

EE325 Ch.8 Autocorrelation

Read Gujarati Ch. 12



Outline

- 1 Nature of Autocorrelation
- 2 Consequence of Autocorrelation
- 3 Detection of Autocorrelation
- 4 Remedial Measures



Nature of Autocorrelation



Autocorrelation may be defined as correlation between members of series of observations ordered in time (as in time series data) or space (as in cross-sectional data)

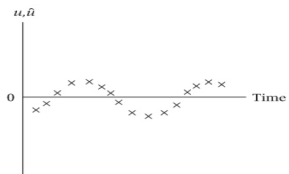


- Autocorrelation defines as lag correlation of a given series with itself, lagged by a number of time units
- The correlation between two time series such as u_1, u_2, \dots, u_{10} and u_2, u_3, \dots, u_{11} , where the former is the latter series lagged by one time period

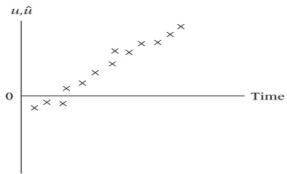


Serial correlation is the correlation between time series such as u_1, u_2, \dots, u_{10} and $\nu_2, \nu_3, \dots, \nu_{11}$ where u_i and ν_i are two different time series

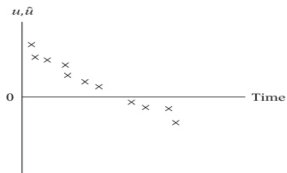




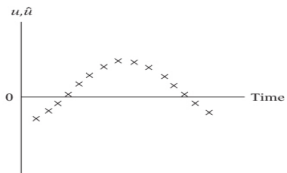
(a)



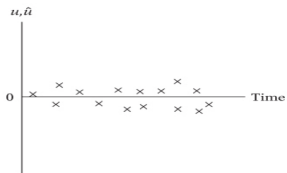
(b)



(c)



(d)



(e)

- Inertia

A salient feature of most economic time series is sluggishness

Example: GNP, price indexes, production, employment, and unemployment exhibit business cycles



- Specification error – excluded variables

True model:

Y_t = quantity of beef demand

X_{2t} = price of beef

X_{3t} = consumer income

X_{4t} = price of pork

Instead, we model:

where

The price of pork affects the consumption of beef, the error or disturbance term (ν) will reflect a systematic pattern, thus creating (false) autocorrelation

- Specification error – incorrect functional form

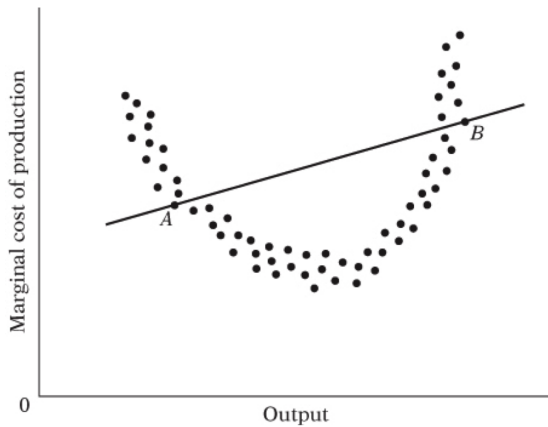
True model for cost-output study:

Instead, we model:

where

The marginal cost curve corresponding to the true model is shown in the next figure along with the incorrect linear cost curve





- Cobweb Phenomenon

The supply of many agricultural commodities reflects the so-called cobweb phenomenon, where supply reacts to price with a lag of one time period because supply decisions take time to implement



