
Subject: Weighting data after merging survey rounds with different levels of representation

Posted by [jswindle](#) on Wed, 11 Jan 2017 06:11:07 GMT

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

Hello DHS experts and other forum users,

My problem relates to weighting after merging survey rounds.

I have merged the IR for the Malawi 2000, 2004, and 2010 surveys. When I weight this data, I am unsure how to best create the strata given different levels of representation across the surveys. The 2000 and 2004 surveys were representative across the 10 largest districts and across the three regions. The 2010 survey was representative across all 27 districts, as well as across the three regions.

I am combining these survey data with district level, time-varying data for foreign aid.

Currently I am doing the following, but am unsure about the commands. I am especially looking for guidance about the fifth through seventh lines below, which begin with "egen mw_00_strata..."

Thank you.

* Weight the dataset

```
generate weight = v005/10000000
```

```
recode survey (2000=1) (2004=2) (2010=3)
```

```
egen clusters=group(survey v021), label
```

```
egen mw_00_strata = group(survey region urban), label
```

```
egen mw_04_strata = group(survey region urban), label
```

```
egen mw_10_strata = group(survey region district urban), label
```

```
gen strata = .
```

```
replace strata = mw_00_strata if year==2000
```

```
replace strata = mw_04_strata if year==2004
```

```
replace strata = mw_10_strata if year==2010
```

```
svyset clusters [pweight=weight], strata(strata) singleunit(centered)
```
