Subject: Melogit and Weights
Posted by Yawo on Thu, 23 Jul 2020 13:38:17 GMT
View Forum Message <> Reply to Message

Hello,

I am trying to fit a multilevel model examining HIV risk behaviors among those living with HIV using a pooled data from 23 countries.

My model nests individuals (level 1) into clusters (level 2). For melogit, I am aware I need to set weights at the two levels of my analysis, yet DHS only provides one weight.

From my review of forum here (https://userforum.dhsprogram.com/index.php?t=msg&goto=13522&&srch=multilevel+weights#msg_13522), I understand I can svyset my data, using the following:

gen weight=hiv05/1000000
gen weight2=1
svyset idhspsu, weight(weight) strata(idhsstrata) singleunit(centered) || _n, weight(weight2)

When I run my basic melogit, <svy: melogit condomless_spouse sex || idhspsu: >

I get the following error: "weights in variable weight not constant within groups defined by: idhspsu
an error occurred when svy executed melogit"

I am sure I may not be doing something wrong here. I would be very grateful if any of you could point me in te right direction.

Thanks in advance of your assistance.

cY

Subject: Re: Melogit and Weights
Posted by Bridgette-DHS on Wed, 05 Aug 2020 12:53:04 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Senior Analysis & Research Manager, Shireen Assaf:

I suspect your error is due to pooling 23 surveys. The code you provided is used for one survey. You may need to create a new psu and strata variable that takes into account you are using a pooled dataset. So you should run the command below and use those variables in your svyset.

egen psupool= group(idhspsu v000)
egen stratapool= group(strata v000)

Also, you actually have 3 levels with the survey as the third level. I suggest you search the forum for pooled analysis as this was discussed extensively by users and you may get more answers and suggestions there.

Subject: Re: Melogit and Weights
Posted by Yawo on Thu, 06 Aug 2020 04:47:15 GMT
View Forum Message <> Reply to Message

Brigette.

Thanks. I followed your example and generated the new psu/strata, but getting the same error. Below is a listing of some PSU's and the non-constant weights, therein:

Thanks - cY

. list sample idhspsu psupool weight weight2 v005 in 200/250

	sample	idhspsu	psupool	weight	weight2	v005
200.	Angola 2015	2401000011	11	3.262758	1	2.964534
201.	Angola 2015	2401000011	11	2.894893	1	2.836539
202.	Angola 2015	2401000011	11	2.894893	1	2.836539
203.	Angola 2015	2401000011	11	3.262758	1	2.964534
204.	Angola 2015	2401000011	11	3.262758	1	2.964534
205.	Angola 2015	2401000011	11	2.894893	1	2.836539
206.	Angola 2015	2401000011	11	2.894893	1	2.836539
207.	Angola 2015	2401000011	11	3.262758	1	2.964534
208.	Angola 2015	2401000011	11	3.262758	1	2.964534
209.	Angola 2015	2401000011	11	3.262758	1	2.964534
210.	Angola 2015	2401000011	11	3.262758	1	2.964534
211.	Angola 2015	2401000011	11	3.262758	1	2.964534
212.	Angola 2015	2401000011	11	2.894893	1	2.836539
213.	Angola 2015	2401000011	11	3.262758	1	2.964534
214.	Angola 2015	2401000011	11	2.894893	1	2.836539
215.	Angola 2015	2401000011	11	2.894893	1	2.836539
216.	Angola 2015	2401000011	11	2.894893	1	2.836539
217.	Angola 2015	2401000011	11	3.262758	1	2.964534
218.	Angola 2015	2401000011	11	3.262758	1	2.964534
219.	Angola 2015	2401000011	11	2.894893	1	2.836539

```
220. | Angola 2015   2401000012       12    .396283       1    .41734 |
221. | Angola 2015   2401000012       12     .40295       1   .437145 |
222. | Angola 2015   2401000012       12     .40295       1   .437145 |
223. | Angola 2015   2401000012       12     .40295       1   .437145 |
224. | Angola 2015   2401000012       12    .396283       1    .41734 |
     |-----------------------------------------------------------------|
225. | Angola 2015   2401000012       12    .396283       1    .41734 |
226. | Angola 2015   2401000012       12    .396283       1    .41734 |
227. | Angola 2015   2401000012       12     .40295       1   .437145 |
228. | Angola 2015   2401000012       12    .396283       1    .41734 |
229. | Angola 2015   2401000012       12     .40295       1   .437145 |
     |-----------------------------------------------------------------|
230. | Angola 2015   2401000012       12     .40295       1   .437145 |
231. | Angola 2015   2401000012       12     .40295       1   .437145 |
232. | Angola 2015   2401000012       12     .40295       1   .437145 |
233. | Angola 2015   2401000012       12     .40295       1   .437145 |
234. | Angola 2015   2401000012       12    .396283       1    .41734 |
     |-----------------------------------------------------------------|
235. | Angola 2015   2401000012       12     .40295       1   .437145 |
236. | Angola 2015   2401000012       12    .396283       1    .41734 |
237. | Angola 2015   2401000012       12     .40295       1   .437145 |
238. | Angola 2015   2401000012       12     .40295       1   .437145 |
239. | Angola 2015   2401000013       13    .586726       1   .595693 |
     |-----------------------------------------------------------------|
240. | Angola 2015   2401000013       13    .547652       1   .604053 |
241. | Angola 2015   2401000013       13    .586726       1   .595693 |
242. | Angola 2015   2401000013       13    .547652       1   .604053 |
243. | Angola 2015   2401000013       13    .586726       1   .595693 |
244. | Angola 2015   2401000013       13    .547652       1   .604053 |
     |-----------------------------------------------------------------|
245. | Angola 2015   2401000013       13    .547652       1   .604053 |
246. | Angola 2015   2401000013       13    .586726       1   .595693 |
247. | Angola 2015   2401000013       13    .547652       1   .604053 |
248. | Angola 2015   2401000013       13    .547652       1   .604053 |
249. | Angola 2015   2401000014       14    .503401       1   .477899 |
     |-----------------------------------------------------------------|
250. | Angola 2015   2401000014       14    .465081       1   .491306 |
     +-----------------------------------------------------------------+
```

