



Considerations for Analysis of SASS and NTPS Data

Module Objectives

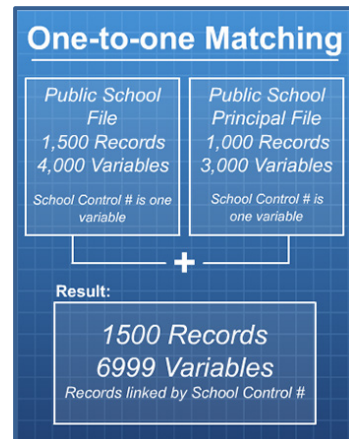
- Describe the interrelated data in NTPS and how that affects analysis
- Describe merging files for analysis
- Discuss the appropriate unit of analysis and how that affects weighting
- Present information on low response rates and nonresponse bias analysis for NTPS
- Describe considerations for analyses of SASS and NTPS data

Data Files and File Merging

- NTPS focuses on flexibility, timeliness, and integration with other ED data which allows for school, principal, and teacher characteristics to be analyzed in relation to one another
- NTPS is the primary source of data on teacher and principal labor markets, and on the state of K-12 school staffing
- Data files can be analyzed separately or merged
- Data files are merged by the NTPS school control number (CNTLNUMS)
- One-to-one or one-to-many matching/merging
- To analyze public and private data together (e.g., all schools), data files are concatenated or stacked
 - Merges are only used with data files within the same sector (public, private)
 - For example – it is not possible to merge private teachers to the public-school file by CNTLNUMS

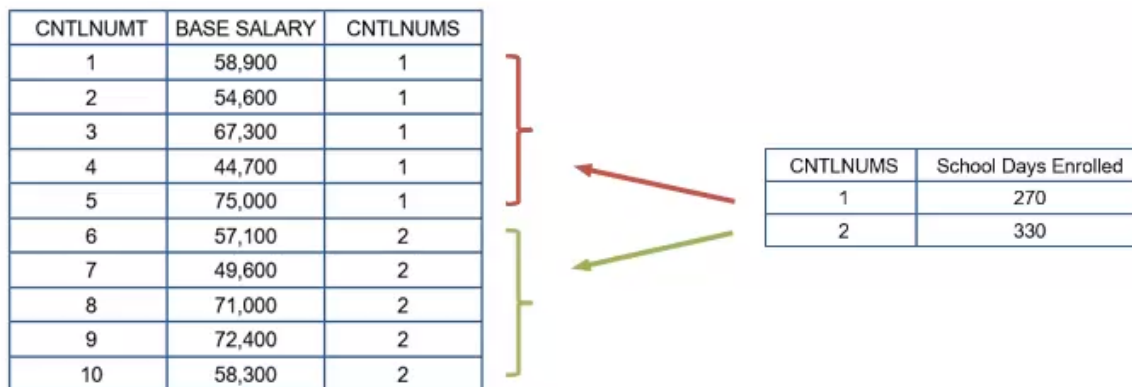
Merging NTPS Data Files (one-to-one matching)

- Each data record is matched by a unique identifying variable (CNTLNUMS – School control number for NTPS)
- The variables from the second data file as listed in the merge statement that do not match are added onto the matched record from the first data file listed in the merge statement
- This would be used to match principals to schools



Merging NTPS Data Files (one-to-many matching)

- Used when the identifying variable is unique in one file, but not unique in the other
- Each data record from one file can be matched with many data records in another file
- This would be used to match teachers to schools



Sample Stata Syntax for Merging Files Within NTPS

Merging Restricted-use Data Files Using the School Control Number (CNTLNUMS)

- Default merge variable is created ("merge")
- Both data files being merged must be sorted by the school control number prior to performing the merge
- Words in italics are meant to be replaced by file or variable names that the user specifies

```

use "filepath\dataset1.dta"
sort CNTLNUMS
save "filepath\dataset1.dta", replace
use "filepath\dataset2.dta"
sort CNTLNUMS
save "filepath\dataset2.dta", replace
merge CNTLNUMS using "filepath\dataset1.dta" /*merges the two files */
[optional statement: merge CNTLNUMS using "filepath\dataset1.dta", nolabel]
drop if _merge==1 /*specifies dataset1 as unit of analysis*
save "filepath\newfilename.dta", replace /*saves a new file keeping all dataset1
records and only matching dataset2 records*/
    
```

Sample SPSS Syntax for Merging Files Within NTPS

Both data files being merged must be sorted by the variable listed in the “by” statement prior to performing the merge

- In SPSS value labels are attached automatically during the extraction process
- Words in italics are meant to be replaced by the file or variable names that the user specifies

When merging any of the NTPS files together, the school’s control number, CNTLNUMS, is used to merge data files. The SPSS syntax is provided below

get file = “*filepath\dataset1.sav*”.

sort cases by CNTLNUMS(A).

save outfile = “*filepath\dataset1.sav*”.

get file = “*filepath\dataset2.sav*”.

sort cases by CNTLNUMS(A).

save outfile = “*filepath\dataset2.sav*”.

match files file = “*filepath\dataset1.sav*” * merges the two files and specifies dataset1 as unit of analysis* /table “*filepath\dataset2.sav*” /by CNTLNUMS.

save outfile = “*filepath\mergeddatafile.sav*”. *creates new merged filename*

Sample SPSS Syntax for Merging Files Within NTPS (Continued)

