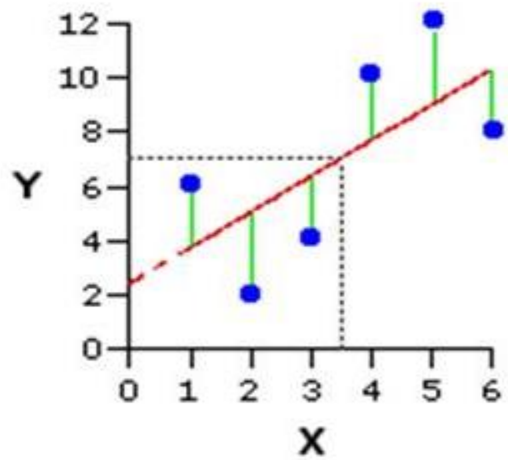
ORDINARY LEAST SQUARES (OLS) REGRESSION

1 Theoretical Considerations

Definition:

Method of estimating parameters, while minimizing the error observed between observed values and linear approximation of the data.

For example, consider a simple relationship between x and y.



A linear approximation of the data will generate error, e_i.

In the case above, OLS would estimate the following model:

$$y_i = \alpha + \beta x_i + \epsilon_i$$

Where the LS estimate is:

$$\hat{\beta} = \frac{\sum x_i y_i - \frac{1}{n}\sum x_i y_i}{\sum x_i^2 - \frac{1}{n}(\sum x_i)^2}$$

OLS seeks to minimize the sum of square of the error terms generated from estimating $\hat{\beta}$. That is, $\min \sum e_i^2$.

Assumptions:

OLS is subject to the following assumptions:

i) Model is correctly specified. E.g. linear relationship between x and y is true.
If violated: estimation is invalid.

ii) Strict exogeneity. Expected value (mean) of the error term is zero. $E[\epsilon] = 0$, furthermore, error terms and regressors (independent variable) should be uncorrelated, that is, $E[x, \epsilon] = 0$.
If violated,i.e. error terms and regressors are endogenous (related to each other)OLS estimates are invalid.

iii) No linear dependence between regressors. E.g. $E\left[x_i, x_j\right] = 0$
If violated, regressors are linearly dependent (multicollinear). Estimated coefficients, β_i are not accurate.

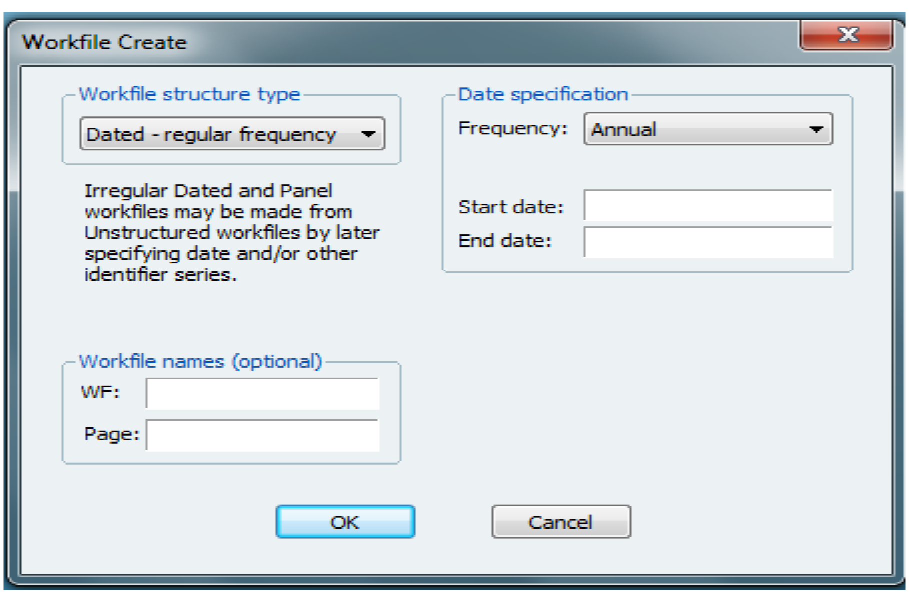
iv) Homoscedasticity. Constant variance of error terms, that is, $E\left[\epsilon_i^2 \middle| \boldsymbol{X}\right] = \boldsymbol{\sigma}$.
If violated, errors are heteroskedastic. OLS not suitable for the data.

v) No autocorrelation. Error terms between observations are not related to each other. This means that $E\left[\epsilon_i, \epsilon_j \middle| \boldsymbol{X}\right] = 0$.

