ES1004 Econometrics by Example

Lecture 6: Autocorrelation

Dr. Hany Abdel-Latif

Swansea University, UK

Gujarati textbook, second edition

16th July 2016



Dr. Hany Abdel-Latif (2016)

ES1004ebe Lecture 6

Autocorrelation 1 / 41

CLRM Assumptions

- A₁: model is linear in parameters
- A₂: 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
- A7: no specification bias



Image: A matrix and a matrix

Basic Idea I

• CLRM assumes the covariance between u_i and u_j is zero

$$E(u_iu_j) = 0$$
 for $i \neq j$

• 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

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

- 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



(日) (周) (日) (日)

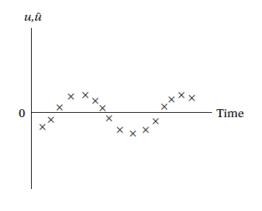
Autocorrelation II

- 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



A B A B
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A

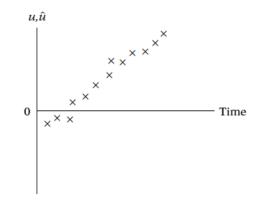
• a discernible pattern among the *u*'s [cyclical pattern]





Dr. Hany Abdel-Latif (2016)

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

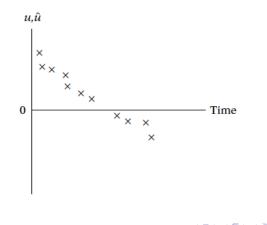


Dr. Hany Abdel-Latif (2016)

< 47 ▶

The Economic Soc

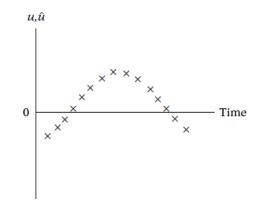
• a discernible pattern among the *u*'s [downward linear trend]



Dr. Hany Abdel-Latif (2016)

The Economic Soc

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



EC1004-1

Dr. Hany Abdel-Latif (2016)

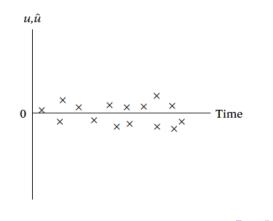
The Economic Soc

Nature of Autocorrelation

No Autocorrelation: Example

• no discernible pattern among the *u*'s [no systematic pattern]

- no autocorrelation in this case



Dr. Hany Abdel-Latif (2016)

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



э

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)



Detection

Graphical Method

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



э

ヨトィヨト

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 LNDPI LNWEALTH INTEREST	-0.467712 0.804873 0.201270 -0.002689	0.042778 0.017498 0.017593 0.000762	-10.93347 45.99838 11.44063 -3.529279	0.0000 0.0000 0.0000 0.0009
R-squared Adjusted R-squared S.E. of regression Sum squared resid Log likelihood F-statistic Prob(F-statistic)	0.999560 0.999533 0.011934 0.007121 164.5880 37832.66 0.000000	Mean dependent var S.D. dependent var Akaike info criterion Schwarz criterion Hannan-Quinn criter. Durbin-Watson stat		7.826093 0.552368 -5.947705 -5.800373 -5.890885 1.289232



э

・ロト ・聞ト ・ヨト ・ヨト

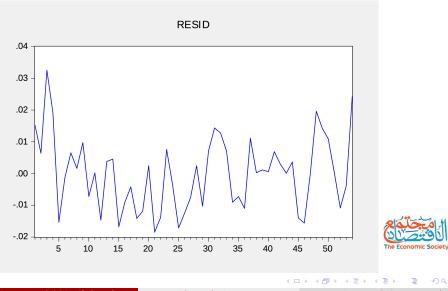
Example: OLS Estimation

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 LNDPI LNWEALTH INTEREST	-0.467712 0.804873 0.201270 -0.002689	0.042778 0.017498 0.017593 0.000762	-10.93347 45.99838 11.44063 -3.529279	0.0000 0.0000 0.0000 0.0009
R-squared Adjusted R-squared S.E. of regression Sum squared resid Log likelihood F-statistic Prob(F-statistic)	0.999560 0.999533 0.011934 0.007121 164.5880 37832.66 0.000000	Mean dependent var S.D. dependent var Akaike info criterion Schwarz criterion Hannan-Quinn criter. Durbin-Watson stat		7.826093 0.552368 -5.947705 -5.800373 -5.890885 1.289232

