[Academic Script]

Autocorrelation

Subject:	Business Economics
Course:	B. A. (Hons.), 5th Semester, Undergraduate
Paper No. & Title:	Paper – 531 Elective Paper Q1 – Advanced Econometrics
Unit No. & Title:	Unit – 1 Relaxing the Assumptions of The Classical Linear Model
Lecture No. & Title:	Lecture – 5 Autocorrelation

Academic Script

1. What is the problem of autocorrelation?

Hello friends nice meeting you. We want to discuss today the problem of autocorrelation. Generally we will consider the GLM then the disturbances are not correlated also they are not correlated with the explanatory variables. But sometimes it happens that the disturbances are correlated themselves by means of serial correlation. So such problem will be called problem of autocorrelation and we want to study in detail. First of all we should know what are the test of autocorrelation. Once we know this test then we should know how we can tackle the problem of autocorrelation and that is how we can deal with this situation whenever it occurs in practise. So now we will study in detail.

While dealing with GLM, one of the basic assumption was about homoscadasticity- that is uniform variance for the disturbances and pairwise the disturbances are uncorrelated.

If we violate this condition of uniform variance of disturbances and pair wise statistical independence among disturbances, we have the phenomena of heteroscadasticity.

In addition if we put the condition that any pair of disturbances is serially correlated then this phenomena is called autocorrelation. Before we go into further details, let us know about the nature of disturbances as shown by the following diagrams.

We may plot u (or $\hat{u}$ $i.e.e$) against time. We can have different situations as shown in the above diagrams.

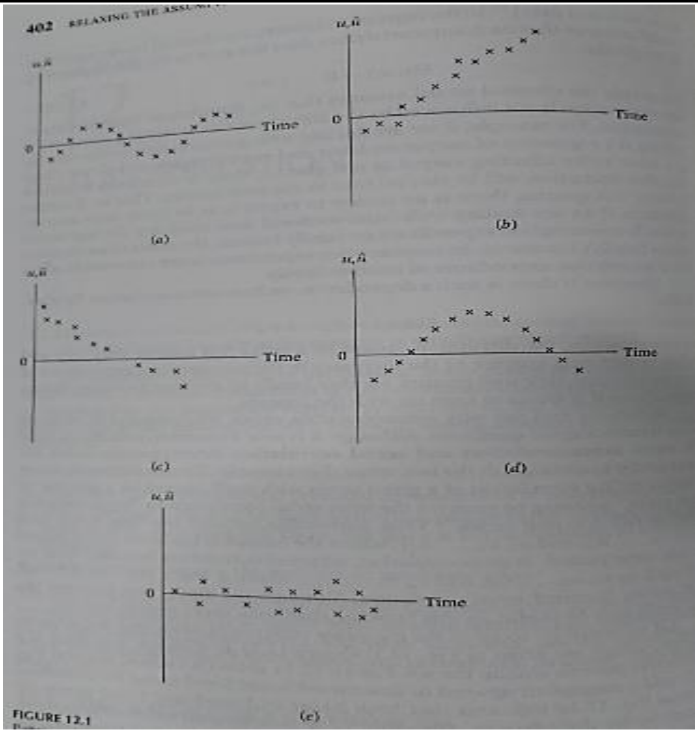


Fig (a) shows that there is a cyclic pattern

Fig (b) shows that there is an upward linear trend

Fig (c) shows that there is a downward linear trend

Fig (d) shows that there is both a linear and quadratic trend

Fig (e) shows that there is no systematic pattern.

Here fig(a) to fig(d) exhibits that there is some sort of autocorrelation between disturbances and fig(e) shows that there no autocorrelation among the disturbances. Briefly we may say that in fig(a) to fig(d), the disturbances are seen to be serially related.

What are the reasons for autocorrelation?

Some reasons for observing autocorrelation can be briefly summarized as under.

(1) Inertia

In some time series like GNP, production, employment etc. the series moves steadily Up or down. But in the longer span period the series remains almost stagnant or steady and then it starts ahead. This is called inertia and it causes autocorrelation.

