ES1004 Econometrics by Example

Lecture 6: Autocorrelation

Dr. Hany Abdel-Latif

Swansea University, UK

Gujarati textbook, second edition

16th July 2016

CLRM Assumptions

A$_1$: model is linear in parameters

A$_2$: regressors are fixed non-stochastic

A$_3$: the expected value of the error term is zero $E(u_i|X) = 0$

A$_4$: homoscedastic or constant variance of errors $var(u_i|X) = \sigma^2$

A$_5$: no autocorrelation, $cov(u_i, u_j) = 0, i \neq j$

A$_6$: no multicollinearity; no perfect linear relationships among the Xs

A$_7$: no specification bias

Basic Idea I

- CLRM assumes the covariance between u_i and u_j is zero

$$E(u_i u_j) = 0 \quad for \quad i \neq j$$

  - the disturbance term relating to any observation is not influenced by the disturbance term relating to any other observation

No Autocorrelation

- in time series
  - the disruption due to a labour strike affecting output in one quarter will not be carried over to the next quarter
- in cross section
  - the effect of an increase of one family's income on its consumption expenditure is not expected to affect the consumption expenditure of another family

Autocorrelation I

$$E(u_i u_j) \neq 0 \quad for \quad i \neq j$$

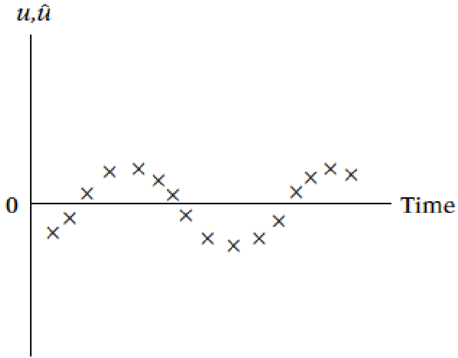
- the disruption caused by a strike this quarter may very well affect output next quarter
- the increases in the consumption expenditure of one family may very well prompt another family to increase its consumption expenditure

Autocorrelation II

- this is likely to be the case with time series data
  - the possible strong correlation between the shock in time t with the shock in time $t+1$

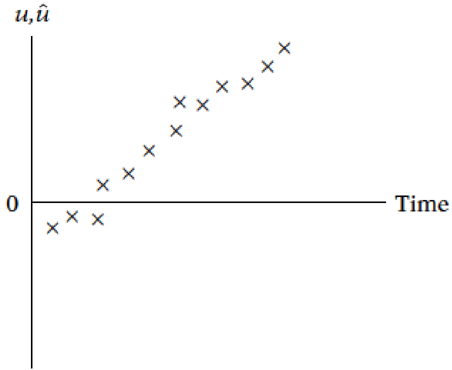
Autocorrelation: Example

- a discernible pattern among the u's [cyclical pattern]



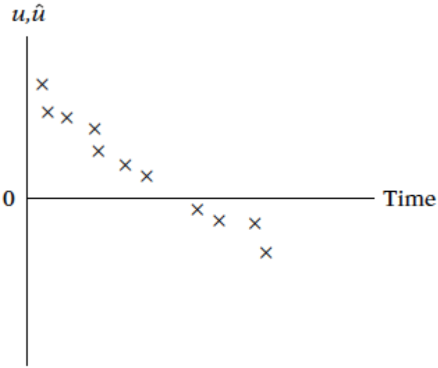
Autocorrelation: Example

- a discernible pattern among the u's [upward linear trend]



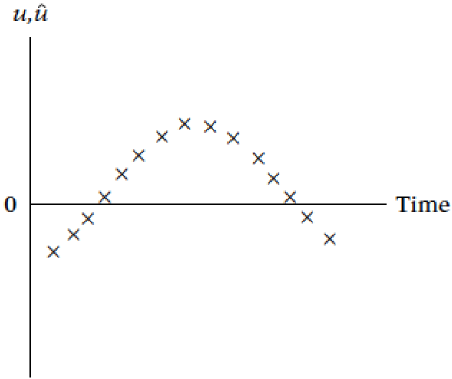
Autocorrelation: Example

- a discernible pattern among the u's [downward linear trend]



