Subject: Couple data: sampling and weights
Posted by Isha on Fri, 13 Aug 2021 06:13:12 GMT
View Forum Message <> Reply to Message

Hello,

I am using the couple recode file for India (NFHS4, 2015-16). I had three questions related to sampling/weighting.
1. I understand that the couple recode file is created from a sub-sample of currently married women and their husbands and was not sure how nationally representative this data is. Further, can the analyses reflect state-wise differences (of course after using state weights).
2. Also, I wanted to create an indicator of "desired sex ratio at the community level". I had created such an indicator from the women's recode file by creating the ratio of ideal number of boys per 100 girls for each rural and urban area of all states. This value of desired sex ratio at the community level was assigned to all women belonging to that community. However, can one create the same for couples in each community (state x place of residence) in the couple recode file? I was not sure if it is appropriate for each case to take on a community level value as the couple data is not representative of each administrative state (or can it be?).
3. In Stata, when I used the "svy" command for a regression, I get this error: "Missing standard errors because of stratum with single sampling unit". Could you explain how I can use weights or account for the sample design while carrying out a regression in such a dataset which seems to have this issue of a stratum with a single PSU.

Thank you so much!

Best,
Isha

Subject: Re: Couple data: sampling and weights
Posted by Isha on Mon, 16 Aug 2021 17:13:43 GMT
View Forum Message <> Reply to Message

Hello,

I wanted to follow up to my last message on using the svy command in Stata and getting an error an error that there is a stratum with just one PSU. This is the command I used to create the weight:

svyset [pw=wgt], psu (v021) strata (v023)

It would be good to know if I should create a survey weight differently or ignore this error.

Hope to hear from you regarding my other queries, i.e. using state weights for analysis and using a community level variable in couple data.

Thank you!

Best,
Isha

Subject: Re: Couple data: sampling and weights
Posted by Shireen-DHS on Mon, 16 Aug 2021 18:25:08 GMT
View Forum Message <> Reply to Message

Hello,

All you need to do in this case is add the option singleunit(centered) as follows:

svyset [pw=wgt], psu (v021) strata (v023) singleunit(centered)

For couples data, please use the men's weight (mv005). You should divide this weight by 1000000.

I am uncertain about your question about the desired sex ratio at the community level. Perhaps you can take the average ratio in each cluster? You can get this with the following code assuming your sex ratio variable is idealratio. Be sure to recode non-numeric responses as missing before using variables related to ideal number of children/boys/girls.

bysort v001: egen idealratioclus=mean(idealratio)

Thank you.

Best,
Shireen Assaf
The DHS Program

Subject: Re: Couple data: sampling and weights
Posted by Isha on Wed, 18 Aug 2021 18:01:41 GMT
View Forum Message <> Reply to Message

Dear Shireen,

Thanks A LOT for your prompt help. Your advice worked instantly.

Regarding my other question on ideal sex ratio, let me see if I can try multilevel modeling. I found the DHS guide (in pdf) to that on this forum and will use it. I will get back if I have any questions. I will also look into the command you suggested.

Hope you have a great week!

Best,
Isha

Subject: Re: Couple data: sampling and weights
Posted by Isha on Fri, 20 Aug 2021 16:39:27 GMT
View Forum Message <> Reply to Message

Dear Shireen,

Thank you for your advice. I have got a lot of help from you all experts here. I also searched the forum to help answer some more questions, which was useful.
As of now, I have completed most of my analysis and wanted to check that these statements make sense in terms of the sampling and do not look analytically illogical:
- My aim is to see how state-level norms (level-2) effect of predictors (level-1) of son preference in India. My binary dependent variable is a "couple level" variable of son preference (merged responses of women and their husbands).
- I am carrying out a multilevel logistic regression using couple data from India (CR recode file)
- I have used national men's weights for all the level-1 variables (these include individual level attitudes and socioeconomic background)
- For level two, I am only using state (since that is how I set up my argument). The state variable (v024) includes some smaller administrative units but I do not want to drop them to avoid missing data).
- At level two for state, I assign all observations the same value (=1) since all states (approx for smaller administrative units) have an equal probability of being selected in the final sample.
- I ran the code ("melogit") and the results look believable (I cross checked with a logistic regression model).

I hope all this makes sense for couple data and the way it is sampled. I wanted to check since it is a couple recode file for which there are fewer resources to validate.

Thank you for the bysort idealsex ratio code. That was useful!

Best,
Isha

Subject: Re: Couple data: sampling and weights
Posted by Isha on Wed, 25 Aug 2021 20:47:14 GMT
View Forum Message <> Reply to Message

Dear Shireen,

I hope all is well. I am following up to my last query on using weights to carry out multilevel analysis with couple data. It seems that when I use weights at both levels to carry out multilevel

logistic regression, Stata does not consider them as the results are completely the SAME when I run the same command without weights. Hence, I thought to use the svy command instead of separate weights for the two levels.

For my multilevel analysis, I am using couples at level-1 (mens' weight) and state (v024, weight=1) at level 2. But the svyset command I have been using for all other analysis is: svyset [pw=wgt], psu (v021) strata (v023) singleunit(centered)
This svy command does not take state into account and I was wondering if I could create a different svy command with state (and use that same command to weight the multilevel logistic regression data). Currently, Stata tells me that the svy command is not allowed with this analysis but other social scientists do use it. So I was wondering if it makes sense to create a svy command which considers state into the design and could be used for all final analysis.

In other words, can we assume that the sampling/analysis is two level with couples NESTED in states? (instead of using cluster and strata).

Thanks a lot!

Best,
Isha

Subject: Re: Couple data: sampling and weights
Posted by Shireen-DHS on Wed, 25 Aug 2021 21:11:20 GMT
View Forum Message <> Reply to Message

Dear Isha,

I think the following DHS report may answer your questions on multi-level models and weights: https://dhsprogram.com/pubs/pdf/MR27/MR27.pdf

Best,
Shireen