.

Subject: Re: Melogit and Weights
Posted by Bridgette-DHS on Thu, 06 Aug 2020 12:43:39 GMT

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

Your listing indicates that within each value of "psupool" you have exactly two different values of the weight variable. This is what you would see if two clusters are sharing the same ID code. The weights should be exactly the same within a cluster. I suggest that you take a closer look at all the id information for the cases within one of these values of psupool. Somehow or other, I suspect that you actually have two distinct clusters when you think you have one, and the two clusters correspond with the different values of the weight variable.

Subject: Re: Melogit and Weights
Posted by Yawo on Thu, 06 Aug 2020 16:15:30 GMT

Thanks for your responses.

I think we have double two psu's because I appended male and female cases in my pooled dataset. As the listing below shows, the weight values are constant within each psu by gender.

. list sample sex weight weight2 psupool stratapool in 115431/115440

```
      +-------------------------------------------------------------------+
      |      sample     sex    weight   weight2   psupool   strata~l |
      |-------------------------------------------------------------------|
115431. | Lesotho 2014    male   .515554         1      4753        336 |
115432. | Lesotho 2014  female    .48581         1      4753        350 |
115433. | Lesotho 2014  female    .48581         1      4753        350 |
115434. | Lesotho 2014    male   .515554         1      4753        336 |
115435. | Lesotho 2014  female    .48581         1      4753        350 |
      |-------------------------------------------------------------------|
115436. | Lesotho 2014    male   .515554         1      4753        336 |
115437. | Lesotho 2014  female    .48581         1      4753        350 |
115438. | Lesotho 2014  female    .48581         1      4753        350 |
115439. | Lesotho 2014  female    .48581         1      4753        350 |
115440. | Lesotho 2014  female   .981849         1      4754        340 |
      +-------------------------------------------------------------------+
```

to resolve this issue, should I recreate the psu and strata variables by taking gender into account, ie:

egen psupool= group(idhspsu v000 sex)
egen stratapool= group(strata v000 sex)

sincerely, Yawo

Subject: Re: Melogit and Weights
Posted by Bridgette-DHS on Fri, 07 Aug 2020 13:29:00 GMT
View Forum Message <> Reply to Message

Following is another response from DHS Research & Data Analysis Director, Tom Pullum:

Men and women have different weights because they have different levels of nonresponse and therefore different corrections for nonresponse. I didn't think about you having both women and men in the same file.

