

---

Subject: WEIGHT SAMPLING REPRESENTATIVITY

Posted by [natiesim@yahoo.fr](mailto:natiesim@yahoo.fr) on Thu, 15 Oct 2015 19:52:47 GMT

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

---

Hello, Please I need your suggestions. I want to use data of DHS to analyse the effect of the consumption of a product on the probability to work. Unfortunately, I have only 100 people who consume the product. I thought I could use the sampling weight to allow my 100 people to be representative of the country population. However when I use `tab var [aw=hv005]` I have about 300 000 000 people who consume the product, that doesn't make sense since the real population of the country is about 11 000 000. when I use the recommendation of the DHS material `gen weight=hv005/1 000 000; tab var [aw=weight]` I get less than 100 people who consume the product. so I would like to know how I can use sample weights in my descriptive statistics to have a representativeness of the population with my few data. Thanks.

---

---

Subject: Re: WEIGHT SAMPLING REPRESENTATIVITY

Posted by [Reduced-For\(u\)m](#) on Thu, 15 Oct 2015 23:13:11 GMT

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

---

100 people who have chosen to consume a product can never represent the entire population (probably not even the entire population that consumes the product). That may or may not be a huge problem in your analysis - you don't need that sub-group to be population-representative. But any analysis on whether or not use of some product affects labor supply would need to include the observations of those that do NOT consume it anyway, so as to have a comparison rate of labor force participation. If you want to do this analysis knowing there are only 100, self-selected people who use the product, you can (though you are not likely to have sufficient precision to get a statistically significant result), but you should include everyone in the sample in your analysis, and compare those that did take it and those that did not.

When you use the whole sample, you should use the regular weights. Also, which types of weights are you specifying them as? If you are using them as "frequency weights", which I'm guessing might be the case, that would be the wrong type. The DHS weights are modified probability weights: in Stata for example you could specify them as "aweight" (analytic weights) or "pweight" (probability weights).

---