- Lags



- Data Transformation

$$\text{Level form} \quad Y_t = \beta_1 + \beta_2 X_t + u_t$$

$$Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + u_{t-1}$$

$$\text{Difference form} \quad \Delta Y_t = \beta_2 \Delta X_t + \Delta u_t$$

$$\Delta Y_t = \beta_2 \Delta X_t + \nu_t$$

$$\text{where} \quad \nu_t = \Delta u_t = (u_t - u_{t-1})$$

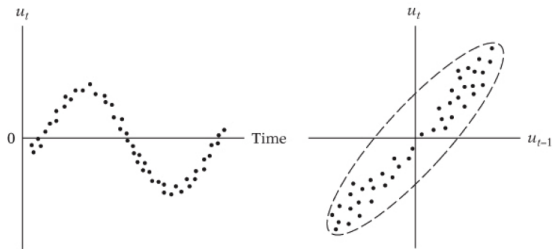
$$\text{Suppose} \quad E(u_t) = 0$$

$$\text{Then} \quad E(\nu_t) = E(u_t - u_{t-1}) = E(u_t) - E(u_{t-1})$$

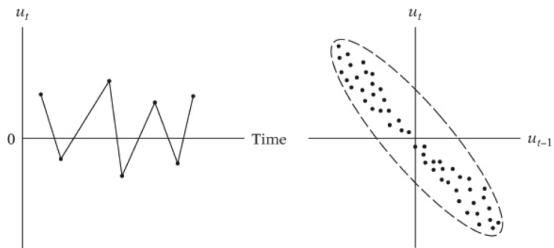
$$\begin{aligned} \text{var}(\nu_t) &= \text{var}(u_t - u_{t-1}) \\ &= \text{var}(u_t) - \text{var}(u_{t-1}) = 2\sigma^2 \end{aligned}$$

$$\begin{aligned} \text{cov}(\nu_t, \nu_{t-1}) &= E(\nu_t - \nu_{t-1}) \\ &= E[u_t - u_{t-1}][u_{t-1} - u_{t-2}] = -\sigma^2 \end{aligned}$$





(a)



(b)

Consequence of Autocorrelation



Characteristics of the Error Terms:

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

$$u_t = \rho u_{t-1} + \epsilon_t; \quad -1 < \rho < 1$$

where ρ is coefficient of autocovariance

Given

where ϵ is white-noise error term and $s \neq 0$



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

where $-1 < \rho < 1$

This equation is known as a Markov first order autoregressive scheme or first-order autoregressive scheme or AR(1)



$$\begin{aligned} E(u_t) &= \rho E(u_{t-1}) + E(\epsilon_t) = 0 \\ \text{var}(u_t) &= E(u_t^2) \\ &= \rho^2 \text{var}(u_{t-1}) + \text{var}(\epsilon_t) = \frac{\sigma_\epsilon^2}{1-\rho^2} \end{aligned}$$

Note

$$\begin{aligned} \text{var}(u_t) &= \text{var}(u_{t-1}) = \sigma^2 \\ \text{var}(\epsilon_t) &= \sigma_\epsilon^2 \end{aligned}$$

Covariance between error term s periods apart is

$$\text{cov}(u_t, u_{t+s}) = E(u_t, u_{t+s}) = \rho^s \frac{\sigma_\epsilon^2}{1-\rho^2}$$

The symmetry property of covariances follows

$$\text{cov}(u_t, u_{t+s}) = \text{cov}(u_t, u_{t-s})$$



By definition, the (population) coefficient of correlation between u_t and u_{t-1} is

$$\rho = \frac{E[u_t - E(u_t)]E[u_{t-1} - E(u_{t-1})]}{\sqrt{\text{var}(u_t)\text{var}(u_{t-1})}} = \frac{E(u_t u_{t-1})}{\text{var}(u_{t-1})}$$

Since $E(u_t) = 0$ for each t and $\text{var}(u_t) = \text{var}(u_{t-1})$ because we are retaining the assumption of homoscedasticity



Correlation between error terms s periods apart is

$$\text{cor}(u_t, u_{t+s}) = \rho^s$$

The symmetry property of correlations follows

$$\text{cor}(u_t, u_{t+s}) = \text{cor}(u_t, u_{t-s})$$

Since ρ is a constant between -1 and $+1$, variance equation shows that under the AR(1) scheme, the variance of u_t is still homoscedastic, but u_t is correlated not only with its immediate past value but its value several periods in the past