However, I don't think doubling the number of clusters and strata would be the correct way to adjust for that effect. Would you get the same error message if you included sex as a covariate in the model?

Subject: Re: Melogit and Weights
Posted by Yawo on Fri, 07 Aug 2020 14:04:40 GMT
View Forum Message <> Reply to Message

Thanks again, I am able to avoid the error if I recreate the psu and strata taking sex/gender into account
as below:

egen psupool= group(idhspsu sample sex)
egen stratapool= group(idhsstrata sample sex)

Here is my model: svy: melogit condomless_spouse sex || psupool:

Here again is a brief listing showing constant weights within each psu.

. list sample sex idhspid idhspsu idhsstrata psupool stratapool weight in 1/20

	sample	sex	idhspid	idhspsu	idhsstrata	psupool	strata~l	weight
1.	Angola 2015	male	2401	00010003	3	2401000001	2401000018	1	321.979475
2.	Angola 2015	male	2401	00010001	1	2401000001	2401000018	1	321.979475
3.	Angola 2015	male	2401	00010002	1	2401000001	2401000018	1	321.979475
4.	Angola 2015	male	2401	00010008	3	2401000001	2401000018	1	321.979475
5.	Angola 2015	male	2401	00010004	1	2401000001	2401000018	1	321.979475
6.	Angola 2015	male	2401	00010014	1	2401000001	2401000018	1	321.979475
7.	Angola 2015	male	2401	00010011	2	2401000001	2401000018	1	321.979475
8.	Angola 2015	male	2401	00010003	1	2401000001	2401000018	1	321.979475
9.	Angola 2015	male	2401	00010023	2	2401000001	2401000018	1	321.979475
10.	Angola 2015	male	2401	00010001	3	2401000001	2401000018	1	321.979475
11.	Angola 2015	male	2401	00010012	3	2401000001	2401000018	1	321.979475
12.	Angola 2015	male	2401	00010026	1	2401000001	2401000018	1	321.979475
13.	Angola 2015	female	2401	00010001	02	2401000001	240100018	2	18 1.001989
14.	Angola 2015	female	2401	00010002	03	2401000001	240100018	2	18 1.001989
15.	Angola 2015	female	2401	00010002	02	2401000001	240100018	2	18 1.001989
16.	Angola 2015	female	2401	00010008	02	2401000001	240100018	2	18 1.001989
17.	Angola 2015	female	2401	00010003	02	2401000001	240100018	2	18 1.001989
18.	Angola 2015	female	2401	00010014	02	2401000001	240100018	2	18 1.001989
19.	Angola 2015	female	2401	00010026	02	2401000001	240100018	2	18 1.001989
20.	Angola 2015	female	2401	00010006	01	2401000001	240100018	2	18 1.001989

+----------+

Subject: Re: Melogit and Weights
Posted by Yawo on Mon, 10 Aug 2020 15:01:33 GMT
View Forum Message <> Reply to Message

As a follow-up.

One way I think we could resolve this issue is to create two data files..one for men and the other for women and run separate models by gender.

Is there no other way to have one master file for both women and men, and control for gender?

Thanks..cY

Subject: Re: Melogit and Weights
Posted by BillC on Tue, 22 Dec 2020 02:12:00 GMT
View Forum Message <> Reply to Message

Hi all,

I am trying to replicate the code in "Multilevel Modeling Using DHS Surveys...:Methodological Report 27" on a combined men and women data set from Afghanistan and also trying to adjust the weights of men and women in this combined dataset.

But am running into the dreaded "weights in variable wt2 not constant within groups defined by: v001" error when attempting to run melogit. When I run the diagnostic code at the end to see how many cases have different weights, I get dif==15,433.

Any suggestions on how to do this correctly?

Thanks!

