

CHILDREN'S HEALTHY LIVING (CHL) CENTER OF EXCELLENCE

Complex Survey Weights for Data Analysis

Developed by the CHL Data Work Group for use in the CHL Pacific Region

For further information, please contact: CHL Data Center Children's Healthy Living Center of Excellence Ag Sciences 302, 1955 East West Road, Honolulu, HI 96822 Email: <u>chldata@hawaii.edu</u> Website: <u>www.chl-pacific.org</u>

This project is supported by the Agriculture and Food Research Initiative Grant no. 2021-68012-35899 and Grant no. 2018-69001-27551 from the USDA National Institute of Food and Agricultural Science Enhancement Coordinated Agricultural Program and Grant no. 1P20GM139753-01A1 from the Centers of Biomedical Research Excellence (COBRE).



United States Department of Agriculture National Institute of Food and Agriculture

BACKGROUND

A community-randomized intervention was conducted in the 5 jurisdictions of Alaska, Hawai'i, American Samoa, Guam and CNMI (called intervention jurisdictions or INT). Child and community level data were collected at baseline (time 1) and post-intervention at 24 months (time 2) and 72 months (time 3). A prevalence survey was conducted in the Freely Associated States (FAS) of Federated States of Micronesia (FSM), Republic of Palau, and the Republic of the Marshall Islands (RMI). In all jurisdictions, CHL data were collected within community clusters. For more details on the CHL's study design, see paper by <u>Wilkens et al</u>.¹

Due to this complex survey design, specific adjustments must be made for analyses with CHL individual-level data. For studies including 2 or more CHL jurisdictions in analyses (e.g., studies including the whole CHL Pacific region), there must be adjustments on the clustering of participants in communities (variable: community_cluster) within jurisdiction strata (variable: jurisnum) and with survey weights (variable: wt_anthro_adj). For studies including only 1 jurisdiction in analyses, there must be adjustments on the clustering of participants in communities only (variable: community_cluster).

Sample program codes for SAS, R, and STATA and instructions for SPSS to include this complex sample design are shown in the next sections.

For any questions, please email the CHL Data Center at chldata@hawaii.edu.

¹Wilken LR, Novotny R, Fialkowski MK, et al. Children's Healthy Living (CHL) Program for remote underserved minority populations in the Pacific region: rationale and design of a community randomized trial to prevent early childhood obesity. *BMC Public Health*. 2013;13:944. Published 2013 Oct 9. doi:10.1186/1471-2458-13-944

SAMPLE PROGRAM CODE FOR SAS, R, & STATA

Highlighted are codes for adding complex survey weights.

1. <u>SAS</u>

```
/*Survey design for analyses that include 2 or more jurisdictions*/
proc surveyreg/surveylogistic data=dataset;
class covariates;
model outcome = exposures covariates;
weight wt_anthro_adj;
strata jurisnum;
cluster community_cluster;
run;
/*Survey design for analyses that include only 1 jurisdiction*/
proc surveyreg/surveylogistic data=dataset;
class covariates;
model outcome = exposures covariates;
cluster community_cluster;
run;
```

2. <u>R</u>

By using the svyglm function from the survey package and setting up your survey design, # you can adjust for the survey weights, strata, and clusters. This provides more accurate # estimates and standard errors given the survey design.

Survey design for analyses that include 2 or more jurisdictions
design <- svydesign(ids =~ id, cluster = ~community_cluster, strata = ~jurisnum, weights
= ~wt_anthro_adj, data = dataset)</pre>

Survey design for analyses that include only 1 jurisdiction
design <- svydesign(ids =~ id, cluster = ~community_cluster, data = dataset)</pre>

outcome <- svyglm(exposure ~ covariate + covariate, design = design, family =
"binomial")</pre>

NOTE:

To use *svydesign*, you must install the survey package and load the library with the same name.