Normal OLS estimation:

$$Y_t = \beta_1 + \beta_2 X_t + u_t$$

OLS estimators are

and its variance is given by



Now under the AR(1) scheme, it can be shown that the variance of this estimator is

$$\text{var}(\hat{\beta}_2)_{AR(1)} = \frac{\sigma^2}{\sum x_t^2} \left[1 + 2\rho \frac{\sum x_t x_{t-1}}{\sum x_t^2} + 2\rho^2 \frac{\sum x_t x_{t-2}}{\sum x_t^2} + \dots + 2\rho^{n-1} \frac{\sum x_t x_n}{\sum x_t^2} \right]$$

where

$$\begin{aligned}x_t &= (X_t - \bar{X}) \\x_{t-1} &= (X_{t-1} - \bar{X}) \\&\vdots \\x_n &= (X_n - \bar{X})\end{aligned}$$



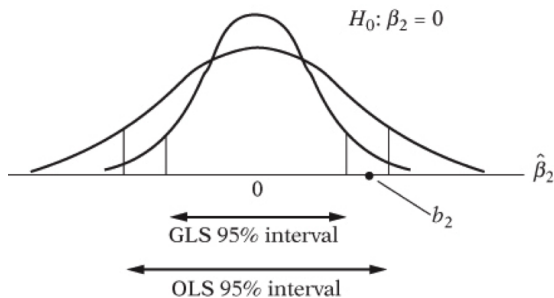
The difference between $var(\hat{\beta}_2)$ and $var(\hat{\beta}_2)_{AR(1)}$, assuming that the regressor X also follows the first order autoregressive scheme with a coefficient of autocorrelation of r , is

$$var(\hat{\beta}_2)_{AR(1)} = \frac{\sigma^2}{\sum x_t^2} \left(\frac{1+r\rho}{1-r\rho} \right) = var(\hat{\beta}_2)_{OLS} \left(\frac{1+r\rho}{1-r\rho} \right)$$

As in the case of heteroscedasticity in the presence of autocorrelation, the OLS estimators are still linear unbiased as well as consistent and asymptotically normally distributed, but they are no longer efficient (i.e. minimum variance)



$\hat{\beta}_2$ is not BLUE and even if we use $\text{var}(\hat{\beta}_2)_{AR(1)}$, the confidence intervals derived from it are likely to be wider than those based on the GLS procedure.



The situation is potentially very serious if we not only use $\hat{\beta}_2$ but also continue to use $var(\hat{\beta}_2)$, which completely disregards the problem of autocorrelation, that is, we mistakenly believe that the usual assumptions of the classical model hold true. Errors will arise for the following reasons:



- The residual variance $\hat{\sigma}_u^2 = \frac{\sum \hat{u}_i^2}{n-2}$ is likely to underestimate the true σ^2 (biased estimator)
- As a result, we are likely to overestimate R^2
- Even if σ^2 is not underestimated, $var(\hat{\beta}_2)$ may be underestimated compared to $var(\hat{\beta}_2)_{AR(1)}$ its variance under AR(1), even though the latter is inefficient compared to $var(\hat{\beta}_2)_{GLS}$
- Therefore, the usual t and F tests of significance are no longer valid



To see how OLS is likely to underestimate σ^2 and the variance of $\hat{\beta}_2$, let's conduct the following Monte Carlo Experiment

Suppose in the two-variable model

$$\begin{aligned} Y_t &= 1.0 + 0.8X_t + u_t \\ E(Y_t|X_t) &= 1.0 + 0.8X_t \end{aligned}$$

$$\begin{aligned} u_t &= 0.7u_{t-1} + \epsilon_t \\ \epsilon &\sim N(0, 1) \end{aligned}$$