Here is my code (apologies for its length):

do "C:\Users\Afghanistan\Appending men and women datasets.do"

*This is adapted from the Zimbabwe code at the end of the DHS Report #27

* a_c_h completed clusters by strata

```
gen a_c_h=.
levelsof v022, local(lstrata)
foreach ls of local lstrata {
tab v021 if v022==`ls', matrow(T)
scalar stemp=rowsof(T)
replace a_c_h=stemp if v022==`ls'
}
* A_h total number of census clusters by strata; from Table A2 of Afghanistan Final Report; pg 311.
gen A_h = 0
*Urban # EAs
replace A_h = 1870 if v022 == 1
..etc...
replace A_h = 119 if v022 == 68

* M_h average number of households per cluster by strata - from Table A2 of Afghanistan Final Report; pg 311.
gen M_h = 0

*urban - avg # households per EA
replace M_h = 239.8 if v022 == 1
...etc.....
replace M_h = 181.5 if v022 == 68

* m_c total number of completed households - Section 1.9, pg 5 of Afghan Final Report
gen m_c= 24395

* M total number of households in country - Table A1, pg 310 in Afghan Final Report
gen M = 4269415

* S_h households selected per stratum - Section A3, pg 312 of Afghan Final Report
gen S_h = 27


*adjusting weights of men and women in combined dataset
gen wtfactor=0
replace wtfactor=(16727000/29461) if sex==2   //dividing population of women by number interviewed in 15-49 yrs
replace wtfactor=(17686000/10760) if sex==1  //dividing population of men by number of men interviewed

gen wt=v005/1000000
gen newwt=wt*wtfactor
label variable newwt "Population adjusted sample weight"
gen DHSwt = newwt/1000000


* Steps to approximate Level-1 and Level-2 weights from Household or Individual Weights
```

```
* Step 1. De-normalize the final weight, using approximated normalization factor
gen d_HH = DHSwt * (M/m_c)

*Step 2. Approximate the Level-2 weight
* f the variation factor
gen f = d_HH / ((A_h/a_c_h) * (M_h/S_h))

scalar alpha=0.5

gen wt2 = (A_h/a_c_h)*(f^alpha)

gen wt1 =  d_HH/wt2


* Svyset

svyset v001, strata(v022) weight(wt2) singleunit(centered) || _n, weight(wt1)

svy: melogit v474 i.sex ||v001:, or

**for testing
bysort v001: gen dif = 0
replace dif = 1 if v001 == v001[_n-1] & wt2 != wt2[_n-1]
browse if dif == 1
count if dif==1
```

Subject: Re: Melogit and Weights
Posted by Yawo on Tue, 22 Dec 2020 09:32:11 GMT
View Forum Message <> Reply to Message

BillC:

I am not sure if this would help, but ensure that you do NOT append men/women datasets together before svysetting the data

Try the weighting procedure on the men and women data separately.

best, Yawo

Subject: Re: Melogit and Weights
Posted by Bridgette-DHS on Tue, 22 Dec 2020 14:47:07 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

Yes, we too have found that the methods in MR27 lead to this problem. The solution we use is to replace the level 2 weights by their geometric mean within each cluster (the difference from the arithmetic mean is negligible). The following Stata lines will calculate "rev_wt2", a revised wt2.

```
gen ln_wt2=log(wt2)
egen mean_wt2=mean(ln_wt2), by(v001)
gen rev_wt2=exp(mean_wt2)
drop ln_wt2
```

For the arithmetic mean you would just have
```
egen rev_wt2=mean(wt2), by(v001)
```

This may not be the final recommendation for this issue--stay tuned.

Subject: Re: Melogit and Weights
Posted by BillC on Tue, 22 Dec 2020 16:32:45 GMT
View Forum Message <> Reply to Message

Thanks Tom (and Yawo) for your quick replies. (It seems I am not the only one slaving away during the holidays.... :-))

Your suggestion worked Tom- many thanks for it. I tried weighting by both the arithmetic and geometric mean and the estimates in my test regression were very close (but that's probably due to the fact I have a large sample).

Am curious to know more about the issues you have noticed with the MR27 dataset. Can you shed some more light on that - and on the enigmatic statement at the end to stay tuned? Does that mean you will revise the code in DHS Report 27?

Once again, thanks for the reply. Merry Christmas and Happy New year to all - may 2021 be an Annus Mirabilis for all.

