Stata tip 14: Using value labels in expressions

Kenneth Higbee
StataCorp
khigbee@stata.com

Did you know that there is a way in Stata to specify value labels directly in an expression, rather than through the underlying numeric value? You specify the label in double quotes (" "), followed by a colon (:), followed by the name of the value label. If we read in this dataset and see what it contains

```
. webuse census9
  (1980 Census data by state)
. describe
    obs: 50 1980 Census data by state
    vars: 5 16 Jul 2002 18:29
    size: 1,550 (99.9% of memory free)
```

we notice that variable `region` has values labeled by the `cenreg` value label. The correspondence between the underlying number and the value label is shown by

```
. label list
  cenreg:
     1 NE
     2 N Cntrl
     3 South
     4 West
```

[R] `regress` uses this dataset to illustrate weighted regression. To obtain the regression of `drate` and `medage` restricted to the “South” region, you could type

```
. regress drate medage [aweight=pop] if region == 3
```

But, if you do not remember the underlying region number for “South”, you could also obtain this regression by typing
Typing the value label instead of the underlying number makes it unlikely that you will obtain an unintended result from entering the wrong region number. An added benefit of using the value label is that, when you later review your results, you will quickly see that the regression is for the “South” region, and you will not need to remember what region was assigned number 3.

See [U] 16.9 Label values for further information about specifying value labels in expressions.