NIS-4 Analyses
Sample Stata Syntax

This syntax was tested using Stata version 14 with the Survey Replication based Standard Errors package.
Because of the complex sample design and use of paired jackknife replicate weighting (JK2):
You MUST use the weights when working with NIS-4 data.

Data File

Use the example syntax below to set up your dataset.

To analyze NIS-4 in Stata (and to obtain the correct standard errors) you must first install a Stata package called **svr**: Survey Replication based Standard Errors. The install command below will download the package from the Boston College Statistical Software Components (SSC) archive.

The data file in your NDACAN NIS-4 #147 dataset is named DS_147.dta. Place your DS_147 Stata file in a folder with this pathname: "C:\temp\data\DS_147.dta" or change the path and/or file in the **use** command(s) for the syntax examples below.

The **list** command will provide output verifying you have declared the weights and method properly. Some **svr** commands are based on official Stata **svy** commands with which you may be familiar.

```
. ssc install svr
. use c:\temp\data\DS_147.dta
. svrset clear
. svrset set pw CHAWT rw CHAWT1-CHAWT62
. svrset set meth jk2
. svrset list
```

Frequencies

The example syntax below first creates a binary indicator (flag) that is positive 1 when "Not countable as physically abused under Harm Standard" is true. It then uses the **svrtotal** procedure to obtain frequencies (incidence estimates) for variable PAH (Physical Abuse: Harm Standard). It will output weighted frequency, standard error, confidence interval, and design effects (Deff). Missing data are deleted casewise by default.

```
. gen PAHno
. replace PAHno=1 if PAH ==0
. svrtotal PAH PAHno
```

Because some **svr** commands are based on official Stata **svy** commands, for instructions to change your syntax or output options, see the Survey Data Reference Manual in the **Resources** section below.

Cross Tabulations

The example syntax below uses the **svrtab** procedure to obtain cross tabulations (subgroup incidence estimates) for variable PAH (Physical Abuse: Harm Standard) by variable CHSEX (Child Sex). It will output formatted weighted frequencies (incidence estimates), standard errors of weighted counts, confidence intervals, chi-square statistics, and design effects (Deff). Missing data are deleted casewise by default.

> . **svrtab** PAH CHSEX, count se ci deff fcou(%12.2f) fse(%12.2f) fci(%12.2f)

For instructions to change your syntax or output options, see the Survey Data Reference Manual in the **Resources** section below.

Rates

The formula below calculates rates for overall incidence per 1,000 children. This can be done and recorded in your software package, in MS Excel, or manually.

> Rate = (weighted frequency estimate) / (population denominator)

Table 2-1. *Population Totals for Computing Rates per 1,000 for All Maltreated Children and for Specific Subgroups* in the *NIS-4 PUF Manual (User's Guide)* lists the available population denominators, variables 537 to 560 in the *Codebook.* For overall incidence use ALLPOP (Population Total for All Maltreated Children) as the population denominator. For cross tabulation subgroup calculations, use the appropriate census denominator variable in the data file.

Logistic Regression

The example syntax below first creates a binary indicator so the probability modeled is PAH = 'Not countable as physically abused under Harm Standard'. It then uses the **svrmodel** procedure to run a bivariate logistic regression analyses. The **logit** command will output the intercept (_cons), estimate, and standard error, with t-tests, p values, and confidence intervals. The **logistic** command will output odds ratio point estimates and standard errors, with t-tests, p values, and confidence intervals. Missing data are deleted casewise by default.

> . **gen** PAHno= 0
> . **replace** PAHno= 1 if PAH == 0
> . **svrmodel** PAHno CHSEX, cmd(**logit**)
> . **svrmodel** PAHno CHSEX, cmd(**logistic**)

For instructions to change your syntax or output options, see the Logit and Logistic Stata references in the **Resources** section below.

Resources

Winter, N. (2004, August). *Replication methods for analysis of complex survey data in Stata* [Slide Presentation], Cornell University. Retrieved from http://faculty.virginia.edu/nwinter/papers/Winter.NASUG.pdf

Stata: Data Analysis and Statistical Software. (2016, September 27). I*nstalling programs from SSC.* Retrieved from http://www.stata.com/support/ssc-installation/

Stata: Data Analysis and Statistical Software. (2016, September 27). *Survey Data Reference Manual.* Retrieved from http://www.stata.com/bookstore/survey-data-reference-manual/

Stata: Data Analysis and Statistical Software. (2016, September 27). *Logit - logistic regression, reporting coefficients.* Retrieved from http://www.stata.com/manuals14/rlogit.pdf

Stata: Data Analysis and Statistical Software. (2016, September 27). *Logistic - logistic regression, reporting odds ratios.* Retrieved from http://www.stata.com/manuals14/rlogistic.pdf

Lee, E. S., Forthofer, R. N., & Lorimor, R. J. (1989). *Analyzing complex survey data.* Sage University papers series 71: Quantitative applications in the social sciences. Sage.

This document was prepared by NDACAN Research Analyst Diane Wach, MA, MSEd, LPC. Email questions to NDACANsupport@cornell.edu.