
Subject: Re: Using weights in regression analysis
Posted by [Bridgette-DHS](#) on Wed, 20 Feb 2013 16:50:09 GMT
[View Forum Message](#) <> [Reply to Message](#)

Here is a response from one of our DHS experts Tom Pullum, that should answer your question.

Future versions of the Guide to DHS Statistics will modify that recommendation. Not using weights is a minority viewpoint here at DHS. Almost all of us now advocate the use of weights. How you use them will depend somewhat on your statistical package. Most of us here use Stata.

If you do not use weights, the coefficients will be biased toward the over-sampled sub-populations.

For the HR and PR files, use hv005, for the IR, KR, and BR files, use v005, for the MR file use mv005. The CR file contains both v005 and mv005. It makes very little empirical difference which you use, but we prefer mv005 because it is adjusted for male non-response, which is typically more serious than female non-response. Some people (e.g. Stan Becker) have proposed a composite couples weight, but as I said the effect of alternatives is trivial.

If you use the AR file, the weight is hiv05, and if you form a couples file using the AR data, the weight is hiv05 for males.

For some purposes it is convenient to divide the weights by 1,000,000, but in Stata, for example, pweight is unaffected by that, and for regressions you use pweight.

You also need to adjust for the clusters (the primary sampling units) and the strata. In Stata that would be done with svyset and svy. These adjustments do not alter the coefficients but they do alter the standard errors, usually in opposite directions.

I hope this helps.

Bridgette-DHS
