
Subject: Weighting Combined Individual Data for Logistic Regression Analysis
Posted by [cudis](#) on Mon, 02 Feb 2015 03:32:20 GMT

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

We would like to conduct analysis of determinants of employment by estimating a logistic regression of the dichotomous employment outcome variable on various individual characteristics (e.g., education, age, age², sex, etc.).

From reading previous topics, our understanding is that DHS recommends weighting data before estimating regressions. However, although certain subpopulations are over- or under-represented in the sample, we cannot see how this would affect a regression onto individual characteristics. Could this please be explained in further detail? We also wonder whether the results of our regression analysis will be affected by the much different sample sizes for the two genders (but perhaps that is another issue entirely).

If we should weight the data, what would be the appropriate weight to use for a combined individual file (i.e., all men and women interviewed), where the unit of analysis is the individual?