(2) Manipulation of data

Due to interpolation and extrapolation based upon time series data, it is observed that the data are manipulated. Many times the quarterly data on time series are obtained by aggregating the data for 3 months to visualize quarterly variation. This is called smoothing of the time series data which cause autocorrelation.

(3) Specification bias

When one or more variables are excluded from the model we have the problem of specification. As per example, we have the model

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + U_t$$

Now if we exclude X_{4t} from the model and write

$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \nu_t$ then actually

$$\nu_t = \beta_4 X_{4t} + U_t$$

If we write disturbance term as U_t alone it is a specification bias which causes error due to autocorrelation.

(4) Incorrect specification of the model

Sometimes incorrect specification of the model also causes autocorrelation.

e.g. We have the model $Y_t = \beta_1 + \beta_2 X_t + \beta_3 X_t^2 + U_t$

If we write it as $Y_t = \beta_1 + \beta_2 X_t + \nu_t$

Then actually $\nu_t = \beta_3 X_t^2 + U_t$

So if we treat disturbance as U_t alone then it is not proper. This is also specification bias which leads to the presence of autocorrelation.

(5) Cobweb phenomena

Cobweb model represents the supply at time t in terms of the price at time$(t-1)$, but if the price at time t is more or less than the price at time (t-1) then accordingly supply will change and that in turn results in change in price and so on.

Due to such cobweb phenomena the effect will be to change the disturbances so that we find the presence of autocorrelation.

(6) Due to lag structure in the model

Consider a model as under

$$C_t = \beta_1 + \beta_2 I_t + \beta_3 C_{t-1} + U_t$$

Where C_t = Consumption at time t

I_t = Income at time t

C_{t-1} = Consumption at time t-1

Here there is a lag in the model (current consumption is affected by earlier consumption). Note that one of the explanatory variable is the dependent variable in the earlier period. If this fact is neglected then its resulting effect will be on the autocorrelation.

2. What are the practical consequences of autocorrelation?

We want to illustrate the case for the presence of autocorrelation by means of our two variables linear regression model.

We write this model in the standard form

$Y_t = \beta_1 + \beta_2 X_t + U_t \;\; (t = 1,2,\ldots,n)$.................................... (1)

Along with the usual assumptions for this model, we impose some more conditions for the disturbances in the model.

We assume that $U_t = \rho U_{t-1} + \epsilon_t$ .... (2)

$(t = 1,2,\ldots\ldots n)$

Which shows that the disturbances U_t follows first order autoregressive scheme

Here ρ is called coefficient of autocovariance or autocorrelation or autocorrelation coefficient $(-1 \leq \rho \leq 1)$ and ϵ_t is the stochastic disturbance term.

we further state that ϵ_t satisfies the following assumptions

$$\left.\begin{aligned} E\,(\epsilon_t) &= 0 \text{ for all t} \\ V\,(\epsilon_t) = E(\epsilon_t^2) &= \sigma^2 \text{ for all } t \\ COV\,(\epsilon_t, \epsilon_{t+s}) &= 0 \text{ for all } t\ (s \neq 0) \end{aligned}\right\} \quad \text{.................. (3)}$$

[Thus E $(\epsilon_t \cdot \epsilon_{t+s}) = \sigma^2$ if s $= 0$

$= 0$ if s $\neq 0$]

Note that stochastic disturbance term ϵ_t satisfies all the basic assumptions for standard model.

From equation (2)

$$U_t = \rho U_{t-1} + \epsilon_t$$

$$= \rho(\rho U_{t-2} + \epsilon_{t-1}) + \epsilon_t$$

$$= \rho^3 U_{t-3} + \rho^2 \epsilon_{t-2} + \rho\, \epsilon_{t-1} + \epsilon_t \text{ and so on.}$$

This shows that

$U_t = \epsilon_t + \rho\, \epsilon_{t-1} + \rho^2\, \epsilon_{t-2} + \rho^3\, \epsilon_{t-2} + \text{.........}$ upto ∞

Hence $U_t = \sum_{j=0}^{\infty} \rho^j\, \epsilon_{t-j}$

Thus $E(U_t) = \sum \rho^j E(\epsilon_{t-j}) = 0$ for all t

Now $U_t^2 = (\epsilon_t + \rho\, \epsilon_{t-1} + \rho^2\, \epsilon_{t-2} + \rho^3\, \epsilon_{t-3} + \cdots)^2$

