Subject: Appropriate stratification variable for NFHS data
Posted by econ_stata on Fri, 04 Sep 2015 18:17:43 GMT
View Forum Message <> Reply to Message

I am working with the NFHS-3 survey KR dataset in Stata, studying the determinants of child health at the all-India level, and have a doubt about the appropriate stratification variable for the survey declaration. Looking at various answers on the Forum, I see recommendations by experts for both

1. v022, "sample stratum number" (73 distinct values)

e.g, by Trevor-DHS here

   http://userforum.dhsprogram.com/index.php?t=msg&goto=326
8&S=c572d650be37c4375bff4d100ab9ced0&srch=NFHS+strat a#msg_3268

2. a grouping of v024, "state" (29 values) with v025, "type of place of residence" (2 values), i.e.,

 egen float stratum = group(v024 v025), resulting in 58 distinct values,

e.g., from Ruilin Ren here

   http://userforum.dhsprogram.com/index.php?t=msg&goto=213
4&S=c572d650be37c4375bff4d100ab9ced0&srch=NFHS+strat a#msg_2134

I was using option 2 so far, but revisiting this issue, I checked the NFHS publication

'National Family Health Survey (NFHS-3), 2005-06: India: Volume I', section on Sample Design, p61, (copy uploaded)

and it seems the stratification scheme is more complex than just the simple grouping of State & place of residence.

I would really appreciate a definitive answer to this from one of the experts familiar with the NFHS data- I think this would serve many researchers now & in future as well.

File Attachments
1) 2SamplingDesign_from_NFHS_FRIND3-Vol1AndVol2 copy.pdf,
downloaded 514 times

Subject: Re: Appropriate stratification variable for NFHS data
Posted by Bridgette-DHS on Fri, 18 Sep 2015 18:20:31 GMT
View Forum Message <> Reply to Message

See below the answer from Senior DHS Sampling Specialist, Ruilin Ren.

The exact stratification of the NFHS-III was very complicated, it is state specific. We do not have

the complete information for the exact stratification. I am not aware of the variable HV022 with 73 codes, it might be inserted to the data file much later (on 2010). I will check with other colleagues to figure out what HV022 stands for in NFH-III.

Since the exact stratification is not available, as usual, we always suggest to use HV024 (state) cross HV025 (urban/rural) as approximate stratification. Since there are 8 cities in India which have slum areas, the slum areas were oversampled in NFHS-III. If you want and if your analysis is sensitive to slum and non-slum areas, you can further split the HV024 by HV025 stratification by adding SHSLUMC (variable indentifies slum areas defined by census) at HH level, or using V024 by V025 by SSLUMC at individual level. The eight slum cities are as follows:

- Delhi in the state of Delhi
- Kolkata in the state of West Bangal
- Hyderabad in the state of Andhra Pradesh
- Chennai in the state of Tamil Nadu
- Indore in the state of Madhya Pradesh
- Merrut in the state of Uttar Pradesh
- Mumbai in the state of Maharashtra
- Nagpur in the state of Maharashtra

Subject: Re: Appropriate stratification variable for NFHS data
Posted by econ_stata on Wed, 23 Sep 2015 12:24:25 GMT
View Forum Message <> Reply to Message

Thank you very much, Dr. Ren (and Bridgette!) for your very helpful & informative answer.

Subject: Re: Appropriate stratification variable for NFHS data
Posted by Bridgette-DHS on Tue, 29 Sep 2015 18:14:18 GMT
View Forum Message <> Reply to Message

You are welcome.