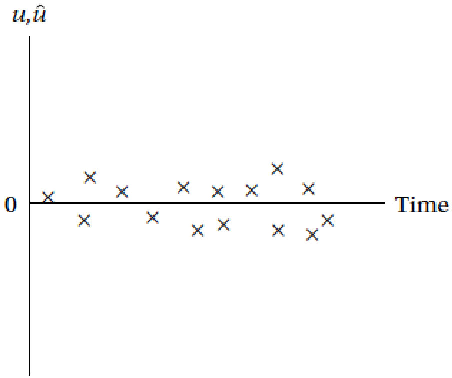
Autocorrelation: Example

- a discernible pattern among the u's [linear and quadratic trend]



No Autocorrelation: Example

- no discernible pattern among the u's [no systematic pattern]
  - no autocorrelation in this case



Inertia - Partial Adjustment

- most time series variables (e.g., gnp, price indexes, production, employment, and unemployment) exhibit business cycles
- there is a momentum built into them, and it continues until something happens (e.g., increase in interest rate or taxes or both) to slow them down

Misspecification - Specification Bias

- excluded variables: the omission of a relevant variable which is itself positively or negatively autocorrelated over time, and whose influence is then absorbed by u_i
- incorrect functional form

Cobweb Phenomenon and Lags

- the supply of agriculture commodities react to price with a lag of one time period because supply decisions take time to implement
- consumption level this year depends on income this year and consumption level last year

Data Manipulation: Smoothing

- raw data is often manipulated by taking average, which introduces smoothness into the data by dampening the fluctuation in the raw data
- interpolation and extrapolation of the data can be introducing autocorrelation

Nonstationarity

- a time series is stationary if its characteristics (mean, variance and covariance) are time invariant
- that is, they do not change over time

Autocorrelation and OLS Estimation

- if autocorrelation exists, several consequences ensue
  - OLS estimators still unbiased and consistent
  - still normally distributed in large samples
  - no longer efficient, meaning that they are not longer BLUE
  - in most cases standard errors are underestimated
  - hypothesis testing procedure becomes suspect, since the estimated standard errors may not be reliable, even asymptotically (i.e., in large samples)

Graphical Method

- plot the values of the residuals e_t chronologically
- if discernible pattern exists, autocorrelation likely a problem

Example: US Consumption Function

- `table6_1` time series data 1947-2000
- real consumption expenditure, real disposable personal income, real wealth, real interest rate
- the term real means adjusted for inflation

Example: OLS Estimation

Dependent Variable: LNCONSUMP
Method: Least Squares
Date: 07/16/16 Time: 15:16
Sample: 1 54
Included observations: 54

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.467712	0.042778	-10.93347	0.0000
LNDPI	0.804873	0.017498	45.99838	0.0000
LNWEALTH	0.201270	0.017593	11.44063	0.0000
INTEREST	-0.002689	0.000762	-3.529279	0.0009

R-squared	0.999560	Mean dependent var	7.826093
Adjusted R-squared	0.999533	S.D. dependent var	0.552368
S.E. of regression	0.011934	Akaike info criterion	-5.947705
Sum squared resid	0.007121	Schwarz criterion	-5.800373
Log likelihood	164.5880	Hannan-Quinn criter.	-5.890885
F-statistic	37832.66	Durbin-Watson stat	1.289232
Prob(F-statistic)	0.000000		

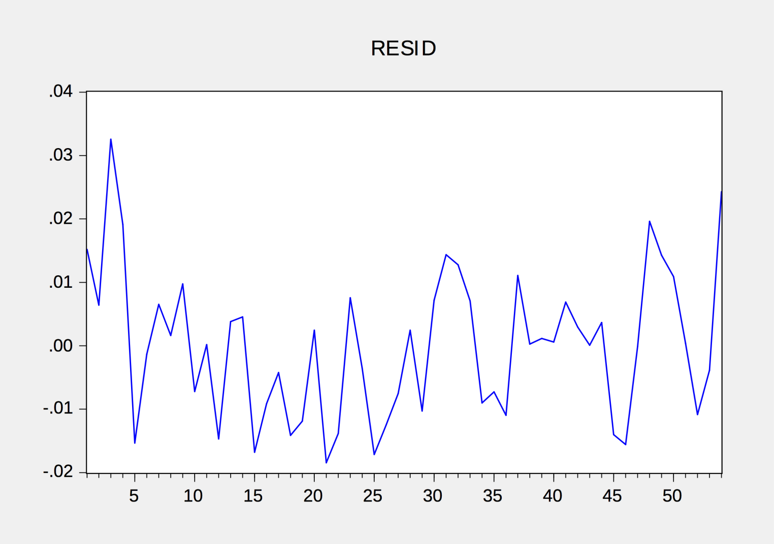
Example: OLS Estimation

Dependent Variable: LNCONSUMP
Method: Least Squares
Date: 07/16/16 Time: 15:16
Sample: 1 54
Included observations: 54