Hence E $(U_t^2) = E(\epsilon_t^2) + \rho^2\, E(\epsilon_{t-1}^2) + \rho^4 E(\epsilon_{t-2}) + \ldots..$

Since $E(\epsilon_t^2) = \sigma^2$ for all t

We get E $(U_t^2) = \sigma^2(1 + \rho^2 + \rho^4 + \cdots \infty)$

Which is infinite geometric series

Hence V $(U_t) = \frac{\sigma^2}{1-\rho^2} = \sigma_u^2$

$$E(U_t \cdot U_{t-1}) = \rho\sigma^2$$

$$E(U_t \cdot U_{t-2}) = \rho^2\, \sigma^2$$

$$\ldots\ldots\ldots$$

$$E(U_t \cdot U_{t-s}) = \rho^s\, \sigma^2$$

We define $\rho_s =$ Serial Correlation coefficient of order s

Which is correlation coefficient between U_t and U_{t-s}

Then $\rho_s = \frac{cov(U_t, U_{t-s})}{\sqrt{V(U_t)\cdot V(U_{t-s})}}$

After simplification we get $\rho_s = \rho^s$

This is a relation for Serial Correlation Coefficient of order s in terms of the auto Correlation Coefficient (Note that for $s = 1$, $\rho_1 = \rho$ thus Series Correlation Coefficient of first order is the auto Correlation Coefficient. (Hence in practice we may use both the words conveniently).

Thus the dispersion matrix for the disturbances will be given by

V $(\underline{U}) = E\,(\underline{U}\,,\,\underline{U}') = V$ Where V is given by

$$V = \sigma_u^2 \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{pmatrix}$$

Here σ_u^2 is the variance of U which given by

$$\sigma_u^2 = \frac{\sigma^2}{(1-\rho^2)}$$

e.g. For 4 variables model,

$$V = \sigma_u^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{pmatrix}$$

Which is a real symmetric matrix with all the elements in the principle diagonals as unity

We may compare these situations by using OLS and GLS method.

Model $\underline{Y} = X\underline{\beta} + \underline{U}$

GLS	OLS
$E\left(\underline{U}\right) = \underline{0}$	$E\left(\underline{U}\right) = \underline{0}$
V $(\underline{U}) = \sigma_u^2 \cdot V$	V $(\underline{U}) = \sigma_u^2\, \mathrm{I}_n$

$\rho(X) = k < n$	$\rho(X) = k < n$
$\underline{\hat{\beta}} = (X'V^{-1}X)^{-1}X'V^{-1}\underline{Y}$	$\underline{\hat{\beta}} = (X'X)^{-1}X'\underline{Y}$
E $(\underline{\hat{\beta}_g}) = \underline{\beta}$	E $(\underline{\hat{\beta}}) = \underline{\beta}$
V $(\underline{\hat{\beta}_g}) = (X'V^{-1}X)^{-1}\sigma_u^2$	V $(\underline{\hat{\beta}}) = (X'X)^{-1}\sigma_u^2$

(Note that if $V = \mathrm{I}_n$, GLS and OLS are exactly identical.)

Now We consider the phenomena of auto Correlation where $u's$ are auto Correlated as shown above. So here if we use OLS method for estimation, then

$$\underline{\hat{\beta}} = (X'X)^{-1}X'\underline{Y}$$

$$=(X'X)^{-1}X'(X\underline{\beta} + \underline{U}) = \underline{\beta} + (X'X)^{-1}X'\underline{U}$$

Thus $\underline{\hat{\beta}} - \underline{\beta} = (X'X)^{-1}X'\underline{U}$

$$V(\underline{\hat{\beta}}) = E[\left(\underline{\hat{\beta}} - \underline{\beta}\right)\left(\underline{\hat{\beta}} - \underline{\beta}\right)']$$

$$= (X'X)^{-1}X'\{E(\underline{U}\,\underline{U}')\}X\,(X'X)^{-1}$$

V $(\underline{\hat{\beta}}) = [(X'X)^{-1}X'VX(X'X)^{-1}]$ where matrix V is as indicated above under auto correlation phenomena. (If V $= \sigma_u^2\,\mathrm{I}_n$, V $(\underline{\hat{\beta}})$ $= (X'X)^{-1}\sigma_u^2$ However if we use GLS estimator then V $(\underline{\hat{\beta}_g}) = (X'V^{-1}X)^{-1}\sigma_u^2$) Here it may be worthwhile as to which estimator can be considered to be preferable?

