

{ STATA CHEAT SHEET }

MONIKA AVILA MARQUEZ, OSKA SHELDON-FENTEM
(UNIVERSITY OF BRISTOL)

BASICS:

1.1. Notes:

- In order to use the time series commands of STATA, it is crucial to declare the dataset as time series. More information on time series in Stata can be found in <https://www.stata.com/manuals/ts.pdf>.

1.2. Data uploading:

`use dataset`

`describe`

1.3. Declare dataset as time series:

`tsset timevariable`

Note: timevariable has to be an integer.

Hint:

- Generate new variable using `generate name_newvariable`
- Count number of observations in the dataset using `_n`

1.4. Basic commands of time series in Stata:

Get the lag of variable `y`: `L. y`

Get the 2-period lag of variable `y`: `L2. y`

Get the lead of variable `y`: `F. y`

Get the 2-period lead of variable `y`: `F2. y`

Get the first-difference of variable `y`: `D. y`

Get the difference of difference variable `y`: `D2. y`

AUTOCORRELATION

2.1. Regression:

- Include continuous variables in interactions with a continuous variable in a linear regression model: `c.name_variable`

- Predict residuals: `predict name_residuals, resid`

- Obtain R-squared and store it in a scalar object: `scalar R2 = e(r2)`

2.2. Testing for autocorrelation:

- Obtain Breusch-Godfrey test using the scalar command: use `scalar name_scalar = formula`

2.3. Other useful commands:

- Obtain the number of observations used to estimate the specified linear model: `e(N)`

- Obtain the number of observations used to estimate the specified linear model: `e(N)`

- Display created scalar: `display name_scalar`

HETEROSKEDASTICITY

3.1. Data simulation:

To simulate data from a normal distribution with mean 0 and variance 1: `gen x = rnormal(0, 1)`

3.2. OLS estimation with heteroskedastic-robust standard errors:

Regressing `y` on `x` with heteroskedasticity robust standard errors can be done using `regress y x, vce(robust)`

3.3. Heteroskedasticity diagnostics: graphical

We can perform an informal test of heteroskedasticity by plotting the fitted residuals, \hat{u} against our regressors x_i . By the definition of heteroskedasticity, we shouldn't observe any patterns between \hat{u} and our regressors. After performing a regression such as `regress y x` we can obtain the fitted residuals using `predict uhat, res`

3.3. Heteroskedasticity diagnostics: testing

We can formally test for heteroskedasticity by using a Breusch-Pagan test. In stata this can be done by using the command `estat hettest x_1 x_2` after fitting our regression model.