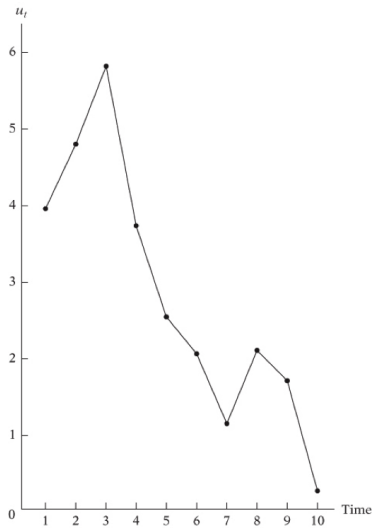
TABLE 12.1

**A Hypothetical
Example of Positively
Autocorrelated Error
Terms**

	ε_t	$u_t = 0.7u_{t-1} + \varepsilon_t$
0	0	$u_0 = 5$ (assumed)
1	0.464	$u_1 = 0.7(5) + 0.464 = 3.964$
2	2.026	$u_2 = 0.7(3.964) + 2.026 = 4.8008$
3	2.455	$u_3 = 0.7(4.8010) + 2.455 = 5.8157$
4	-0.323	$u_4 = 0.7(5.8157) - 0.323 = 3.7480$
5	-0.068	$u_5 = 0.7(3.7480) - 0.068 = 2.5556$
6	0.296	$u_6 = 0.7(2.5556) + 0.296 = 2.0849$
7	-0.288	$u_7 = 0.7(2.0849) - 0.288 = 1.1714$
8	1.298	$u_8 = 0.7(1.1714) + 1.298 = 2.1180$
9	0.241	$u_9 = 0.7(2.1180) + 0.241 = 1.7236$
10	-0.957	$u_{10} = 0.7(1.7236) - 0.957 = 0.2495$

Note: ε_t data obtained from *A Million Random Digits and One Hundred Thousand Deviates*, Rand Corporation, Santa Monica, Calif., 1950.





Now suppose the values of X are fixed at 1,2,3, . . . , 10 and generate estimated Y

TABLE 12.2
Generation of Y
Sample Values

X_t	u_t	$Y_t = 1.0 + 0.8X_t + u_t$
1	3.9640	$Y_1 = 1.0 + 0.8(1) + 3.9640 = 5.7640$
2	4.8010	$Y_2 = 1.0 + 0.8(2) + 4.8008 = 7.4008$
3	5.8157	$Y_3 = 1.0 + 0.8(3) + 5.8157 = 9.2157$
4	3.7480	$Y_4 = 1.0 + 0.8(4) + 3.7480 = 7.9480$
5	2.5556	$Y_5 = 1.0 + 0.8(5) + 2.5556 = 7.5556$
6	2.0849	$Y_6 = 1.0 + 0.8(6) + 2.0849 = 7.8849$
7	1.1714	$Y_7 = 1.0 + 0.8(7) + 1.1714 = 7.7714$
8	2.1180	$Y_8 = 1.0 + 0.8(8) + 2.1180 = 9.5180$
9	1.7236	$Y_9 = 1.0 + 0.8(9) + 1.7236 = 9.9236$
10	0.2495	$Y_{10} = 1.0 + 0.8(10) + 0.2495 = 9.2495$

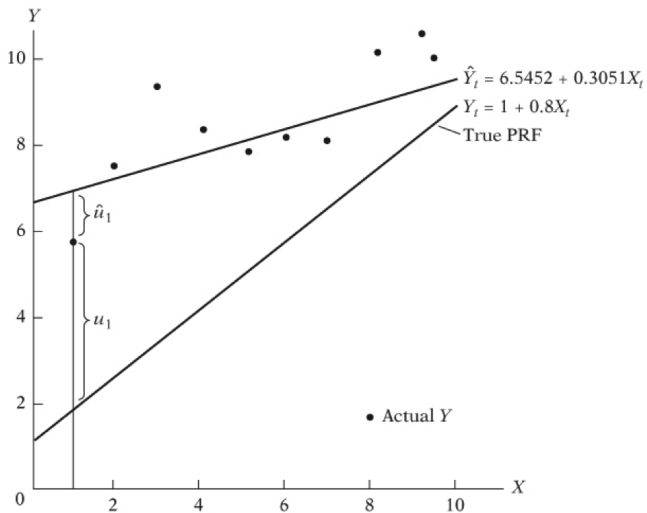
Note: u_t data obtained from Table 12.1.