Note that due to matrix V, we may have the estimator with larger Variance but this all depends upon the autocorrelation coefficient. (Generally V $(\underline{\hat{\beta}})$ under autocorrelation will be larger than that without autocorrelation).

We want to understand this by means of two variable model.

We denote the model by $y_t = \beta x_t + u_t$, t = 1,2,…n

Where the original Variables are deviated from their means, thus we consider the reduced form, we have the relation

$u_t = \rho u_{t-1} + \epsilon_t$ for the case of autocorrelation.

In matrix notations $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} : n \times 1$

Hence $(X'X)^{-1} = \frac{1}{\sum_1^n x_i^2}$

$\underline{\hat{\beta}} = \hat{\beta}$

Thus using the relation V $(\underline{\hat{\beta}}) = [(X'X)^{-1}X'VX(X'X)^{-1}]$

We get $V(\hat{\beta}) = \left(\frac{1}{\sum x_i^2}\right)(x_1 \quad x_2 \quad \cdots \quad x_n)V\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\left(\frac{1}{\sum x_i^2}\right)$

Which gives after some simplification the result

$$V(\hat{\beta}) = \left(\frac{\sigma_u^2}{\sum x_i^2}\right)\left[1 + 2\rho\left(\frac{\sum_1^{n-1} x_i \cdot x_{i+1}}{\sum_1^n x_i^2}\right) + 2\rho^2\left(\frac{\sum_1^{n-2} x_i \cdot x_{i+2}}{\sum_1^n x_i^2}\right) + \cdots + 2\rho^{n-1}\left(\frac{x_1 x_n}{\sum_1^n x_i^2}\right)\right]$$

(Check that when $\rho = 0$, that is when there is no autocorrelation, we get $V(\hat{\beta}) = \left(\frac{\sigma_u^2}{\sum x_i^2}\right)$ as before)

Let us suppose that $\rho > 0$ and also that $x's$ are auto correlated with auto correlation coefficient ρ then we can write $\rho = \left(\frac{\sum_1^{n-1} x_i \cdot x_{i+1}}{\sum_1^n x_i^2}\right)$

$\rho^2 = \sum_1^{n-2} x_i \cdot x_{i+2} \Big/ \sum x_i^2$ and so on and finally $\rho^{n-1} = \frac{x_1 x_n}{\sum x_i^2}$ so that

after some simplication we get the expression in the case of sufficiently large n,

$$V(\hat{\beta}) \approx \left(\frac{\sigma_u^2}{\sum x_i^2}\right)\left(\frac{1+\rho^2}{1-\rho^2}\right) > \left(\frac{\sigma_u^2}{\sum_1^n x_i^2}\right)$$

Since $\rho > 0, \frac{1+\rho^2}{1-\rho^2} > 1$

Hence we may conclude that variance will be larger in the case of autocorrelation as compared to that in the case of no autocorrelation. Thus we conclude that if we ignore the phenomena of autocorrelation, variance of the estimator remains under estimated.

Estimation of $\underline{\hat{\sigma}^2}$ in the case of autocorrelation

We know that for 2 variables model under no autocorrelation,

$\hat{\sigma}^2 = \frac{\sum_1^n e_i^2}{(n-2)}$

In the case of autocorrelation, it can be shown that the estimation of σ^2 is given by the formula

$$E(\hat{\sigma}^2) = \sigma^2 \left[\frac{n-\left(\frac{2}{1-\rho}\right)-2\rho r}{(n-2)}\right]$$

Where $r = \frac{\sum_1^{n-1} x_i \cdot x_{+1}}{\sum_1^n x_i^2}$ and $\rho =$ autocorrelation coefficient

This also indicates that under autocorrelation the variance gets under estimated.