If the package is already installed in your environment then you can load it: install.packages("survey") library(survey)

3. <u>STATA</u>

// svyset [psu] [weight] [, strata(varname) fpc(varname)]

// Survey design for analyses that include 2 or more jurisdictions
svyset community_cluster [pweight = wt_anthro_adj], strata(jurisnum)

// Survey design for analyses that include only 1 jurisdiction svyset community_cluster

svy: regress outcome exposures covariates

INSTRUCTIONS FOR PREPARING COMPLEX SAMPLES IN <u>SPSS</u>

ta Uni	titled1 [[DataSet0]	- IBM SP	SS Statistics Da	ta Editor									
<u>F</u> ile	<u>E</u> dit	<u>V</u> iew	<u>D</u> ata	Transform	<u>A</u> nalyze	<u>G</u> raphs	<u>U</u> tilities	E <u>x</u> tens	sions	<u>W</u> in	dow	<u>H</u> elp		
					Po <u>w</u> er	Analysis			>					Qs
					Meta /	Analysis			>			1		
					Report	s			>					
	vai	r	var	var			A			ar	Va	ar	var	var
1					Bayes	ian Statist	ics		>					
2					Tables				>					
3					Comp	aro Moane	and Propo	rtions	\$					
5					Conor	al Linear N	and Tropo	nuons	Ś					
6					Gener		lodel							
7					Gener	alized Line	ar Models		>					
8					Mi <u>x</u> ed	Models			>					
9					<u>C</u> orrela	ate			>					
10					Regres	ssion			>					
12					L <u>o</u> gline	ear			>					
13					Neural	Networks			>					
14					Classi	fy			>					
15					Dimen	sion Redu	ction		>					
16					Scale				>					
17					Nonna	rametric T	octe		\$					
18					<u>N</u> onpa		6515		ĺ					
20					Foreca	as <u>t</u> ing			,					
21					<u>S</u> urviva	al			>					
22					M <u>u</u> ltip	le Respons	se		>					
23					ジ Missin	g <u>V</u> alue Ar	nalysis							
24					Mult <u>i</u> p	le Imputati	on		>					
25					Comp	ex Sample	es		>	🔳 Se	elect a	Sample		
26					🖶 Simula	tion			Г		oparo f	or Anal	veie	
28					Quality	v Control			, L		epare i		yərə	
29					Snatia	l and Tem	noral Mode	lina	>	<u>123 F</u> r	equenc	:les		
30					Direct	Markating		g	\$	<u>Пара</u> е	escripti	ves		
31					Direct	mar <u>k</u> eung	∇		<i>`</i>	🛃 <u>C</u> r	osstab	s		
32							*			12 <u>R</u> a	atios			
33										CS G	eneral I	inear N	lodel	
34											nintin f	Degrad	ion	
36											gistic I	Regress	sion	
37										<u> </u>	dinal R	legress	on	
38										🔡 Co	o <u>x</u> Regr	ression.		
39														
40														
_	<	_												
Data	View	Variab	le View											
_														

Select Complex Samples \rightarrow Prepare for Analysis.

Use the Analysis Preparation Wizard to set you complex survey weights.

🕼 Analysis Preparation Wizard		×
Welcome to the Analysis Prepa The Analysis Preparation Wizard sample weights and other inform Your selections will be saved to	ration Wizard I helps you describe your complex sample and choose an estimation method. You will be asked to pro ation needed for accurate estimation of standard errors a plan file that you can use in any of the analysis procedures in the Complex Samples Option.	vide
	What would you like to do? • <u>C</u> reate a plan file Choose this option if you have sample data but have not created a plan file. • <u>E</u> dit a plan file Choose this option if you want to add, remove, or modify stages of an existing olan. • <u>If</u> you already have a plan file you can skip the Analysis Preparation Wizard and go directly to a of the analysis procedures in the Complex Samples Option to analyze your sample.) Inny
	< <u>B</u> ack <u>N</u> ext > Finish Cancel Help	

Name your Complex Sample (CS) Analysis Plan and Save.