Then we run the regression:

$$\begin{array}{rcl} \hat{Y}_t & = & 6.5452 + 0.3051X_t \\ se & & (0.6153) \quad (0.0992) \\ t & & (10.6366) \quad (3.0763) \\ R^2 & = & 0.5419 \\ \hat{\sigma}^2 & = & 0.814 \end{array}$$





Keeping the X_t and ϵ_t given in Table 12.1 and 12.2, let's assume $\rho = 0$, that is, no autocorrelation.

TABLE 12.3
Sample of Y Values
with Zero Serial
Correlation

X_t	$\epsilon_t = u_t$	$Y_t = 1.0 + 0.8X_t + \epsilon_t$
1	0.464	2.264
2	2.026	4.626
3	2.455	5.855
4	-0.323	3.877
5	-0.068	4.932
6	0.296	6.096
7	-0.288	6.312
8	1.298	8.698
9	0.241	8.441
10	-0.957	8.043

Note: Since there is no autocorrelation, the u_t and ϵ_t are identical. The ϵ_t are from Table 12.1.



The regression based on Table 12.3 is as following:

$$\begin{array}{rcl} \hat{Y}_t & = & 2.5345 + 0.6145X_t \\ se & & (0.6796) \quad (0.1087) \\ t & & (3.7910) \quad (5.6541) \\ R^2 & = & 0.7997 \\ \hat{\sigma}^2 & = & 0.9752 \end{array}$$



Detection of Autocorrelation



- Informal method
 - graphical method
- Formal method
 - t-test for autocorrelation
 - Durbin-Watson d test
 - Breusch-Godfrey test



t test for autocorrelation

we can examine AR(1) model or $u_t = \rho u_{t-1} + \epsilon_t$ and the independent variables in the model have to be **strictly exogenous**. That is,

$$E(u_t | X_{2t}, X_{3t}, \dots, X_{kt}) = 0$$

or

$$\text{cov}(u_t, X_{jt}) = 0 \text{ where } j = 2, 3, \dots, k$$



When the independent variables are strictly exogenous, we can set the following hypothesis as



With the following test procedure.

Step 1: estimate the model of interest through OLS method to obtain \hat{u}_t

Step 2: construct AR(1) model with dependent variable u_t and independent variable u_{t-1} and, then, estimate the model to obtain the estimated value of ρ .

Step 3: calculate t-statistic of the estimator $\hat{\rho}$ and test for statistical significance. If we can reject the null hypothesis, that means there is the first order autocorrelation among disturbance terms.



This approach can be applied to the test for higher order of autoregressive model like $u_t = \rho u_{t-3} + \epsilon_t$ which includes setting hypothesis, computing t-statistic, comparing it with the critical value in statistical table, and drawing the conclusion of whether the null hypothesis should be rejected.



Durbin-Watson d test

$$d = \frac{\sum_{t=2}^{t=n} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{t=n} \hat{u}_t^2}$$

The ratio of the sum of squared differences in successive residuals to the RSS. Note that in the numerator of the d statistics the number of observations is n-1 because one observation is lost in taking successive difference



- The regression model includes the intercept term
- The explanatory variables, X's, are non stochastic, or fixed in the repeated sampling
- The disturbances are generated by the AR(1):

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

Therefore, it cannot be used to detect higher-order AR schemes

- The error term, u_t , is assumed to be normally distributed
- The regression model does not include the lagged value(s) of the dependent variable as one of the explanatory variables
- There are no missing observations in the data



$$d = \frac{\sum_{t=2}^{t=n} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{t=n} \hat{u}_t^2}$$

$$d = \frac{\sum \hat{u}_t^2 + \sum \hat{u}_{t-1}^2 - 2 \sum \hat{u}_t \hat{u}_{t-1}}{\sum \hat{u}_t^2}$$