2 Using Eviews

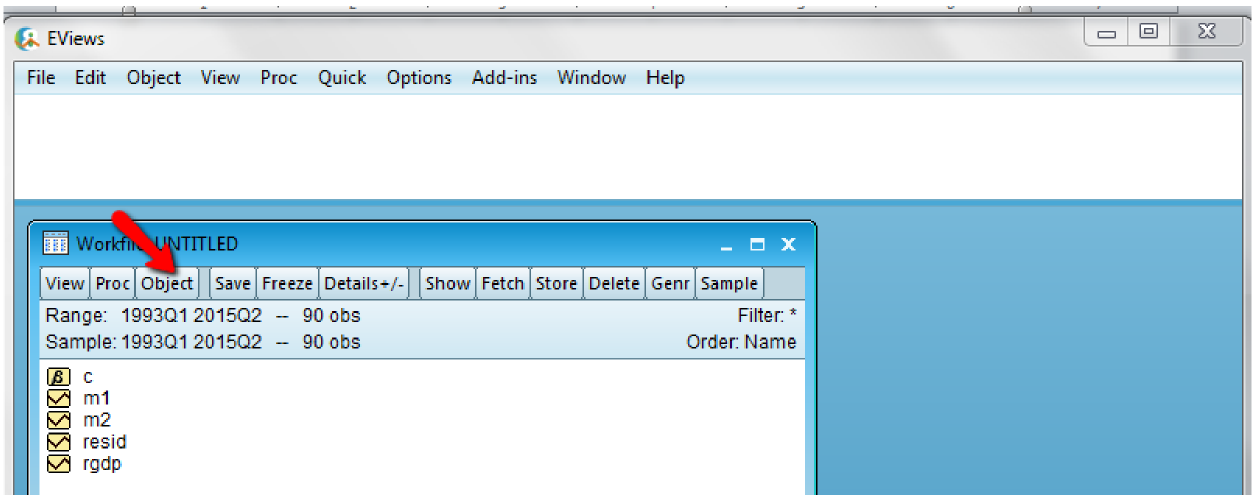
2.1 Create Workfile

- -
- -
- -

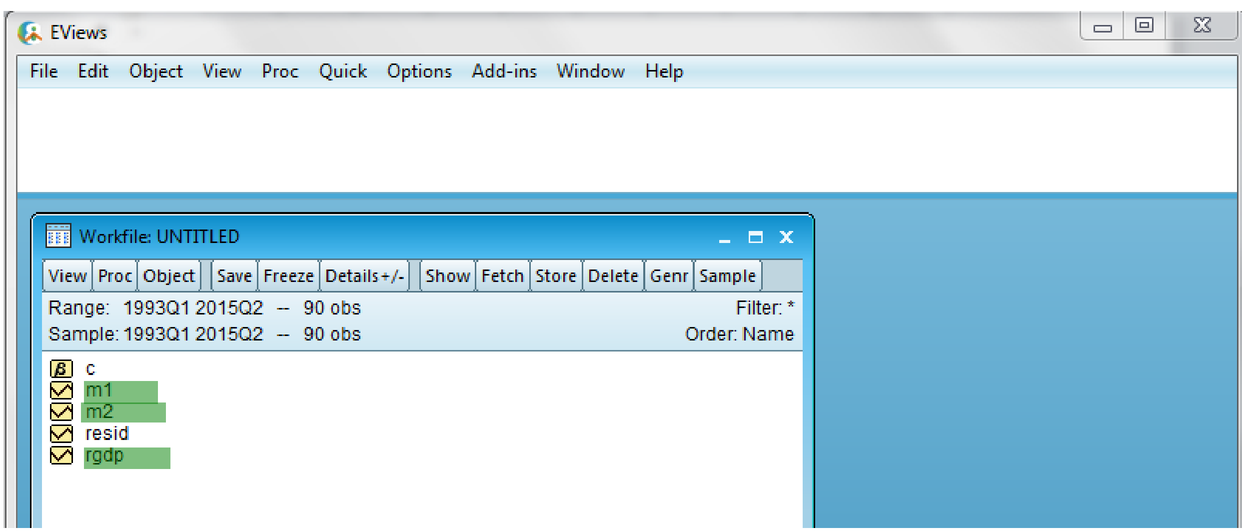


2.2 Import Data

- -
- -
- -

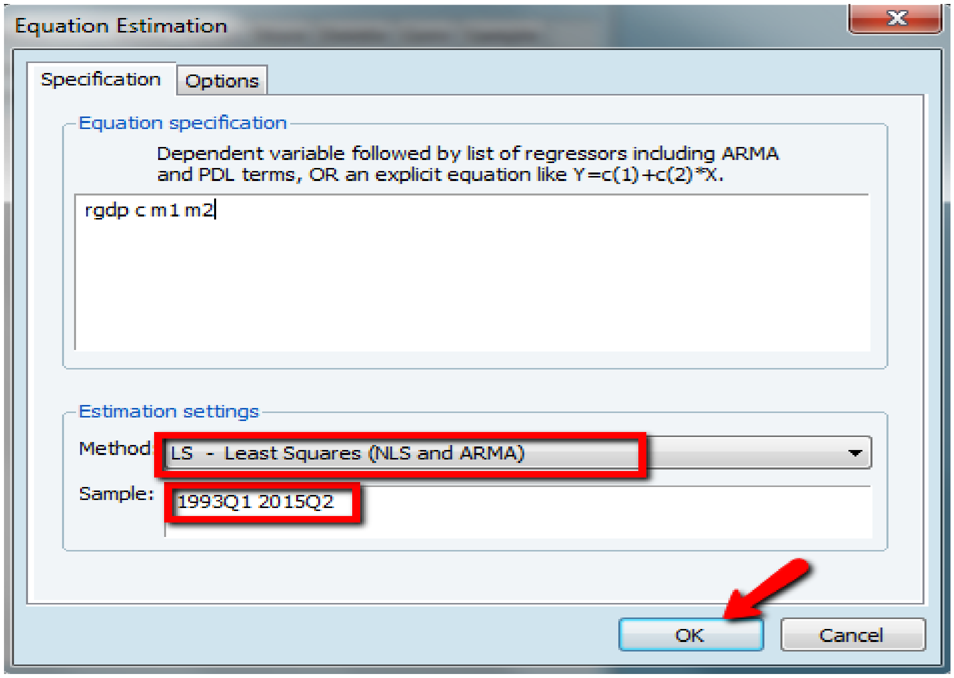


- -
- -
- -

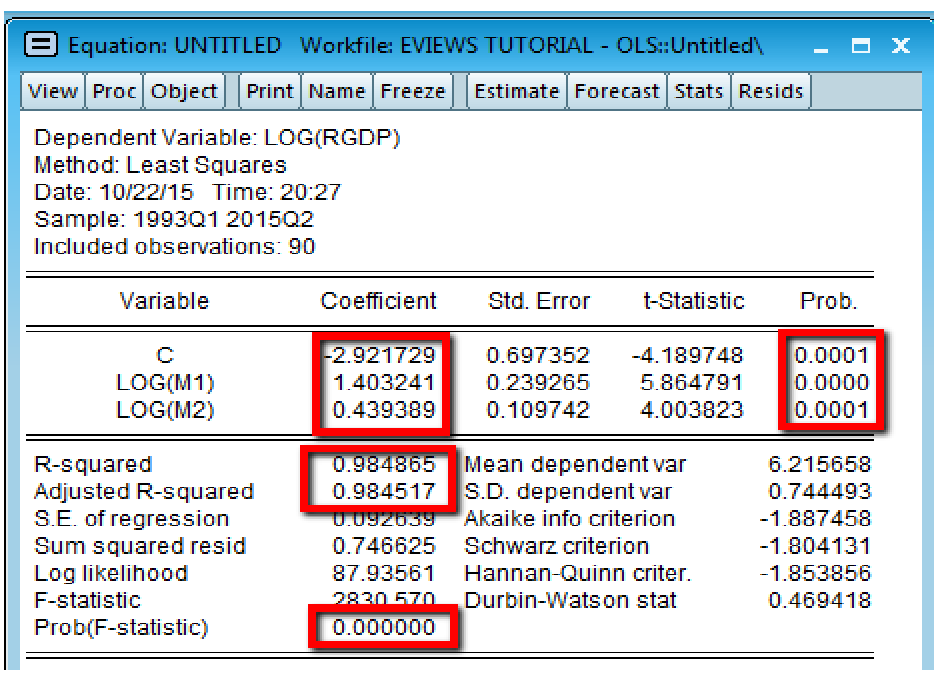


2.3 Run the OLS Regression

- -
- -
- -



Upon clicking “OK”, Eviews will run the data and produce the estimation output:



Dependent Variable: LOG(RGDP)
Method: Least Squares
Date: 10/22/15 Time: 20:27
Sample: 1993Q1 2015Q2
Included observations: 90

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.921729	0.697352	-4.189748	0.0001
LOG(M1)	1.403241	0.239265	5.864791	0.0000
LOG(M2)	0.439389	0.109742	4.003823	0.0001

R-squared	0.984865	Mean dependent var	6.215658
Adjusted R-squared	0.984517	S.D. dependent var	0.744493
S.E. of regression	0.092639	Akaike info criterion	-1.887458
Sum squared resid	0.746625	Schwarz criterion	-1.804131
