
Subject: Re: How to control for the year of sample
Posted by [Bridgette-DHS](#) on Fri, 08 Jan 2021 14:45:10 GMT
[View Forum Message](#) <> [Reply to Message](#)

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

There have been many posts on this topic (pooling or comparing surveys), but mostly in terms of Stata because mostly we use Stata.

The strata in successive surveys should have different codes for the purpose of adjusting for the survey design, even when the strata are the same in each survey. You need some mechanism for assigning different ID codes in the different surveys. For example, if you number the surveys 1, 2, 3, you could construct a variable "stratumid" as "V022 + 1000*(survey-1)". In Stata we would use "egen stratumid=group(survey v022)" to construct distinct numbers.

There are different possibilities for the weights. Say that the number of cases in each survey is n_1 , n_2 , n_3 . You could construct a new weight "v005rev" that would be proportional to v005 within each survey but would add to $(n_1+n_2+n_3)/3$ within each survey. This is spelled out in other posts. Alternatively, you could leave the weights alone, but then the estimates would be biased toward the largest survey.

You should still include region as a predictor in the regression. Including it (via V022) in the survey adjustments will only adjust for the design. It does not control for region in the analysis.