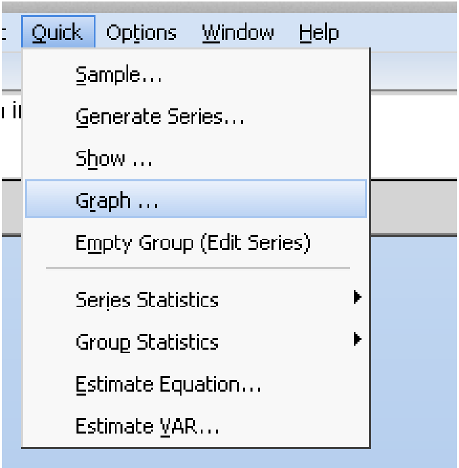
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.467712	0.042778	-10.93347	0.0000
LNDPI	0.804873	0.017498	45.99838	0.0000
LNWEALTH	0.201270	0.017593	11.44063	0.0000
INTEREST	-0.002689	0.000762	-3.529279	0.0009

R-squared	0.999560	Mean dependent var	7.826093
Adjusted R-squared	0.999533	S.D. dependent var	0.552368
S.E. of regression	0.011934	Akaike info criterion	-5.947705
Sum squared resid	0.007121	Schwarz criterion	-5.800373
Log likelihood	164.5880	Hannan-Quinn criter.	-5.890885
F-statistic	37832.66	Durbin-Watson stat	1.289232
Prob(F-statistic)	0.000000		

