

ECON 4848

STATA HANDOUT 2

Commands.

1. **reg**: the command `reg` performs ordinary least squares estimation (or method of moments estimation) under classical assumptions. In addition, its output lists a series of statistics associated with a linear regression. Example: `reg regressand regressor1 regressor2` . Let's try this small code using the data set `wage1.dta`.

```
capture log close
set logtype text, permanently
log using first.txt, replace
clear
set memory 100m
set more on
use wage1.dta, clear
describe
reg wage educ
log close
exit
```

2. **predict**: after an estimation command, obtains predictions, residuals and a number of additional statistics based on the previously run estimation. General syntax is:

```
predict newvar, options
```

Example:

```
regress y x
predict yhat, xb
predict uhat, residuals
```

Some of the options are:

xb: linear prediction

residuals: linear regression residuals.

pr(a,b): the probability $P(y|a < y < b)$
e(a,b): the expectation $E(y|a < y < b)$
stdp: standard error of the linear prediction
stdf: standard error of the forecast
stdr: standard error of the residual

3. **scatter**: scatter draws scatterplots. It is both a command and a plottype in Stata. The general syntax is,

scatter variable list, options

Scatter may be combined with other plot types such as **line**, **lfit** as in,

scatter y x || **lfit** y x || **line** y x

Example:

regress y x

scatter y x || **lfit** y x

If more than two variables are specified, all but the last are given the interpretation of being y (vertical axis) variables. For example,

scatter y1var y2var xvar

would plot y1var versus xvar and overlay that with a plot of y2var versus xvar. It is the same as typing

scatter y1var xvar || **scatter** y2var xvar

For a list of all the options on **scatter** type **help scatter** in Stata.

It is easy to run regressions in Stata and conveniently present the results. Try using the following commands:

quietly regress y x

estimates store model1

quietly regress y1 x1

estimates store model2

estimates table model1 model2, **stats(r2 rmse) b se**