From the above expressions that we have studied, it is apparently useful and important to detect the presence of autocorrelation in the model when we deal actually with our data analysis. Our next two sections will consider this important aspects of detecting the problem of autocorrelation and hence after tackling the problem of autocorrelation.

3. How to detect autocorrelation?

(1) Graphical Method

As already explained earlier we can detect the autocorrelation situation by plotting the values of $U\left(or\ \widehat{U}\right)$ against time t. However this method just gives some preliminary knowledge only, but for more exactness we have to use the tests for autocorrelation.

(2) Certain tests for detecting autocorrelation

1. Von Neumann Ratio test

Here we want to test the hypothesis

$H_0: \rho = 0$ against $H: \rho \not> 0$

(Where $\rho =$ autocorrelation in the given series)

Von Neumann considered the ratio also known as VNR which is given by

$$\text{VNR} = \Delta = \frac{\delta^2}{S^2} = \frac{\sum_{i=2}^{n}(e_i - e_{i-1})^2 / (n-1)}{\sum_{i=1}^{n}(e_i - \bar{e})^2 / n} \quad \text{where } \bar{e} = \frac{\sum_{i=1}^{n} e_i}{n}$$

$e_i = i^{th}$ error term

It can be shown that for large samples

$E(\Delta) = \frac{2n}{n-1}$ = Mean Value of Δ

$V(\Delta) = \frac{4n^2(n-2)}{(n+1)(n-1)^2}$ = variance of Δ

Then we can apply normality test given by

$|z| = \frac{|\Delta - E(\Delta)|}{\sqrt{V(\Delta)}}$

Hence if |z| is significant, reject H_0 (which means that there is autocorrelation).

This can be tested as usual at 5% as well as 1% level of significance.

(2) Durbin Watson Test (DW test)

(I) Assumptions

(1) Regression model assumes the intercept term. If the deviated form is used, compute the error term based upon the regression with intercept term.

(2)The explanatory variables are non stochastic or fixed in the repeated sampling

(3) The disturbances U_t are generated by the first order autoregressive scheme given by the equation $u_t = \rho u_{t-1} + \epsilon_t$

(4) Regression model does not include Lagged Value (s) of the dependent variable as one of the explanatory variable

(II) Test Statistic

Define statistic d given by the formula

$d = \frac{\sum_{i=2}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} e_i^2}$

Here $e_i = i^{th}$ error term in the sample of n observations.

The exact distribution of d is difficult due to expression on the right hand side because the errors depend upon the explanatory variables.

Durbin and Watson (DW) have obtained a Lower bound d_L and an upper bound d_U to apply this test for detecting serial correlation in the given series .

(e.g. $H_0: \rho^s = 0 \Rightarrow \rho = 0$ for all S)

Writing d = $\left(\sum_{i=1}^{n} e_i^2 + \sum_{i=2}^{n} e_{i-1}^2 - 2\sum_{i=2}^{n} e_i \cdot e_{i-1}\right) \Big/ \sum_{i=1}^{n} e_i^2$

Since only one observation is missing in the numerator we can set $\sum_{i=1}^{n} e_i^2 \doteqdot \sum_{i=2}^{n} e_{i-1}^2$

This shows that $d \approx 2\left(1 - \sum_{i=2}^{n} e_i\, e_{i-1} \Big/ \sum_{i=1}^{n} e_i^2\right)$

Let us denote Sample first order autocorrelation coefficient by

$\rho^* = \sum_{i=2}^{n} e_i\, e_{i-1} \Big/ \sum_{i=1}^{n} e_i^2$ then $d \approx 2(1 - \rho^*)$

(i) If $\rho^* =$ 0, $d \approx 2$

(ii) If $\rho^* = 1, d \approx 0$

(iii) If $\rho^* = -1, d \approx 4$ Hence d lies between 0 to 4

Working Rule

(i) If d is very near to 2, there is no positive auto correlation

(ii) If d is very close to 0, there is positive auto correlation (perfect if $d = 0$)

(iii) If d is very near to 4, there is negative autocorrelation (perfect if $d = 4$)