Log likelihood	87.93561	Hannan-Quinn criter.	-1.853856
F-statistic	2830.570	Durbin-Watson stat	0.469418
Prob(F-statistic)	0.000000		

2.3 Run Diagnostic Tests

2.3.1 Multicollinearity (Coefficient Diagnostics)

To check for multicollinearity:

- -
- -
- -

Equation: UNTITLED Workfile: EVIEWS TUTORIAL - OLS::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Variance Inflation Factors
Date: 10/22/15 Time: 20:29
Sample: 1993Q1 2015Q2
Included observations: 90

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	0.486300	5099.920	NA
LOG(M1)	0.057248	15302.94	58.37098
LOG(M2)	0.012043	2851.784	58.37098

Rule of thumb: If VIF is greater than 5, than severe multicollinearity exists. In this case, multicollinearity exists between M1 and M2. Remedial measure?

2.3.2 Autocorrelation (Residual Diagnostics)

To check for autocorrelation:

- -
- -
- -

Equation: UNTITLED Workfile: EVIEWS TUTORIAL - OLS::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	67.18080	Prob. F(2,85)	0.0000
Obs*R-squared	55.12607	Prob. Chi-Square(2)	0.0000

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 10/22/15 Time: 20:35
Sample: 1993Q1 2015Q2
Included observations: 90
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.116695	0.462360	2.415208	0.0179
LOG(M1)	-0.381671	0.158476	-2.408385	0.0182
LOG(M2)	0.171298	0.072475	2.363534	0.0204
RESID(-1)	0.835041	0.105078	7.946880	0.0000
RESID(-2)	-0.035861	0.108267	-0.331228	0.7413

R-squared	0.612512	Mean dependent var	1.31E-15
Adjusted R-squared	0.594277	S.D. dependent var	0.091592
S.E. of regression	0.058341	Akaike info criterion	-2.791084
Sum squared resid	0.289308	Schwarz criterion	-2.652206
Log likelihood	130.5988	Hannan-Quinn criter.	-2.735080

If $\rho < \alpha$, where α is a specified significance level, reject null hypothesis of no serial correlation.