Since $\sum \hat{u}_t^2$ and $\sum \hat{u}_{t-1}^2$ differ in only one observation, they are approximately equal. Therefore, we set $\sum \hat{u}_t^2 \approx \sum \hat{u}_{t-1}^2$

$$d \approx 2 \left(1 - \frac{\sum \hat{u}_t \hat{u}_{t-1}}{\sum \hat{u}_t^2} \right)$$

$$\hat{\rho} = \frac{\sum \hat{u}_t \hat{u}_{t-1}}{\sum \hat{u}_t^2}$$

$$\therefore d \approx 2(1 - \hat{\rho})$$

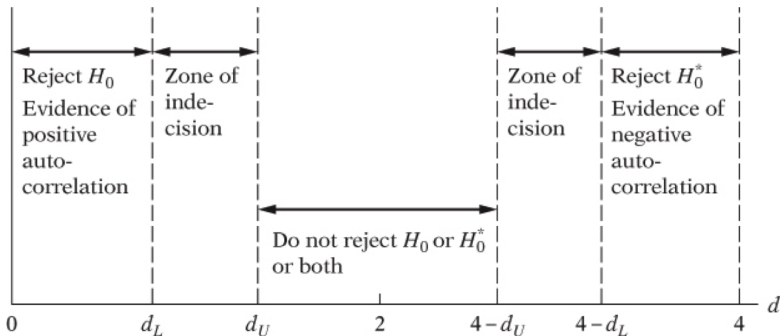
Since $-1 \leq \rho \leq 1$, thus $-4 \leq d \leq 4$



The procedure of the Durbin-Watson test are

- Run the OLS regression and obtain the residuals
- Compute d
- For the given sample size and given number of explanatory variables, find out the critical d_L and d_U values
- Follow the decision rules as following





Legend

 H_0 : No positive autocorrelation H_0^* : No negative autocorrelation

TABLE 12.6
Durbin-Watson d
Test: Decision Rules

Null Hypothesis	Decision	If
No positive autocorrelation	Reject	$0 < d < d_L$
No positive autocorrelation	No decision	$d_L \leq d \leq d_U$
No negative correlation	Reject	$4 - d_U < d < 4$
No negative correlation	No decision	$4 - d_U \leq d \leq 4 - d_L$
No autocorrelation, positive or negative	Do not reject	$d_U < d < 4 - d_U$



Example:

U.S. Consumption expenditure between 1947 - 2000

$$\ln Consumption_t = \beta_1 + \beta_2 \ln Income_t + \beta_3 \ln Wealth_t + \beta_4 \ln Interest_t + u_t$$



Source	SS	df	MS			
Model	16.1637474	3	5.3879158	Number of obs =	54	
Residual	.007120721	50	.000142414	F(3, 50) =	37832.66	
Total	16.1708681	53	.305110719	Prob > F =	0.0000	
				R-squared =	0.9996	
				Adj R-squared =	0.9995	
				Root MSE =	.01193	

Inc	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnincome	.8048728	.0174978	46.00	0.000	.7697273	.8400182
lnwealth	.2012702	.0175926	11.44	0.000	.1659345	.236606
i	-.0026891	.0007619	-3.53	0.001	-.0042194	-.0011587
_cons	-.467712	.042778	-10.93	0.000	-.5536342	-.3817899

$$\ln Consumption_t = -0.4677 + 0.8049 \ln Income_t + 0.2013 \ln Wealth_t - 0.0027 \ln Interest_t$$



Hypothesis Testing



```
. tsset year
      time variable: year, 1947 to 2000
      delta: 1 unit

. estat dwatson

Durbin-Watson d-statistic( 4, 54) = 1.289232
```

$H_o :$

$H_a :$

With $n = 54$, $k = 4$ at 0.05 level of significance;

$d_L =$

$d_U =$

Reject null hypothesis, evidence of positive correlation



Breusch-Godfrey test

This method can be used to test for any order, from 1 or AR(1) to p or AR(p) of autocorrelation as illustrated below from the simple regression model

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



We can form the hypothesis about the relationship among these disturbance terms as

$H_o :$

$H_a :$



With the following procedure.

Step 1: establish the regression model of interest. The example shown is the simple one.