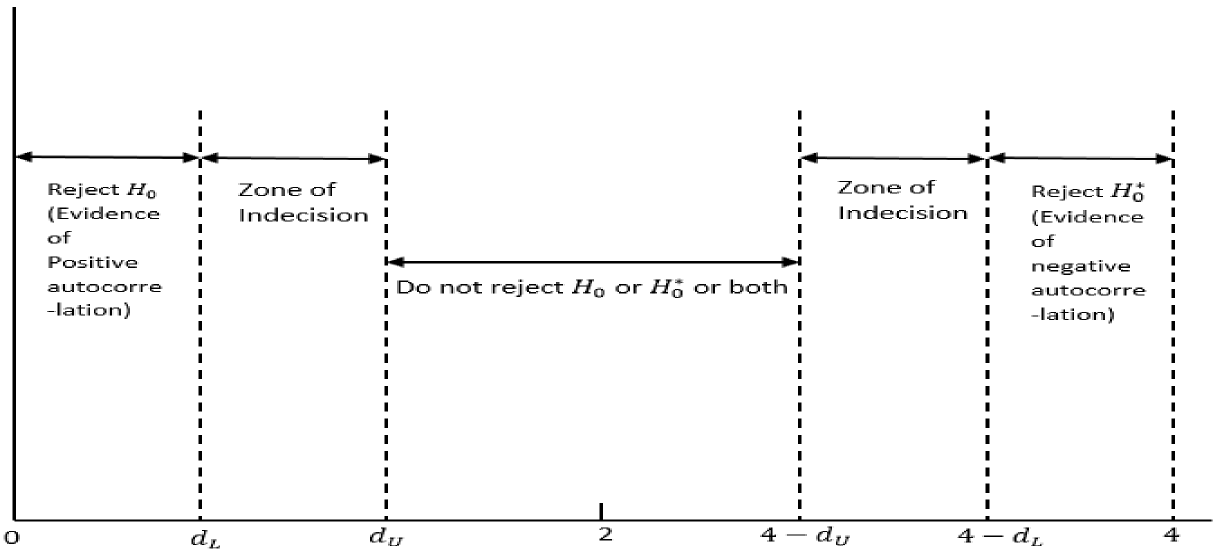
Computational procedure

(1) Run the OLS regression as usual and obtain the residuals

(2) Compute d statistic by the formula as given above

(3) For the given sample size n and the given number of explanatory variables k, obtain values of d_L and d_U from the tables prepared by DW. (DW test tables).

Different tests can be considered as shown below, Diagrammatic presentation helps in locating the test procedure.



H_0: *No Positive Correlation*

H_0^* : *No Negative Correlation*

DW test (Decision rules)

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

Comments

(1) This test is very popular for all applications.

(2) To overcome the problem when decision cannot be taken due to indecisive region, modified DW test can be used.

(3) There are other tests also for detecting auto correlation. They are named as Run test, Durbin's h statistic test, Breusch-Godfrey (BG) test of higher order auto correlation, Durbin's m test etc.

(4) For DW test, the Sample size should be as least 15

4. What are the remedial measures to tackle the problem of auto correlation?

Case: 1 When ρ is known

(1) Method of Generalised Difference Equation

Consider the two variables model

$$Y_t = \beta_1 + \beta_2 X_t + U_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

and $U_t = \rho U_{t-1} + \epsilon_t$ .... (2)

where $|\rho| < 1$

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

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

Which gives Generalised difference equation (3)

$$Y_t - \rho Y_{t-1} = \beta_1(1-\rho) + \beta_2 (X_t - \rho X_{t-1}) + \epsilon_t \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

Where $\epsilon_t = U_t - \rho U_{t-1}$

Since ϵ_t satisfies all the assumptions for OLS, we can run this regression given in (3) above.

Here due to differences one observation on X and Y is lost. To avoid this we can transform the first observation on Y as $Y_1\sqrt{1-\rho^2}$ and first observation on X as $X_1\sqrt{1-\rho^2}$, Which solves the difficulty.

Note How to know ρ ?

(1) Since $d \approx 2(1-\rho) \Rightarrow \rho \approx 1 - \frac{d}{2}$

which gives the values of ρ for large Samples

(2) We can use Theil Nagar formula for obtaining ρ for small Samples.