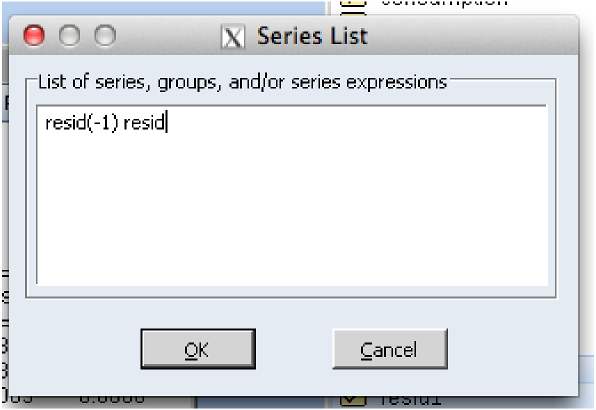
Example: Graphical Method I



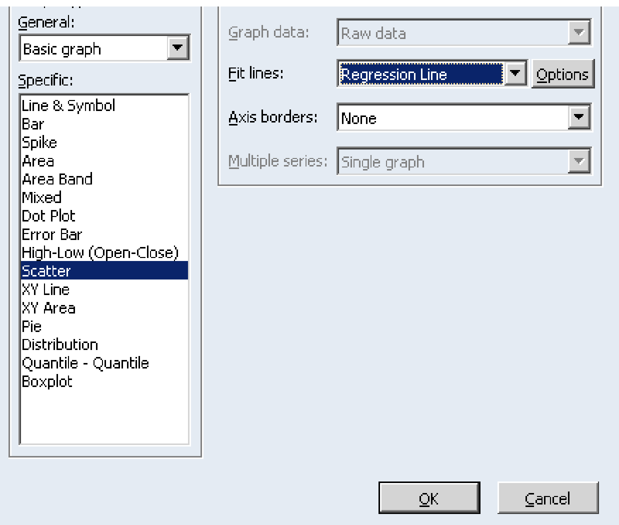
Example: Graphical Method II



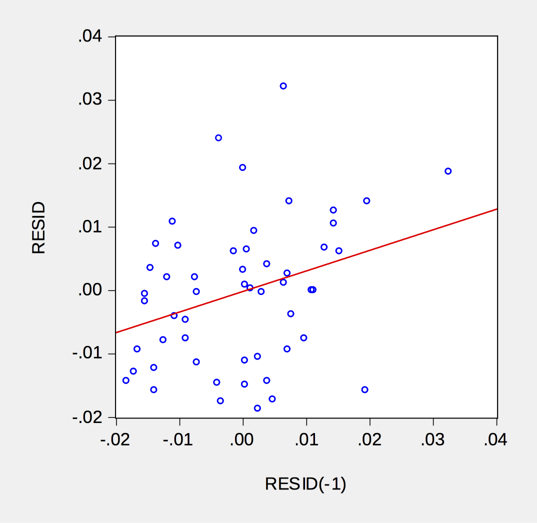
Example: Graphical Method II



Example: Graphical Method II



Example: Graphical Method II



Durbin Watson Test: Assumptions I

- the regression model includes an intercept term
- the regressors are fixed in repeated sampling
- the error term is normally distributed
- the regressors do not include the lagged value(s) of the dependent variable Y_t

Durbin Watson Test: Assumptions II

- the error term follows the first order autoregressive (AR1) scheme

$$u_t = \rho u_{t-1} + v_t$$

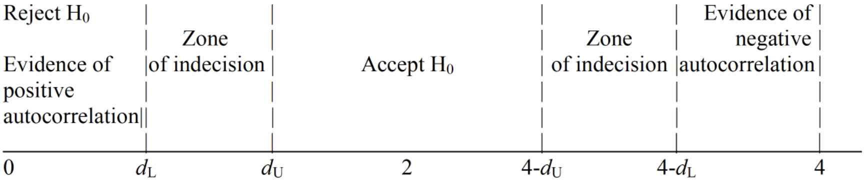
- where ρ (rho) is the coefficient of autocorrelation, a value between -1 and 1

Durbin Watson Test: Decision I

- two critical values of the d statistic, d_L and d_U
- d value always lies between 0 and 4
  - closer to 0 $\rightarrow$ positive autocorrelation
  - closer to 4 $\rightarrow$ negative autocorrelation
  - about 2 $\rightarrow$ no evidence of positive or negative (first order) autocorrelation

The Economic Society

Durbin Watson Test: Decision II



Example: OLS Estimation

Dependent Variable: LNCONSUMP
Method: Least Squares
Date: 07/16/16 Time: 15:16
Sample: 1 54
Included observations: 54

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.467712	0.042778	-10.93347	0.0000
LNDPI	0.804873	0.017498	45.99838	0.0000
LNWEALTH	0.201270	0.017593	11.44063	0.0000
INTEREST	-0.002689	0.000762	-3.529279	0.0009

R-squared	0.999560	Mean dependent var	7.826093
Adjusted R-squared	0.999533	S.D. dependent var	0.552368
S.E. of regression	0.011934	Akaike info criterion	-5.947705
Sum squared resid	0.007121	Schwarz criterion	-5.800373
Log likelihood	164.5880	Hannan-Quinn criter.	-5.890885
F-statistic	37832.66	Durbin-Watson stat	1.289232
Prob(F-statistic)	0.000000		

Durbin Watson Test: Decision II

	k' = 1		k' = 2		k' = 3	
n	d_L	d_u	d_L	d_u	d_L	d_u
34	1,39	1,51	1,33	1,58	1,27	1,65
35	1,40	1,52	1,34	1,58	1,28	1,65
36	1,41	1,52	1,35	1,59	1,29	1,65
37	1,42	1,53	1,36	1,59	1,31	1,66
38	1,43	1,54	1,37	1,59	1,32	1,66
39	1,43	1,54	1,38	1,60	1,33	1,66
40	1,44	1,54	1,39	1,60	1,34	1,66
45	1,48	1,57	1,43	1,62	1,38	1,67
50	1,50	1,59	1,46	1,63	1,42	1,67
55	1,53	1,60	1,49	1,64	1,45	1,68

The Economic Society

Breusch-Godfrey LM Test I

- this test allows for
  - lagged values of the dependent variables to be included as regressors
  - higher order autoregressive schemes, such as AR(2), AR(3), etc
  - moving average terms of the error term, such as u_{t-1}, u_{t-2}, etc

Breusch-Godfrey LM Test II

- the error term in the main equation follows the following AR(p) autoregressive structure