- Specify the two data files; sort each by the school control number
- dataset1 is the unit of analysis; dataset2 is the school file
- Match files is the command for merging
- /table specifies the second data file (‘dataset2’)
- /by provides the merge variable (CNTLNUMS)

Sample SAS Syntax for Merging Data Files and Attaching Variable Labels

Merging Restricted-use Data Files Using the School Control Number (CNTLNUMS)

- Both data files being merged must be sorted by the variable listed in the “by” statement prior to performing the merge
- Comments to explain lines of code are contained within “/* */”
- Words in italics are meant to be replaced by the file or variable names that the user specifies

```
proc sort data = dataset1;  
by CNTLNUMS;  
run;  
proc sort data = dataset2;  
by CNTLNUMS;  
run;  
data newfilename; /*creates new merged file name*/  
merge dataset1 /table (in=a) dataset2; /* merges the two files and  
specifies dataset1 as unit of analysis*/  
by CNTLNUMS;  
if a = 1; /*keeps all dataset1 records and only matching dataset2 records*/  
run;
```

Unit of Analysis

- Each data file in NTPS has a different final weight
- If you are only using one data file use the final weight in that file
 - Example: Public School Teacher data file
 - Choose TFNLWGT as your weight
- If you have merged files, your merged dataset will have more than one final weight
 - Example: A merged school and teacher data file
 - Choose weights carefully to match your analyses
- It is extremely important to identify the appropriate unit of analysis and use the correct weight, or else your analytic approach may be misguided and have misleading results
- Usually, considering the target population is the best way to select which weight variable to use
- Considering which group the analysis is being generalized to represent will help you identify the unit of analysis for your research

Low Unit Response Rates in NTPS Data Files

- Sample surveys require an acceptable level of response
- Unit response rate: Percentage of eligible, sampled cases that are “complete interviews”
- NCES Statistical Standards require a nonresponse bias analysis when the level of response for a survey sample falls below the threshold rate of 85 percent
- NCES does not publish estimates when unit response rate is below 50 percent for a particular subgroup without conducting additional nonresponse bias analysis
- Public School Teacher: Teachers in city schools had response rates below 50 percent in the 2015-16 NTPS
- Public Schools, Principals, and Teachers in Maryland: The 2017-18 NTPS data for Maryland is not published because response rates fell below 50 percent for all of the surveys
- More details on unit response rates for public and private (2017-18 only) surveys are provided in each of the NTPS First Look Reports and the Restricted-Use User’s Manuals

Low Item Response Rates

- Biased data are responses that may not be representative of all the target population
- NCES Statistical Standards require a nonresponse bias analysis when the level of response to a survey item falls below the threshold rate of 85 percent
- A basic overview of the bias analysis results is provided in each of the NTPS First Look Reports
- For a more detailed explanation of response rates and nonresponse bias analyses, the Restricted-Use NTPS User’s Manual for each survey is the best resource

Use Caution Generalizing Beyond Defined Sampling Strata

- NTPS is designed to support sampling strata estimates as documented in "[NTPS Sample Design, Weight, and Variance](#)"
- Subpopulations are subject to high standard errors that should be taken into consideration before reporting whether results are statistically significant
- It is imperative to compute the correct standard errors for all estimates computed to avoid Type I or Type II errors
- Respondents should never be identified

From SASS to NTPS

- NTPS is redesigned from the Schools and Staffing Survey (SASS)
- From 2017-18 onward, NTPS is designed to produce state-level estimates for public schools, and also includes private schools
- Maintains many of the SASS teacher, school, and principal items through core content and a series of rotating modules
- Modules rotate every other cycle

Changes Over Time

- NTPS content and procedures differ from SASS, which also underwent changes over the course of more than 20 years
 - For example, NTPS uses Jackknife weighting and SASS uses Balanced Replicate weighting.
 - SASS had 88 replicate weights on each file and NTPS has 200.
- NTPS is a repeated cross-sectional survey and is not designed for longitudinal analysis
 - Similar to SASS, a small proportion of the same schools may be included in the NTPS sample from one data collection cycle to the next but is not sufficient for any form of longitudinal or pseudo-longitudinal analysis
 - Samples of principals and teachers are independent in each data collection of NTPS
- Background variables such as locale codes may change over time
- Survey variables also change over time

Module Summary and Resources

- Describe the inter-related data in NTPS and how that affects analyses
- Describe merging files for analysis
- Discuss the appropriate unit of analysis and how that affects weighting
- Present information on low response rates and nonresponse bias analysis for NTPS
- Describe considerations for analyses of SASS and NTPS data

Module Resources

- [NTPS Sample Design, Weights, and Variance](#)

First Look Reports

- [Characteristics of Public Elementary and Secondary **Schools** in the United States: Results From the 2015–16 National Teacher and Principal Survey](#)
- [Characteristics of Public Elementary and Secondary School **Teachers** in the United States: Results From the 2015–16 National Teacher and Principal Survey](#)
- [Characteristics of Public Elementary and Secondary School **Principals** in the United States: Results From the 2015–16 National Teacher and Principal Survey](#)
- [Characteristics of Public Elementary and Secondary **Schools** in the United States: Results From the 2017-18 National Teacher and Principal Survey](#)
- [Characteristics of Public Elementary and Secondary School **Teachers** in the United States: Results From the 2017-18 National Teacher and Principal Survey](#)
- [Characteristics of Public Elementary and Secondary School **Principals** in the United States: Results From the 2017-18 National Teacher and Principal Survey](#)