э

・ロト ・聞ト ・ヨト ・ヨト



Dr. Hany Abdel-Latif (2016)

ES1004ebe Lecture 6

Autocorrelation

:	Quick	Options	Window	Help	
	Sa	ample			
İI	Ge	enerate Se	ries		
	St	jow			
	Gŗ	aph			
-	En	npty Group) (Edit Serie	es)	
	Se	eries Statis	tics	I	•
	Gr	oup Statis	tics	I	•
	Es	timate Equ	uation		
	Es	timate <u>V</u> AF	ł		



3

・ロト ・ 日 ・ ・ ヨ ・ ・ ヨ ・

-		
	\varTheta 🔿 🔿 📉 Series List	
ł	List of series, groups, and/or series expressions	
ł	resid(-1) resid	
g		
3	<u>O</u> K <u>C</u> ancel	المجتمع
),	V. 185101	The Economic Society

Dr. Hany Abdel-Latif (2016)

Autocorrelation

イロト イヨト イヨト イヨト

24 / 41

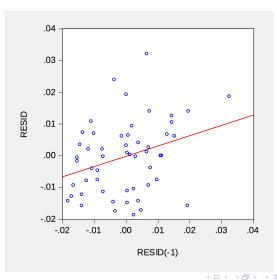
3

General:	Graph data:	Raw data	*	
Basic graph	Eit lines:	Regression Line	<u>Options</u>	
Line & Symbol Bar	<u>A</u> xis borders:	None	•	
Spike Area	Multiple series	Single graph	v	
Area Band Mixed Dot Plot Error Bar High-Low (Open-Close) Scatter XY Line XY Area Pie Distribution Quantile - Quantile				
Boxplot		QK	Cancel	The Economic Society
		< □	・ 〈母〉 〈同〉	∢ ≣ ▶ ≣ প ৭

Dr. Hany Abdel-Latif (2016)

ES1004ebe Lecture 6

Autocorrelation 2





Dr. Hany Abdel-Latif (2016)

ES1004ebe Lecture 6

Autocorrelation

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

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

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



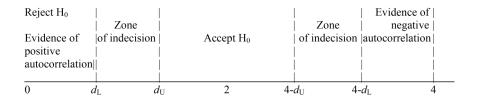
э

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
 - $\bullet~$ closer to 0 $\rightarrow~$ positive autocorrelation
 - $\bullet\,$ closer to 4 $\rightarrow\,$ negative autocorrelation
 - about 2 \rightarrow no evidence of positive or negative (first order) autocorrelation



Durbin Watson Test: Decision II





B> B

Image: A matrix and a matrix

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 LNDPI LNWEALTH INTEREST	-0.467712 0.804873 0.201270 -0.002689	0.042778 0.017498 0.017593 0.000762	-10.93347 45.99838 11.44063 -3.529279	0.0000 0.0000 0.0000 0.0009
R-squared Adjusted R-squared S.E. of regression Sum squared resid Log likelihood F-statistic Prob(F-statistic)	0.999560 0.999533 0.011934 0.007121 164.5880 37832.66 0.000000	Mean dependent var S.D. dependent var Akaike info criterion Schwarz criterion Hannan-Quinn criter. Durbin-Watson stat		7.826093 0.552368 -5.947705 -5.800373 -5.890885 1.289232

