
Subject: Re: variables for svyset in Stata (Bangladesh 2011)
Posted by [mmr-UMICH](#) on Wed, 04 Feb 2015 03:58:29 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thank you, Trevor.

I will cut and paste some lines whenever requires from original message and try to clarify concretely the reasons behind that error/warnings and its possible solution:

```
**now analyzing using poisson for the lower 2 wealth quintiles**  
svy: poisson csect v190 if v190<3
```

The above svy: command is not recommended [and not correctly handle the domain concept] as "if v190<3" in svy: subsets the data (i.e., also deleting design information [aside: full sample design information is important for correctly calculating the sampling errors]) prior to run the poisson regression.

This svy: command uses such a subset data and as a result analysis sample (i.e. "estimation sample" in Stata wording) lacks one strata and 101 PSUs (see below output (cut and paste):

```
-----start-----  
(running poisson on estimation sample)
```

Survey: Poisson regression

```
Number of strata = 19 Number of obs = 3645  
Number of PSUs = 499 Population size = 3867.0893  
Design df = 480  
----- end -----
```

We have to create a variable, say: mydomain = 1 if v190 < 3, otherwise, mydomain = 0, then use svy command:

```
svy, subpop(mydomain): poisson csect v190
```

I hope this run will not encounter such issue and also does not require singleunit(centered)* svyset option. And the output will show the same # of obs and population size, but changed others stats such as # of strata, PSUs and degrees of freedom (df).

I verified that strata 5 and 11 have 5 and 23 PSUs respectively; so this svy, subpop(): that form "analytic" domain/subpopulation/subgroup will not be an issue of singleton-strata from full sample data.

*note that singleunit(method) is kind of practically recommended for "analytic" subgroup and/or subclass analysis which sometimes encounter singleton-strata. This specification also appropriately calculates the degrees of freedom, which is prim important for statistical inferences, e.g, confidence intervals and p-values estimation.

Thank you all again.

Moshiur Rahman
