

---

Subject: Dropping observations

Posted by [shayankhan](#) on Thu, 06 Oct 2016 23:00:36 GMT

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

---

I want to restrict my analysis to children under the age of 5 years using the Pakistan DHS of 2012-13. So, I am using the command of drop to just keep the observations of children that are under 5 years of age and am doing my analysis on them. Is this the valid approach?

---

---

Subject: Re: Dropping observations

Posted by [Bridgette-DHS](#) on Fri, 07 Oct 2016 11:38:47 GMT

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

---

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

Quote:Yes, that's valid. But if you are using the KR file, which contains most of the child-level variables, that file is already limited to children born in the past five years (both alive and dead), so you don't need to do anything.

---

---

Subject: Re: Dropping observations

Posted by [shayankhan](#) on Fri, 07 Oct 2016 11:45:33 GMT

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

---

Yeah, I guess I took the long route on that. But, now using the KR file. I will now just merge the variables that I need from HR and IR files to KR file and use that for my analysis. I'd appreciate if you could also let me know the consequences of not using svy commands in my analysis. I mean I know it is important to use it for representativeness but I would like to know that is there any situation where not using svy command is justified?

Thanks.

---

---

Subject: Re: Dropping observations

Posted by [Trevor-DHS](#) on Sun, 09 Oct 2016 23:52:23 GMT

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

---

You can not use svy if all you want is a point estimate. If you want confidence intervals or significance estimates then you need to use the svy commands. The point estimate will be the same whether you use the svy commands or not, but confidence intervals and tests of significance will be wrong if you do not use the svy commands

---

---

Subject: Re: Dropping observations

Posted by [shayankhan](#) on Mon, 10 Oct 2016 08:08:56 GMT

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

---

Hi Trevor,

What do you mean by point estimate? Are you talking about descriptive statistics? I did my logit model both by weights and unweighted and the significance differs as well as the coefficients.

---

---

Subject: Re: Dropping observations

Posted by [Bridgette-DHS](#) on Mon, 10 Oct 2016 11:29:03 GMT

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

---

Another response from Tom Pullum:

You can omit the svy correction for data quality checks, for confirming the construction of recodes, and for a preliminary or tentative analysis. However, your final results should always include the svy adjustments. The weight adjustment is most important. The cluster and stratum adjustments will only affect the standard errors and confidence intervals. They will not affect cross tabs or estimates of proportions, means, or rates. But if you want to construct confidence intervals or do tests, you really must use the cluster and stratum adjustments.

Before you try to merge IR and HR data onto the KR file, you should check which variables you need. Almost all of the relevant characteristics of the mother (from the IR file) are already included on the child's record. There may be some HR variables worth merging, but first check whether they are already in the KR file.

---

---

Subject: Re: Dropping observations

Posted by [shayankhan](#) on Mon, 10 Oct 2016 11:36:22 GMT

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

---

Thanks Bridgette. Yes, I got that much that I have to use weights with the survey command before any analysis. There is a recent paper by Woolrdidge by the title "What are we weighting for" in which the authors also mention exceptions for not using weights in regressions. The best way they say is to mention both the results, weighted as well as unweighted so that the reader can judge the difference in estimates.

What I don't understand though is that what is the difference between number of observations and population size. For example, when i run the svy command, i get number of observations and also the population size and both values are different. Am I to anlyze descriptives using observation size or population size?

---

---

Subject: Re: Dropping observations

Posted by [Bridgette-DHS](#) on Mon, 10 Oct 2016 12:19:44 GMT

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

---

Another response from Tom Pullum:

Quote:For every parameter there is a point estimate and an interval estimate. I would not describe the point estimate as a descriptive statistic but you can if you want. I have nothing to add to what I said before.

---

---

Subject: Re: Dropping observations

Posted by [shayankhan](#) on Mon, 10 Oct 2016 13:04:11 GMT

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

---

Thanks a lot Tom.

---