Step 2: establish another model to obtain the relationship among disturbance terms with \hat{u}_t as the explained variable and X_{2t} and \hat{u}_{t-1} , \hat{u}_{t-2} until \hat{u}_{t-p} as the explanatory variables.

$$\hat{u}_t = \alpha_1 + \alpha_2 X_{2t} + \hat{\rho}_1 \hat{u}_{t-1} + \hat{\rho}_2 \hat{u}_{t-2} + \dots + \hat{\rho}_p \hat{u}_{t-p} + \epsilon_t \quad (1)$$



Step 3: if the sample size is large, LM-statistic can be computed by

$$LM = \quad (2)$$

Step 4: compare LM-statistic with the critical value of chi-square table to conclude whether the null hypothesis should be rejected. Specifically, we reject the null hypothesis if $(n - p)R^2$ is greater than the critical chi-square at the chosen level of significance and conclude that there exists the autocorrelation problem.



Remedial Measures



1. The method of generalized least squares (GLS)
2. The Newey West method - to obtain standard errors of PLS estimators that are corrected for autocorrelation



The Method of Generalized Least Squares (GLS)

The two-variable regression model:

$$Y_t = \beta_1 + \beta_2 X_t + u_t$$

and assume that the error term follows the AR(1) scheme; namely,

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

where $1 < \rho < 1$

Now we consider two cases:

- ρ is known
- ρ is not known but has to be estimated



When ρ is known

$$\begin{aligned} Y_{t-1} &= \beta_1 + \beta_2 X_{t-1} + u_{t-1} \\ \rho Y_{t-1} &= \rho\beta_1 + \rho\beta_2 X_{t-1} + \rho u_{t-1} \end{aligned}$$

$$(Y_t - \rho Y_{t-1}) = \beta_1(1 - \rho) + \beta_2(X_t - \rho X_{t-1}) + (u_t - \rho u_{t-1})$$

$$(Y_t - \rho Y_{t-1}) = \beta_1(1 - \rho) + \beta_2(X_t - \rho X_{t-1}) + \epsilon_t$$

where $\epsilon_t = u_t - \rho u_{t-1}$

$$Y_t^* = \beta_1^* + \beta_2^* X_t^* + \epsilon_t$$

Since the error term (ϵ_t) satisfies the usual OLS assumptions, we can apply OLS to the transformed variables Y^* and X^* and obtain estimators with all the optimum properties, namely, BLUE



When ρ is not known

- First difference method
- based on Durbin-Watson d statistics



$$\begin{aligned} Y_t - Y_{t-1} &= \beta_2(X_t - X_{t-1}) + (u_t - u_{t-1}) \\ \Delta Y &= \beta_2 \Delta X_t + \epsilon_t \end{aligned}$$

Appropriate when d is quite low or $d < R^2$



$$\hat{\rho} = 1 - \frac{d}{2}$$

In reasonably large sample, one can obtain ρ and use it to transform the data as shown in the generalized difference equation

$$(Y_t - \rho Y_{t-1}) = \beta_1(1 - \rho) + \beta_2(X_t - \rho X_{t-1}) + \epsilon_t$$

where $\epsilon_t = u_t - \rho u_{t-1}$



the Newey West method

The corrected standard errors are known as HAC (Heteroscedasticity and autocorrelation-consistent) standard errors or simply Newey West standard errors



```
. newey lnc lnincome lnwealth i, lag(3)
```

```
Regression with Newey-West standard errors
maximum lag: 3
```

```
Number of obs =      54
F( 3, 50) = 22341.20
Prob > F = 0.0000
```

lnc	Coef.	Newey-West Std. Err.	t	P> t	[95% Conf. Interval]	
lnincome	.8048728	.0171172	47.02	0.000	.7704919	.8392536
lnwealth	.2012702	.0154469	13.03	0.000	.1702441	.2322963
i	-.0026891	.0008798	-3.06	0.004	-.0044563	-.0009219
_cons	-.467712	.0439367	-10.65	0.000	-.5559616	-.3794625



Gujarati, D.N. (2009) Basic Econometrics. 5th ed. Singapore, McGraw-Hill.