
Subject: Stata Help

Posted by [bpanda](#) on Tue, 19 Jan 2016 22:16:30 GMT

[View Forum Message](#) <> [Reply to Message](#)

Hi,

I have used the child recode (KR) data file for Bangladesh (2011 DHS) in one of my papers. The statistical analysis uses the stata "SVY" command to account for survey design and sampling weights to report the unbiased regression coefficients and the appropriate linearized standard errors. We recently got a request to do some additional regressions that are not supported by the svy command. In that case, will it be fine to conduct these additional regressions using the sample weight (v005) (i.e. pweight=v005) and clustering the standard error by primary sample units?

I would really appreciate your help.

Many thanks
Bibhu

Subject: Re: Stata Help

Posted by [Reduced-For\(u\)m](#) on Wed, 20 Jan 2016 01:25:25 GMT

[View Forum Message](#) <> [Reply to Message](#)

In general, that should be fine. In fact, I think that a) these estimates will be very, very close to the ones you would get using "svy", which you can confirm by re-running your original regressions that way; b) they should be more "conservative" estimates, in the sense that, by ignoring the stratification, they are likely to generate standard error and p-value estimates that are weakly too big... which is to say, if you get significance there, you are fine. The estimates will be unbiased for sure, because only the weighting needs to be accounted for to assure unbiasedness - getting appropriately sized SE/p-values requires clustering and stratification, but ignoring the stratification should make almost 0 difference (and like I said, you can check that by re-doing the estimates you've already done with "svy").

For the record -I promise I am not the referee that is making you do the extra analyses (smiley face), but if I were, I would be fine with this approach.

Subject: Re: Stata Help

Posted by [Bridgette-DHS](#) on Wed, 20 Jan 2016 13:23:51 GMT

[View Forum Message](#) <> [Reply to Message](#)

Here is another response from Senior DHS Stata Specialist, Tom Pullum:

You should be able to use weights and clusters with syntax like this: regress y x [pweight=v005], cluster(v001). You cannot make the stratum adjustment if you cannot use svy, but that's all you

will lose.

Stratification makes the sample more efficient. If you ignore the strata, the standard errors will tend to be slightly larger than they should be (that is, slightly larger than they would be if you included the strata adjustment). You can check this by running svy for an estimation command which does allow svy, and doing two runs, one with and one without the strata option, and comparing the standard errors, confidence intervals, and p values for the estimates.

Thus, if you ignore strata, the confidence intervals will tend to be a little too wide and the test statistics will tend to be a little too close to zero. That's good--it means your inferences will be a little on the conservative side, similar to having a slightly smaller sample than you actually have. That's preferable to the reverse, having an artificially inflated sample size, which is what happens if you ignore the clusters. Ignoring the cluster adjustment is generally more damaging, in terms of increased risk of Type I error, than ignoring the stratum adjustment.

Using sample weights will make the estimates unbiased. The cluster and stratum adjustments have no impact on bias, i.e. no effect on the estimates themselves, only on their standard errors. In Stata terminology, the cluster and stratum adjustments produce "robust" standard errors.

Subject: Re: Stata Help
Posted by [Reduced-For\(u\)m](#) on Wed, 20 Jan 2016 21:20:56 GMT
[View Forum Message](#) <> [Reply to Message](#)

I still get that feeling of relief when the other reviewers say the same thing I did. Thanks Tom (and Bridgette)!

Subject: Re: Stata Help
Posted by [Bridgette-DHS](#) on Thu, 21 Jan 2016 13:16:08 GMT
[View Forum Message](#) <> [Reply to Message](#)

You are welcome.

Subject: Re: Stata Help
Posted by [bpanda](#) on Mon, 01 Feb 2016 23:11:44 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thank you very much, Reduced-For(u)m and Bridgette. Sorry for the delayed response.
