

SSLC 2016-17

Stata syntax for weighted estimates and sampling error computation

This annex provides a set of examples of the STATA syntax for computing estimates and the corresponding sampling errors, which account for the full SSLC complex sample features (stratification, clustering and weighting due to differential inclusion probabilities and nonresponse adjustments). For more details, data users are referred to the online STATA manual for the svy command (<http://www.stata.com/manuals14/svy.pdf>).

To specify the BSLC sample design features in any of the dataset files, use the command:

```
svyset psu [pweight=weight], strata (stratum) fpc(psus_N)
```

To estimate the overall mean age of household members in dataset file RT002, use the command:

```
svy: mean q1_04
```

To estimate the overall mean age by gender, use the command:

```
svy: mean q1_04, over(q1_03)
```

To estimate the mean age of household heads, use the command:

```
svy, subpop (if q1_02==1): mean q1_04
```

To estimate the frequency distribution of marital status, use the command:

```
svy: tab q1_09, se ci
```

To estimate the frequency distribution of marital status by gender, use the command:

```
svy: tab q1_09 q1_03, col se ci
```

To estimate the frequency distribution of marital status of household heads, use the command:

```
svy, subpop (if q1_02==1): tab q1_09, se ci
```

To estimate the regression coefficients of a continuous variable y on two continuous variables x_1 and x_2 , use the command:

```
svy: regress y x1 x2
```

To estimate the regression coefficients of a continuous variable y on two continuous variables x_1 and x_2 and two categorical variables x_3 and x_4 , use the command:

```
xi: svy: regress y x1 x2 i.x3 i.x4
```

To estimate the Gini coefficient of variable y , install the `svylorenz` Stata command and use the syntax:

```
svylorenz y, ngp(5)
```