$\widehat{\rho} = \frac{n^2\left(1-\frac{d}{2}\right)+K^2}{n^2-K^2}$ $n =$ Sample size

$d =$ DW statistic

$k =$ Number of Coefficients including intercept.

Case 2 When ρ is unknown

1. First Difference Method

Let $Y_t = \beta_1 + \beta_2 X_t + U_t$

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

$\therefore \Delta Y_t = \beta_2 \Delta X_t + \epsilon_t$ Where $\epsilon_t = U_t - U_{t-1}$

This equation can be regressed. Note that there is no intercept term here.

Similarly if original equation is

$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 t + U_t$$

Then $\Delta Y_t = \beta_2 \Delta X_{t-1} + \beta_3 + \Delta U_t$

Here coefficient of t, i.e. β_3 becomes intercept term in the transformed model. Run this regression as usual.

2. Moving Average Regression Model (MAR Model)

Consider the Generalized Difference Equation (GDE) given by

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

If we assume that $\rho = -1$, i.e. there is perfect negative autocorrelation then the above equation gives

$\frac{Y_t + Y_{t-1}}{2} = \beta_1 + \beta_2 \left(\frac{X_t + X_{t-1}}{2}\right) + \frac{\epsilon_t}{2}$

This is moving average regression model using 2 periods as periodicity to workout.

Run this regression as usual.

Note This method works out if $\rho = -1$. We need to test Whether $\rho = -1$ or not. For that there is an advanced method known as Berenblutt-Webb (BW) test.

3. Cochran Orcutt Iterative procedure

Consider the model

$Y_t = \beta_1 + \beta_2 X_t + U_t$ .... (1)

with auto regression (AR) Scheme given by

$U_t = \rho U_{t-1} + \epsilon_t$ .... (2)

and as in earlier case, We have GDE given by

$Y_t - \rho Y_{t-1} = \beta_1(1-\rho) + \beta_2 (X_t - \rho X_{t-1}) + (U_t - \rho U_{t-1})$.................. (3)

Here β_1, β_2, ρ all are unknown. We apply the following iterative procedure.

(1)Run the regression equation (1), find OLSE of β_1 and β_2 and hence find $\widehat{U}_t$

(2)Now using Values of $\widehat{U}_t$ run the regression (2) and estimate ρ given by $\hat{\rho}_1$.

$(\widehat{U}_t = \hat{\rho} U_{t-1} + U_t)$

(3) Now run the regression (3) using $\hat{\rho}_1$obtained above.

We write equation (3) as $Y_t^* = \beta_1^* + \beta_2^* X_t^* + e_t^*$

Here $\beta_1^* = \beta_1(1-\hat{\rho}_1)$

This regression gives $\hat{\beta}_1^*$ and$\hat{\beta}_2^*$, hence we put in (1) and get new residuals

$U_t^{**} = Y_t - \hat{\beta}_1^* - \hat{\beta}_2^* X_t$..................................... (4)

These can be obtained easily as now Y_t, X_t, $\hat{\beta}_1^*$ and $\hat{\beta}_2^*$ are known.

(4)Now use equation (2) again to obtain new value of $\hat{\rho}$ (denoted by$\hat{\rho}_2$) using the relation

$U_t^{**} = \hat{\rho}_2 U_{t-1}^{**} + W_t$ .... (5)

This is second round estimate of ρ given by $\hat{\rho}_2$. Compare $\hat{\rho}_2$ with $\hat{\rho}_1$ if they are very near, our procedure Stops here. (The difference should be by less than 0.01 or 0.005)

(5) If $\hat{\rho}_2$ differs largely from $\hat{\rho}_1$, repeat the above stated procedure. Thus we get successive iterations which give us final outcome.

Note:

(1) The above iterative procedure ends in a finite number of steps. We can compute d and test about autocorrelation.

(2) Besides this method, there are other models also which are named here as

(i) Cochran-orcutt two stage iterative procedure

(ii) Dubin's two step method of estimating ρ etc.

(3) All these methods work out in 2 or 3 stages. In literature these methods of estimation are called Feasible or Estimated Generalized Least Squares (EGLS) methods.