment as a principal in the Spring of 2007.)

## **Public Comment and Consultations Outside the Agency**

The study repeats, designs developed for related teacher questionnaires that have been fielded repeatedly as part of the Schools and Staffing Survey (SASS). As a result, the

## **Volume I**

questionnaires reflect the cumulative input of many experts in the field and past SASS Technical and Content Review Panels.

### **Recruiting and Paying Respondents**

No monetary incentives are planned for this proposed survey.

To recruit for the principal mover survey, each principal will be contacted in the Fall of 2008 at the school identified as the school the principal moved to in 2007-08 using the school name and address from the QED. This survey will start with a letter that introduces the study and a copy of the survey, with a web option, and telephone follow-up for nonresponse as needed.

To recruit for the principal leaver survey the most recent information available is the principal's name and the name and location of the school where the principal worked in the spring of 2007. Through a separate research project, that will also be used to test the feasibility of tracing respondents without the use of social security numbers, NCES will use the services of 2 or 3 noncredit company tracing services to locate current phone numbers and addresses for a sample of the principal leavers. The survey will be conducted as a computer assisted telephone interview (CATI) that will follow an advance letter that will be used to introduce the survey. A mail survey will also be used as part of the nonresponse follow-up activity. When the cases are sent to the tracing companies, a larger sample than needed for the leaver survey will be drawn to ensure a large enough yield for the principal study, due to the fact that the hit or success rate is not anticipated to be perfect. If a high success rate is attained, the cases will be subsampled to yield a nationally representative sample of the proposed size for the principal leaver survey. If fewer cases than expected are found, a nonresponse bias analysis of the identified cases will be conducted and the weights will be adjusted if necessary.

### **Assurance of Confidentiality**

Participation is voluntary and respondents will be given the same confidentiality pledge that is used in SASS, before interviews are conducted. Mathematica Policy Research is conducting this study for the National Center for Education Statistics (NCES) of the U.S. Department of Education. This study is authorized by law under the Education Sciences Reform Act (Public Law 107-279). Your participation is voluntary. Your responses are protected from disclosure by federal statute (P.L. 107-279, Title 1, Part E, Sec. 183). All responses that relate to or describe identifiable characteristics of individuals may be used only for statistical purposes and may not be disclosed, or used, in identifiable form for any other purpose, unless otherwise compelled by law.

All Mathematica staff members working on the project and having access to the data (including monitoring of interviews) are required to sign the NCES Affidavit of Nondisclosure (exhibit 1). In addition, staff members with access to the data are also required to submit the necessary forms and fingerprints for a federal background investigation.

## Volume I

### Justification for Sensitive Questions

Throughout the item selection and development process, effort has been made to avoid asking for information that might be considered sensitive or offensive. The only items not previously used or tested with teachers are the screeners to confirm the respondent is the correct person and is correctly categorized as a mover or leaver and those having to do with the attainment of an advanced degree in education administration. .

### Estimated Response Burden

Each survey is expected to take 13 minutes to complete. Therefore, the estimated respondent burden will be:

Respondent	Estimated Average Response Time-per respondent	Sample Size	Number of respondents	Total
Principal Leavers	13 minutes	1,200	1,020	221 hours
Principal Movers	13 minutes	1,200	1,020	221 hours
Total		2,400	2,040	442 hours

### Estimate of Cost Burden

There is no direct cost to respondents.

### Cost to Federal Government

Cost to the Federal government to conduct the two surveys, process the data, develop weights, prepare a data file, and produce tables, is estimated to be \$230,000.

### Publication Plans and Project Schedule

Experiences with the data collection will be summarized in a research and development report. The report will focus on experiences with collecting data using the QED as a frame for identifying movers and leavers among principals.

**Volume I**  
**Project Schedule**

August 22, 2008	Mathematica finalizes questionnaires
September 1, 2008	NCES provides sample for mover principals with probabilities of selection
September 8, 2008	Mathematica initiates data collection for movers, with mail surveys with a request for a three week turnaround
September 15 – October 17, 2008	Mathematica programs CATI for leaver questionnaire
October 8, 2008	Mathematica starts telephone follow-up on movers
October 24, 2008	NCES provides sample of leaver principals with probabilities of selection
October 29, 2008	Mathematica initiates CATI data collection for leavers
November 17, 2008	Mathematica closes out mover data collection
December 1, 2008	Mathematica starts mail follow-up for leaver nonrespondents
December 19, 2008	Mathematica finalizes editing and weighting of data for movers and provides a preliminary data file with a variable layout to NCES for disclosure review
December 19, 2008 – January 16, 2009	NCES adds data disclosure protections to mover data file and returns file January 16
January 16, 2009	Mathematica closes out leaver data collection
January 23, 2009	Mathematica provides first draft of data tables for movers to NCES
February 1, 2009	Mathematica provides second draft of data tables for movers to NCES
February 3, 2009	Mathematica provides mover data file and documentation to NCES
February 12, 2009	Mathematica finalizes editing and weighting of data for leavers and provides a preliminary data file with a variable layout to NCES for disclosure review
February 12 – February 22, 2009	NCES adds data disclosure protections to mover data file and returns file 2/22/2009
February 26, 2009	Mathematica provides first draft of data tables for leavers to NCES
March 5, 2009	Mathematica provides second draft of data tables for leavers to NCES
March 12, 2009	Mathematica provides mover data file and documentation to NCES