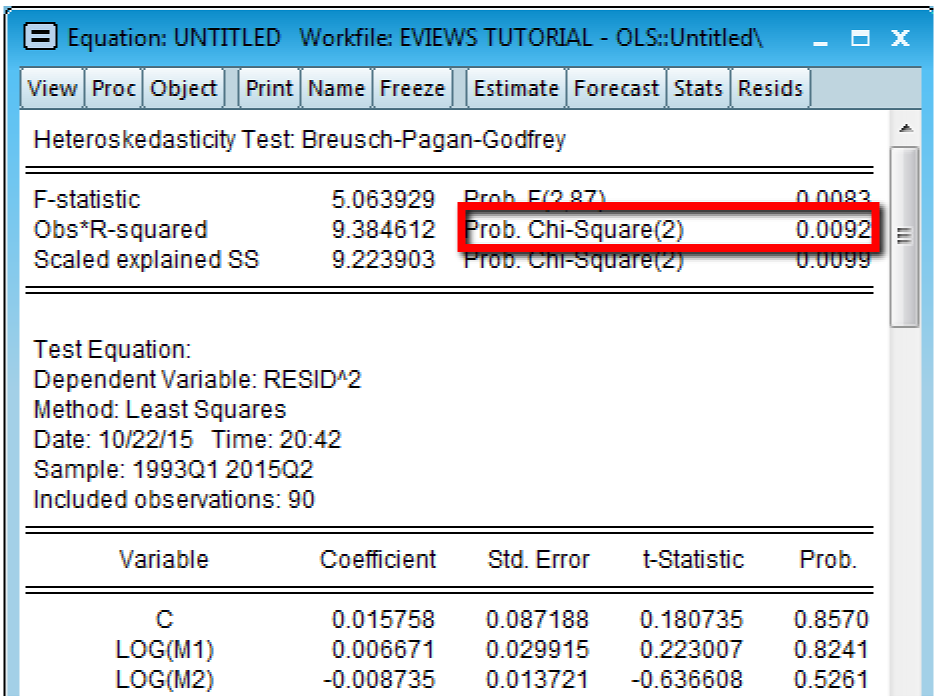
In the case above, ($\rho =$ _______) ($\alpha =$ _____), the model then,

____

2.3.3 Heteroscedasticity

To check for Heteroscedasticity:

- -
- -
- -



Heteroskedasticity Test: Breusch-Pagan-Godfrey

F-statistic	5.063929	Prob. F(2,87)	0.0083
Obs*R-squared	9.384612	Prob. Chi-Square(2)	0.0092
Scaled explained SS	9.223903	Prob. Chi-Square(2)	0.0099

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 10/22/15 Time: 20:42
Sample: 1993Q1 2015Q2
Included observations: 90

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.015758	0.087188	0.180735	0.8570
LOG(M1)	0.006671	0.029915	0.223007	0.8241
LOG(M2)	-0.008735	0.013721	-0.636608	0.5261

If $\rho < \alpha$, where α is a specified significance level, reject null hypothesis of homoscedasticity.

In the case above, ($\rho =$ _______) ($\alpha =$ _____), the model then,

____

3 Interpretation of Regression Results

OLS produced the following estimated regression equation:

$$\log(rgdp) = -2.92 + 1.4\log(M1) + 0.44\log(M2)$$

Our $\hat{\beta}_{\log(M1)}$is 1.4. This means that 1% increase/decrease in M1 money supply would increase/decrease real GDP by 1.4%. Using the same reasoning, 1% increase/decrease in M2 would increase/decrease real GDP by 0.44%.

Exercise:

Interpret the following regression results:

i) $\log y = 1.57 - 0.78\log x_1 + 1.5\log x_2$

ii) $\log y = 1.57 + 0.0097x_1 + 1.5\log x_2$

iii) $\log y = 1.57 + 1.2\log x_1 - 0.0097x_2$

iv) $y = 1.57 + 3.4\log x_1 - 580\log x_2$