

---

Subject: Re: Pooled Cross sections

Posted by [id709nvz](#) on Wed, 07 Oct 2020 01:45:56 GMT

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

---

Dear,

I have issues which I believe is related to your response.

I am using combined Ethiopian KR file (2000, 2005, 2011, and 2016). I went through DHS forum to get some insight on how to weight the data.

To start with my research question:

I want to analyze the effect of X on Y where X started in 2011 but the survey were collected few months before, so in general 2011 is also a kind of pre treatment. My analysis is a sort of Diff-in Diff. could you please how to use weights accurately. to be precise I used the following command.

\*After opening KR file of 2000 survey

```
gen wgt=v005/1000000
```

```
gen weight=(wgt*FEMALE POPULATION)/FEMALE SAMPLE -For KR 2000; FEMALE  
POPULATION=14619 and FEMALE POPULATION=15367
```

```
gen survey=1
```

\*Same procedure for 2005, 2011 and 2016. Except that I used different total number of female population, female sample and with gen survey 2, 3, and 4 respectively. Then I appended the data starting from the latest survey (2016). Afterward, I did the following.

```
egen cluster=group(survey v021) *V021 is PSU and survey identifies year of survey (coded 1,  
2, 3 and 4)
```

```
egen stratum=group(survey v022) *v022 is sample strata for sampling error  
svyset cluster [pw=wgt], strata(stratum) singleunit(centered)
```

Is it correct to use: `tab x1 x2 [iweight=wgt]` or `tab x1 x2 [iweight=weight]`

how about `svy: reg Y x1 x2` (Will this understand the complex nature of the survey).

Added to this: `svy` is not supported with `ivreg2` stata command is there an alternative to this? I simply used `ivreg2` with cluster option at `v001`.

Thank you in advance for any comments.

Kind regards!

---