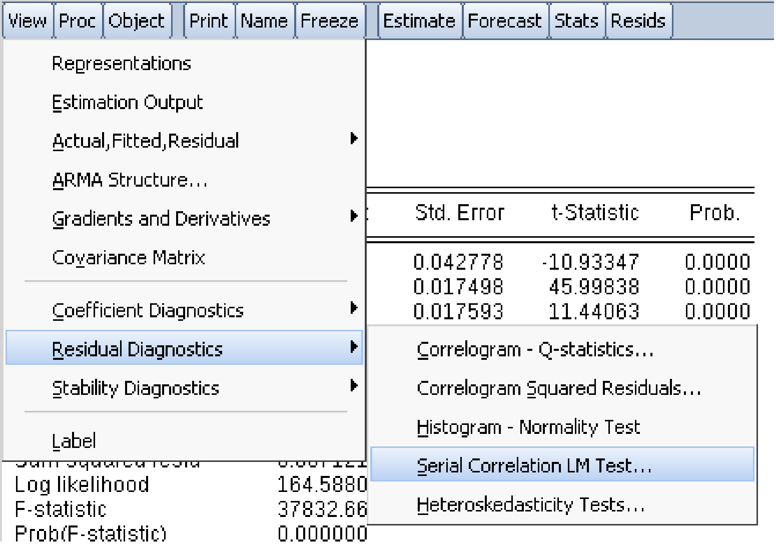
$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_p u_{t-p} + v_t$$

- the null hypothesis of no serial correlation is

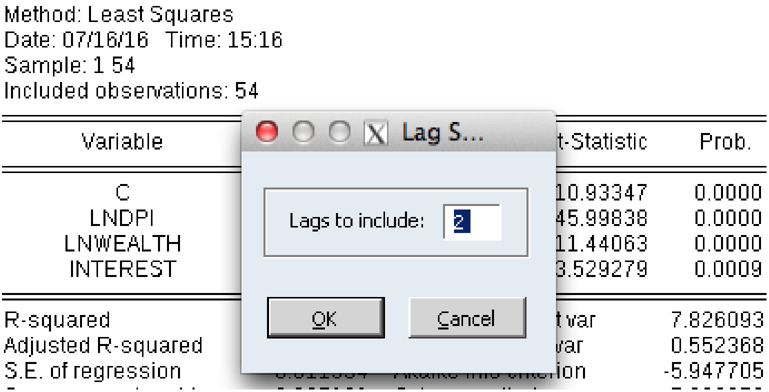
$$\rho_1 = \rho_2 = \cdots = \rho_p = 0$$

The Economic Society

Breusch-Godfrey LM Test: EViews



Breusch-Godfrey LM Test: EViews



Breusch-Godfrey LM Test: EViews

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	3.253931	Prob. F(2,48)	0.0473
Obs*R-squared	6.447226	Prob. Chi-Square(2)	0.0398

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 07/16/16 Time: 16:10
Sample: 1 54
Included observations: 54
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.006513	0.041529	-0.156839	0.8760
LNDPI	-0.004197	0.017158	-0.244607	0.8078
LNWEALTH	0.004191	0.017271	0.242661	0.8093
INTEREST	0.000116	0.000736	0.156970	0.8759
RESID(-1)	0.385178	0.151581	2.541070	0.0143
RESID(-2)	-0.165600	0.154695	-1.070492	0.2898

First Difference Transformation I

- if autocorrelation is of AR(1) type, we have

$$u_t - \rho u_{t-1} = v_t$$

- assume $\rho = 1$ and run the first-difference model
  - taking first difference of dependent variable and all regressors

First Difference Transformation II

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

Dependent Variable: D(LNCONSUMP)
Method: Least Squares
Date: 07/16/16 Time: 16:23
Sample (adjusted): 2 54
Included observations: 53 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.007046	0.003395	2.075001	0.0433
D(LNDPI)	0.714813	0.081689	8.750475	0.0000
D(LNWEALTH)	0.078267	0.038174	2.050292	0.0457
D(INTEREST)	0.000734	0.000801	0.916215	0.3640

R-squared	0.645332	Mean dependent var	0.035051
Adjusted R-squared	0.623617	S.D. dependent var	0.017576
S.E. of regression	0.010783	Akaike info criterion	-6.149264
Sum squared resid	0.005697	Schwarz criterion	-6.000563
Log likelihood	166.9555	Hannan-Quinn criter.	-6.092081
F-statistic	29.71909	Durbin-Watson stat	1.896780
Prob(F-statistic)	0.000000		

Other Methods

- generalised transformation
    - estimate value of ρ through regression of residual on lagged residual
    - use that value to run transformed regression
- Newey-West method
    - generates HAC standard errors
        - i.e., heteroscedasticity and autocorrelation consistent

The Economic Society