🔖 Save Data As			×
Look <u>i</u> n: Doo	cuments	× 👔 🚮	
Custom Offi Fax My SAS Fil	ce Templates es otebooks		
Zoom	ocuments		
File <u>n</u> ame:	CHL Complex Design Weights		Save
Save as <u>t</u> ype:	CS Analysis Plan (*.csaplan)	~	Cancel
	Store <u>File</u> To Repository		<u>H</u> elp

Click Next.

😭 Analysis Preparation Wizard		\times
Welcome to the Analysis Prepa	ration Wizard	
The Analysis Preparation Wizar sample weights and other inform	d helps you describe your complex sample and choose an estimation method. You will be asked to provid Nation needed for accurate estimation of standard errors	de
Your selections will be saved to	a plan file that you can use in any of the analysis procedures in the Complex Samples Option.	
	What would you like to do?	
	● <u>C</u> reate a plan file	
<u> </u>	Choose this option if you have sample data but have not created a plan file.	
	O <u>⊑</u> dit a plan file	
	Choose this option if you want to add, remove, or modify stages of an existing File: Browse plan.	
X of H.	if you already have a plan file you can skip the Analysis Preparation Wizard and go directly to any of the analysis procedures in the Complex Samples Option to analyze your sample.	4
	< Back Next > Finish Cancel Help	

The next window shows what Sample Weight to include depending on if you are including <u>2 or more jurisdictions</u> OR <u>only 1 jurisdiction</u> in your analyses. Carefully follow the next page to decide what Sample Weight to include. **The following step is for analyses including <u>2 or more jurisdictions</u>: Include *wt anthro adj* **as your Sample Weight**

🕼 Analysis Preparation Wizard		×
Stage 1: Design Variables In this panel you can select varial You can also provide a label for the Welcome Stage 1 Design Variables Estimation Method Summary Completion	bles that define strata or clusters. A sample weight variable must be selected in the first stage. The stage that will be used in the output. Variables:	
	 ♣ wic ♣ income ♣ income2 ✔ income3 ♣ owob ♣ CommNum ✓ Stage Label: 	
	Eack Next > Finish Cancel Help	

The following step is for analyses including <u>only 1 jurisdiction</u>:

Include *wt_constant* as your Sample Weight, *wt_constant* is a default variable set to 1 for all observations. If you do not have this variable, you can create it.

🕼 Analysis Preparation Wizard			\times
Stage 1: Design Variables In this panel you can select variab	les that define strata or clusters.	. A sample weight variable must be selected in the first stage.	
You can also provide a label for th Welcome Stage 1: CHL Jurisdicti Design Variables Estimation Method Summary Completion	e stage that will be used in the or Variables: Community Sex Trace_new_omb JurisName V t_anthro V wt_anthro_adj time A agemos bt_snap foodassi Vic income income income commNum	Sample Weight:	
	< <u>B</u> ack <u>N</u> ext >	Finish Cancel Help	

**The remaining steps are applied to analyses including EITHER 2 or more jurisdictions OR only 1 jurisdiction.

Select WR (sampling with replacement).



Click Next.

😭 Analysis Preparation Wizard									×	
Stage 1: Plan Summary	Stage 1: Plan Summary									
This panel summarizes the plan so far. The next step is the Completion panel.										
Velcome	Summar	/ :								
Stage 1	Stage	Label	Strata	Clusters	Weights	Size	Method			
 Design Variables Estimation Method 	1	(None)	jurisnum	community	_wt_anthro_a dj	(n/a)	WR			
Completion										
		1 18:				144 - 14				
	File: C:\Users\Rica Dela Cruz\Desktop\CHL Regional Design Weighted.csaplan									
			1							
		< <u>B</u> ack	<u>N</u> ext >	Finish Ca	ncel Help					

Click Finish.