Subject: Re: Melogit and Weights
Posted by Bridgette-DHS on Wed, 23 Dec 2020 11:37:45 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

Am glad the substitution of the geometric/arithmetic means (of the cluster level weights) worked. The authors of MR27 (led by Mahmoud Elkasabi, plus Ruilin Ren and me) are working on a

potential follow up article for journal submission. The main purpose is to look at the sensitivity of estimates to the choice of alpha. We may put a link or addendum on the website. Check back in a couple of months if you haven't seen something!

Happy Holidays!

Subject: Re: Melogit and Weights
Posted by BillC on Tue, 12 Jan 2021 23:26:24 GMT
View Forum Message <> Reply to Message

Hi Tom,

Another quick clarification on weight for multilevel analyses. To recap, I am combining the mens and womens datsets for a multilevel analyses. (Code is from DHS Report #27 (Multilevel Modeling Using DHS Surveys: A Framework to Approximate Level-Weights)

Two questions:
Q1: Do I need to adjust v005 BEFORE I run this code below?

```
gen wt=v005/1000000
gen DHSwt = v005/1000000
* Steps to approximate Level-1 and Level-2 weights from Household or Individual Weights
* Step 1. De-normalize the final weight, using approximated normalization factor
gen d_HH = DHSwt * (M/m_c)
*Step 2. Approximate the Level-2 weight
* f the variation factor
gen f = d_HH / ((A_h/a_c_h) * (M_h/S_h))
scalar alpha=0.5
gen wt2 = (A_h/a_c_h)*(f^alpha)
gen wt1 =  d_HH/wt2
* Svyset
svyset v001, strata(v022) weight(wt2) singleunit(centered) || _n, weight(wt1)
.....
```

Q2:if yes, then the adjustment factor for women = (population of 15-49yr women in country)/number of 15-49 yr women sampled; and for men = (population of men 15-49 yrs in country)/number of men 15-49 yrs sampled - as per the code below?

```
*adjusting weights of men and women in combined dataset
gen wtfactor=.
replace wtfactor=(xxxxxxxx/29461) if sex==2   //dividing population of women 15-49 yrs in country by number interviewed in 15-49 yrs
replace wtfactor=(yyyyyyy/10760) if sex==1//dividing population of men 15-49 yrs in country by number of men interviewed
label variable wtfactor "Multiplication factor for wt(v005)"
```

gen wt=v005/1000000
gen newwt=wt*wtfactor
label variable newwt "Population adjusted sample weight"
gen DHSwt = newwt/1000000

(then run code above...)

Many thanks!

Bill

Subject: Re: Melogit and Weights
Posted by Bridgette-DHS on Wed, 13 Jan 2021 14:27:58 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Senior Sampling Specialist, Mahmoud Elkasabi:

Yes, your approach is correct. You need to de-normalize v005 and mv005 separately as you described, then you can proceed with Step 2. However, I believe the code you shared is missing the de-normalization for the men weights. See a revised code below.

*adjusting weights of men and women in combined dataset
gen newwt =.
replace newwt =(v005/1000000)* (xxxxxxxx/29461) if sex==2
replace newwt =(mv005/1000000)* (yyyyyyy/10760) if sex==1

*Step 2. Approximate the Level-2 weight
* f the variation factor
.....

Subject: Re: Melogit and Weights
Posted by BillC on Wed, 13 Jan 2021 14:57:25 GMT
View Forum Message <> Reply to Message

Thank you Mahmoud. Appreciate the quick reply.

Bill

Subject: Re: Melogit and Weights
Posted by Alessandra tangianu on Sun, 19 Feb 2023 14:58:04 GMT

View Forum Message <> Reply to Message

Dear All,

I am trying a similar approach but I would require the domestic violence weights. Would I have to adjust the d005 in the same way as v005 and mv005 before I run the multi-level weights code as suggested by Billy C?

Subject: Re: Melogit and Weights
Posted by Bridgette-DHS on Tue, 21 Feb 2023 13:25:55 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS staff member, Tom Pullum:

Yes, if you want to run a multilevel model that includes DV variables, you should use d005 as the weight and go through the same procedure to separate the individual-level and cluster-level weights as you would with the other weights.