Subject: Re: Pooled Cross sections
Posted by cbdolan on Thu, 23 Jun 2016 14:22:05 GMT
View Forum Message <> Reply to Message

Thanks for the detailed follow up and paper link. Both were helpful.

I think I have something wrong with the way I constructed the pooled weights based on the results of my descriptives. The N's shouldn't be this dissimilar. (Note: I previously merged the files with the spatial files so I'm using ADM1_CODE as the province level variable).

I annotated the code to clarify the steps. Please let me know if I've missed a step.

tab ADM1_CODE[iweight=NW]

| ADM1_CODE | Freq. | Percent | Cum. |
|---|---|---|---|
| Bandundu | 329,552,409 | 16.16 | 16.16 |
| Bas-Congo | 90101292.7 | 4.42 | 20.58 |
| Equateur | 264,708,100 | 12.98 | 33.56 |
| Kasai-Occidental | 173,724,497 | 8.52 | 42.07 |
| Kasai-Oriental | 231,513,280 | 11.35 | 53.42 |
| Katanga | 215,915,193 | 10.59 | 64.01 |
| Kinshasa | 182,595,002 | 8.95 | 72.96 |
| Maniema | 72641458.3 | 3.56 | 76.53 |
| Nord-Kivu | 135,628,336 | 6.65 | 83.18 |
| Orientale | 189,207,489 | 9.28 | 92.45 |
| Sud-Kivu | 153,913,256 | 7.55 | 100.00 |
| Total | 2.0395e+09 | 100.00 | |

. tab ADM1_CODE

| ADM1_CODE | Freq. | Percent | Cum. |
|---|---|---|---|
| Bandundu | 272,736 | 12.63 | 12.63 |
| Bas-Congo | 118,666 | 5.49 | 18.12 |
| Equateur | 300,174 | 13.90 | 32.02 |
| Kasai-Occidental | 195,252 | 9.04 | 41.06 |
| Kasai-Oriental | 234,033 | 10.84 | 51.90 |
| Katanga | 259,660 | 12.02 | 63.92 |
| Kinshasa | 156,412 | 7.24 | 71.17 |
| Maniema | 132,302 | 6.13 | 77.29 |
| Nord-Kivu | 140,773 | 6.52 | 83.81 |
| Orientale | 203,672 | 9.43 | 93.24 |
| Sud-Kivu | 145,920 | 6.76 | 100.00 |

Total | 2,159,600 100.00

I did the following to set up the pooled weights:

```
use "Y:\4_DHS_BirthRecode\CDBR61FL.dta"
*Original weight in DHS : v005 (which should preferably be divided by 1000000)
generate n_v005=(v005/1000000)
*note this is the population of 15-49 in DRC (2013) from United Nations, Department of Economic and Social Affairs, Population Division (2015). World Population Prospects: The 2015 Revision, custom data acquired via website.
generate P1549=16167000
*note this is the sample size from the individual recode file of women 15-49 interviewed
generate n1549=18827
*Country specific weight :CSW= P1549/n1549 (population aged 15-49 in the country / sample size of )
generate CSW=(P1549/n1549)
*New weight
generate NW=n_v005*CSW
file Y:\4_DHS_BirthRecode\n_CDBR61FL.dta saved
clear

use "Y:\4_DHS_BirthRecode\CDBR50FL.dta"
*Original weight in DHS : v005 (which should preferably be divided by 1000000)
generate n_v005=(v005/1000000)
*note this is the population of 15-49 in DRC (2013) from United Nations, Department of Economic and Social Affairs, Population Division (2015). World Population Prospects: The 2015 Revision, custom data acquired via website.
generate P1549=13201000
*note this is the sample size from the individual recode file of women 15-49 interviewed
generate n1549=9995
*Country specific weight :CSW= P1549/n1549 (population aged 15-49 in the country / sample size of )
generate CSW=(P1549/n1549)
*New weight
generate NW=n_v005*CSW
file "Y:\4_DHS_BirthRecode\n_CDBR50FL.dta saved
clear

use "Y:\4_DHS_BirthRecode\n_CDBR61FL.dta"
append using "Y:\4_DHS_BirthRecode\n_CDBR50FL.dta"


*generate weight: see code at top
*make unique strata values by region/urban-rural )
egen stratum=group(ADM1_CODE v025)
```

*tell stata the weight (using pweights for robust standard errors, cluster (psu), and strata
svyset [pw=NW],psu(v021)strata(stratum)
*prefix regrss with "svy:stata will now know how to weight your data and compute the right standard errors */