Analysis Preparation Wizard		×
Completing the Analysis Wizard		
You have provided all of the informati	on needed to create a plan.	
You can use the plan file in any Con	plex Samples analysis procedure when you are ready to analyze the data.	
 Welcome Stage 1 Design Variables Estimation Method Summary Completion 	What do you want to do? Save your specifications to a plan file Paste the syntax generated by the Wizard into a syntax window To close this wizard, click Finish.	
	< Back Next > Finish Cancel Help	

Once complex sampling design has been prepared, you can follow the steps below to continue to run your desired analyses:

ta Un	titled1 [[DataSet0]	- IBM SP	SS Statistics Da	ata Editor										
<u>F</u> ile	<u>E</u> dit	<u>V</u> iew	<u>D</u> ata	Transform	<u>A</u> nalyze	<u>G</u> raphs	<u>U</u> tilities	E <u>x</u> tens	ions	<u>W</u> in	dow	<u>H</u> elp			
					Power	Analysis			>			4			0
					Meta	Analysis			>			1		•	
					Report	ts			>						
	vai		var	var			A			ar	Va	ar	var		var
1					Bayes	ian Statist	ics		>						
2					Tables	6			>						
4					Comp	are Means	and Propo	rtions	>						
5					Gener	al Linear M	lodel		>						
6					Conor	alized Line	nouel		ĺ						
7					Gener	all <u>z</u> ed Line	ar wodels								
8					Mixed	Models			>						
9					<u>C</u> orrel	ate			>						
10					<u>R</u> egre	ssion			>						
12					L <u>og</u> lin	ear			>						
13					Neura	l Networks			>						
14					Classi	fy			>						
15					Dimen	ision Redu	ction		>						
16					Scale				>						
17					Nonna	rametric T	oete		>						
18					<u>T</u> onpa		6313		ĺ						
20					Foreca	as <u>t</u> ing									
21					Surviv	al			>						
22					M <u>u</u> ltip	le Respon	se		>						
23					ジ Missin	g <u>V</u> alue A	nalysis								
24					Mult <u>i</u> p	le Imputati	on		>						
25					Comp	lex Sample	es		>	📾 Se	elect a	Sample			
26					🖶 Simula	ation									
27					Qualit	v Control			>	E F	epare f	or Analy	/SIS		
29					Santia		noral Mada	ling		123 <u>E</u> r	equenc	ies			
30					Spatia	and rem	porar wode	ing		<u>D</u> e	escripti	ves		_	
31					Direct	war <u>k</u> eting			,	<u>C</u> r	osstab	s			
32							V			12 Ra	atios				
33									- Г				la dat		
34									-+	<u> G</u> e	eneral L	Linear IV	iodel		
35									-+	<u></u>	gistic F	Regress	ion		
37										🐰 <u>O</u> r	dinal R	legressi	on		
38										🔡 Co	o <u>x</u> Regi	ression.			
39									L						
40															
	<														
Data	View	Variab	le View												
Dutt															

Select Complex Samples \rightarrow Choose statistical method for analysis.

Find your saved complex sample analysis plan (.csaplan) & Select Use default file.

Complex Samples Plan for General Linear Model X							
Plan							
<u>File:</u> p\CHL Regional Design Weighted.csaplan							
If you do not have a plan file for your complex sample, you can use the Analysis Preparation Wizard to create one. Choose Prepare for Analysis from the Complex Samples menu to access the wizard.							
Joint Probabilities							
Joint probabilities are required if the plan requests unequal probability WOR estimation. Otherwise, they are ignored.							
● Use default file (C:\\CHL Regional Design Weighted.sav)							
O An <u>o</u> pen dataset							
SampleDataset.sav [DataSet1]							
○ <u>C</u> ustom file							
File: Browse							
Continue Cancel Help							

Complex Samples Gener	al Linear Model	×
Variables:	Dependent Variable:	Model
CommName		Statistics
<pre> a biii_category id </pre>	<u>F</u> actors:	Hypothesis Tests
jurisdiction	•	Estimated Means
sex sex		Sav <u>e</u>
a race_new_omb	Covariates:	Options
A JurisName		
🖋 wt_anthro		
💑 time		
🛷 agemos		
🗞 ebt_snap	Subpopulation	
o foodassi	Variable:	
🔊 WIC	★	
sincome2	Catagony	
income3	Category.	
💑 owob 🗸 🗸		
OK	Paste Reset Cancel He	elp

Choose your variables of interest and run your model.