э

・ロト ・聞ト ・ヨト ・ヨト

Durbin Watson Test: Decision II

	k ' =	1	$\mathbf{k}^{*} = 2$	2	k ' = 3		[
n	\mathbf{d}_{L}	du	dL	du	\mathbf{d}_{L}	$\mathbf{d}_{\mathbf{u}}$	
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	2122011
50	1,50	1,59	1,46	1,63	1,42	1,67	الافتحلاق
55	1,53	1,60	1,49	1,64	1,45	1,68	The Economic Society

Dr. Hany Abdel-Latif (2016)

Autocorrelation

イロト イヨト イヨト イヨト

≣ •∕) ९ (२ 32 / 41

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

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

• the null hypothesis of no serial correlation is

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



A B A A B A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A
 A

Breusch-Godfrey LM Test: EViews

View Proc Object Print Nan	ne Freeze	Estimate	Forecast	Stats	Resids	
Representations						
Estimation Output						
Actual, Fitted, Residual	•					
ARMA Structure						
Gradients and Derivatives	•	Std. E	Error	t-Stati	istic	Prob.
Covariance Matrix		0.042	2778 -	10.93	347	0.0000
<u>⊂</u> oefficient Diagnostics	•	0.017		45.99 11.44		0.0000
Residual Diagnostics	•	⊆orre	elogram - C)-statis	tics	
Stability Diagnostics	•	Corre	logram Sq	uared I	Residual	s
Label		Histo	gram - Nor	rmality	Test	
Log likelihood	0.007 121 164.5880		Correlatio	on LM T	est	
F-statistic Prob(F-statistic)	37832.66	Heter	roskedasti	city Tes	sts	

The Economic Society

Ξ.

Dr. Hany Abdel-Latif (2016)

イロト イヨト イヨト イヨト

Breusch-Godfrey LM Test: EViews

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

Variable	⊖ ○ ○ 🗙 Lag S	t-Statistic	Prob.
C	Lags to include:	10.93347	0.0000
LNDPI		45.99838	0.0000
LNWEALTH		11.44063	0.0000
INTEREST		3.529279	0.0009
R-squared	QK Cancel	t var	7.826093
Adjusted R-squared		var	0.552368
S.E. of regression		nion	-5.947705



3

(日) (同) (三) (三)

Breusch-Godfrey LM Test: EViews

Breusch-Godfrey Serial Correlation LM Test:

Obs*R-squared 6.447226 Prob. Chi-Square(2) 0.0398	F-statistic Obs*R-squared		Prob. F(2,48) Prob. Chi-Square(2)	0.0473 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.	
		0.0. 200			
С	-0.006513	0.041529	-0.156839	0.8760	
C C	-0.000313	0.041529	-0.100003	0.0700	
LNDPI	-0.004197	0.017158	-0.244607	0.8078	1 the day of
LNWEALTH	0.004191	0.017271	0.242661	0.8093	Robert
INTEREST	0.000116	0.000736	0.156970	0.8759	<u>الاتحال</u>
RESID(-1)	0.385178	0.151581	2.541070	0.0143	The Economic Socie
RESID(-2)	-0.165600	0.154695	-1.070492	0.2898	

Dr. Hany Abdel-Latif (2016)

Autocorrelation

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

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 D(LNDPI) D(LNWEALTH) D(INTEREST)	0.007046 0.714813 0.078267 0.000734	0.003395 0.081689 0.038174 0.000801	2.075001 8.750475 2.050292 0.916215	0.0433 0.0000 0.0457 0.3640
R-squared Adjusted R-squared S.E. of regression Sum squared resid Log likelihood F-statistic Prob(F-statistic)	0.645332 0.623617 0.010783 0.005697 166.9555 29.71909 0.000000	Mean depend S.D. depende Akaike info cri Schwarz criter Hannan-Quin Durbin-Watso	nt var terion 'ion n criter.	0.035051 0.017576 -6.149264 -6.000563 -6.092081 1.896780

Dr. Hany Abdel-Latif (2016)

Autocorrelation

39 / 41

The Economic Society

Other Methods

- generalised transformation
 - $\bullet\,$ 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







Dr. Hany Abdel-Latif (2016)

ES1004ebe Lecture 6

Autocorrelation