Subject: Pooled Survey MeLogit Syntax Query
Posted by cmergenthaler on Tue, 26 Jul 2016 09:16:27 GMT
View Forum Message <> Reply to Message

Dear DHS experts,

I would be grateful to have your insight as to whether my Stata version 14 syntax is logical for the following analysis:

I am attempting to run an melogit model on pooled data from 13 surveys (4 of which are repeated surveys from two countries), in which a number of individual and household variables (ie age category, religion, wealth index, and child health) are used to predict a binary outcome variable of care being provided for a sick child. All data used are collected from moms only.

Based mostly on Tom Pullman's very helpful advice from the thread: "sampling stratification of the Cameroon DHS 2011, 2004, 1998 and 1991 data", I have borrowed some of the following syntax:

```
generate weight = v005/10000000
egen clusters=group(survey v021)
egen strata = group(survey v024 v025)
svyset clusters [pweight=weight], strata(strata) singleunit(centered)

*Then my survey melogit command is:
svy: melogit child_cough_care v476_1 i.v013 i.religion i.v106 i.v190 HIV_accept
```

Can someone kindly clarify whether I have weighted this correctly, and whether it properly accounts for fixed and random effects?

Thank you in advance!

Kind regards,

Christina

Subject: Re: Pooled Survey MeLogit Syntax Query
Posted by cmergenthaler on Tue, 26 Jul 2016 15:03:01 GMT
View Forum Message <> Reply to Message

Apologies, I should have clarified which surveys I am using. They are the individual recode files (merged with some household variables noted above) of:
Cambodia 2005
Ethiopia 2005
Ethiopia 2011
India 2005-06
Kenya 2008-09

Kyrgyzstan 2012
Malawi 2010
Namibia 2006-07
Namibia 2013
Nigeria 2008
Tajikistan 2012
Zambia 2013-14
Zimbabwe 2005-06

Kind regards,
Christina

Subject: Re: Pooled Survey MeLogit Syntax Query
Posted by Bridgette-DHS on Wed, 27 Jul 2016 13:54:51 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Stata Specialist, Shireen Assaf:

The variables and the svyset you specified are correct. However, the model you specified is not a mixed model even though you used melogit. It is in fact a logit model since you did not specify any random intercept or random slope. The melogit would revert to a logit model if you do not specify any random terms. If you are interested in a mixed model with survey having random intercepts, you could do this without the svy as follows:
melogit child_cough_care v476_1 i.v013 i.religion i.v106 i.v190 HIV_accept [pw=weight] || survey:

To use svy with melogit you need to specify two weights. I assume in your case you are interested in a model with perhaps surveys having a random intercept and clusters nested in surveys with random intercept, i.e perhaps something like.
svy: melogit child_cough_care v476_1 i.v013 i.religion i.v106 i.v190 HIV_accept [pw=weight] || survey: || cluster:

In that case you need to generate a country/survey weight and you can see the last response from Dr. Tom Pullum on two ways you can do this:
http://userforum.dhsprogram.com/index.php?t=msg&goto=9698&&srch=%22country+weight%22#msg_9698

After you have your country/survey weight and the weight you computed from v005, you can use svy with melogit as shown in the example in the Stata 14 Multilevel Mixed Effect manual page 83 for how this can be done when there are two weights: http://www.stata.com/manuals14/me.pdf

However, I suspect for the relatively low number of surveys in your analysis (13 surveys), that this model will not converge. You may want to consider performing a meta analysis and looking into the metan command written for Stata (http://www.stata.com/support/faqs/statistics/meta-analysis/).

Alternatively you could just fit a logistic model with a dummy variable for survey with no constant term.

Subject: Re: Pooled Survey MeLogit Syntax Query
Posted by cmergenthaler on Wed, 27 Jul 2016 19:09:55 GMT
View Forum Message <> Reply to Message

Thank you for this very informative response. I was able to make the first suggestion work, using i.survey as a dummy variable. However, I ran into a few new problems when attempting to correct the svy:melogit model, which I understand is the best practice.

I hope I understood your suggestion well to create weights for both the survey and cluster level in order to be able to add in random effects. Clearly I am doing one or several things wrong still, so I'll continue to be grateful for your advice.

From myy original syntax I kept:

generate weight = v005/10000000
egen clusters=group(survey v021)

Then I created a variable called surveyweight for each survey: (female population ages 15-49 of country survey year) / (# females 15-49 in survey)

Then attempted to redo the svyset command correctly:
svyset survey, weight(surveyweight) || clusters, weight(v005)

Then ran the svy: melogit command once again as per your advice and the stata manual to which referred (thank you!):
svy: melogit child_cough_care v476_1 i.v013 i.religion i.v106 i.v190 HIV_accept [pweight=weight] || survey: || clusters:
Then received error message: weights not allowed

So I removed [pweight=weight], although I don't know that this is correct to do, and then ran the model without it:
svy: melogit child_cough_care v476_1 i.v013 i.religion i.v106 i.v190 HIV_accept || survey: || clusters:
And as you rightly predicted, it never converged...

I cannot understand why the weight [pweight=weight] was not allowed, but I suppose that since it still does not converge even after removal of the weight, I should go for the metaan step. Can you kindly clarify?

Thank you in advance for any further assistance!

Christina

Subject: Re: Pooled Survey MeLogit Syntax Query
Posted by Bridgette-DHS on Wed, 27 Jul 2016 20:56:17 GMT
View Forum Message <> Reply to Message

Following is another response from Shireen Assaf:

In fact if you use svy for your model you do not need to add the [pw=weight] as well as the weights are added in the svyset. For the svyset instead of v005 I would put the weight variable you generated from v005/1000000. This will not make a difference I suspect. Yes I think it does not converge due to the small number of surveys you have. You can proceed with meta analysis or simply use an svy: logistic model with i.survey dummy with no constant. If you use the logistic model the svyset would be the one you specified in your first post with only one weight.

Subject: Re: Pooled Survey MeLogit Syntax Query
Posted by cmergenthaler on Thu, 28 Jul 2016 08:41:24 GMT
View Forum Message <> Reply to Message

Thank you very much - by excluding the pweight and using the v005/10000000 as the weight for the cluster, the model converges! Much appreciated.
For others if it's helpful, that's:

generate weight = v005/10000000
egen clusters=group(survey v021)
svyset survey, weight(surveyweight) || clusters, weight(weight)
svy: melogit child_cough_care v476_1 i.v013 i.religion i.v106 i.v190 HIV_accept || survey: || clusters:

This has helped me to move forward with my analysis.