Lecture 16. Autocorrelation

16.1. Introduction

In Lectures 14 and 15, we considered the estimation of reduced form equations by OLS. You will recall that one of the Classical Assumptions of OLS is that the disturbances are not autocorrelated ie. $E[u_t u_s] = 0$ for all $t \neq s$ (see page 6 of the handout for Lecture 14). In this lecture, we consider in greater detail what happens, and what we can do, when the disturbances *are* autocorrelated ie. $E[u_t u_s] \neq 0$ for some $t \neq s$. The material is organised as follows:

Section 16.2. The causes of autocorrelated disturbances
Section 16.3. The consequences of autocorrelation for the OLS estimator
Section 16.4. Some famous models of the processes which generate autocorrelation
Section 16.5. Testing for autocorrelated disturbances
Section 16.6. The Cochrane-Orcutt iterative method of estimation in the presence of autocorrelation
Section 16.7. An interpretation of autocorrelation as 'dynamic misspecification'

You should note that another frequently used term for 'autocorrelation' is 'serial correlation'. We will use these terms interchangeably in what follows. When a variable is not autocorrelated, we can say that it is 'serially independent'.

16.2. The causes of autocorrelated disturbances

It is logical to begin by briefly considering how autocorrelated disturbances can come about. There are three main ways you need to be aware of:

(a). Firstly, recall that one of the reasons for including an error term in an equation is to represent the influence of omitted variables (those that we do not know about, cannot measure, or just do not have data on). If any of these omitted variables exhibit autocorrelation, so will u_t.

(b). Autocorrelation in the errors may also be due to the fact that the effect of a random shock to the system is not *instantaneous* (ie. specific to one particular point in time), but rather persists for several periods after its occurrence.

(c). A third way in which autocorrelated errors may arise is as a result of transformations applied to an equation in deriving the form that is estimated. An illustration of this is the 'trick' applied to stochastic adaptive expectations models in order to derive the final equation (see Lecture 13, Part 2, page 6).

16.3. The consequences of autocorrelation for the OLS estimator

The consequences of autocorrelation in the errors depend on the nature of the explanatory variables in the equation. You will recall from Lecture 14 that, in the case where lagged endogenous variables are present on the right hand side of the equation, OLS estimates are both biased and inconsistent. The situation is very different when all the explanatory variables are exogenous. OLS estimates remain unbiased and consistent in this case. However, several other problems arise:

(a). ***They are no longer the 'best' estimators, in the sense of having minimum variance.*** There is now some other estimator, to be discussed below, which has a smaller sampling variance ie. is more 'efficient'.

(b). ***The way in which the variances and covariances of OLS estimates are calculated assuming serial independence is wrong if the errors are autocorrelated.*** The OLS calculations are based on taking $\sigma_u^2(\mathbf{X}'\mathbf{X})^{-1}$ as the variance-covariance matrix of $\hat{\beta}$, whereas the true variance-covariance matrix in this case is given by

$$E[(\hat{\beta}-\beta)(\hat{\beta}-\beta)'] = (X'X)^{-1}X'E[uu']X(X'X)^{-1} = (X'X)^{-1}X'VX(X'X)^{-1}$$

where V is the variance-covariance matrix of the errors (see Lecture 15, page 6). If the u_t are positively autocorrelated, then the OLS formula is likely to underestimate the true variances. Hence, t-ratios may be larger than they should be, and we may be overoptimistic about the significance of coefficients.

(c). ***The OLS formula for the F-test we considered in Lecture 15 is also incorrect.***

(d). ***The OLS sum of squared residuals (SSR) underestimates the 'true' sum of squared residuals in the presence of positive autocorrelation, so we get an overoptimistic value for R^2***

(Recall from Lecture 14 that R^2 is computed as $R^2 = 1 - \dfrac{\sum_{t=1}^{n} e_t^2}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}$).

16.4. Some famous models of the processes which generate autocorrelation

To simplify the analysis of autocorrelation, it is necessary to assume some formal representation of the process generating the autocorrelation. ***In practice, these representations are usually autoregressive processes, moving-average processes, or a combination of the two known as mixed autoregressive moving-average processes.***

We consider autoregressive processes for u first. The simplest of these has already been considered in Lecture 13 Part 2 (page 2): it is a first-order autoregressive process

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

where $-1 < \rho < 1$ and the ε_t satisfy the Classical Assumptions including that of serial independence. This is denoted by AR(1). A second-order autoregressive process, AR(2), is given by

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \varepsilon_t$$

and a general autoregressive process, AR(p), by

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \cdots + \phi_p u_{t-p} + \varepsilon_t$$

Next, a general moving-average process, MA(q), is given by

$$u_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

and the combination of the two, called a mixed autoregressive moving-average process and denoted by ARMA(p, q), can be written

$$u_t - \phi_1 u_{t-1} - \phi_2 u_{t-2} - \cdots - \phi_p u_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q}$$

Having seen the various representations used, the obvious question is what do they each imply about what we are actually interested in: the correlations between the u's? Since this is an introductory course, we will only consider the implications of AR(1) and MA(1) processes. The basic techniques carry over to more general models. A convenient way of characterising the nature of the correlations is to look at the correlograms implied by the processes. Correlograms were introduced in Supplementary Lecture 4 Part 1, and we now briefly review this material before proceeding. You might find it helpful to re-read the relevant section of Supplementary Lecture 4 Part 1 at this point.

Standard econometric theory will only be applicable if the disturbances are covariance-stationary. This involves firstly the variances being time-invariant:

$$E[u_t^2] = E[u_{t-k}^2] = \sigma_u^2$$

ie. the errors being homoskedastic, and secondly the covariances being time invariant:

$$E[u_t u_{t-k}] = E[u_s u_{s-k}]$$

ie. the covariance between two u's depending only on their *relative* positions in time, and not on their absolute positions. Combining the two, we can define autocorrelations of the form

$$\frac{E[u_t u_{t-k}]}{\sqrt{\{E[u_t^2]E[u_{t-k}^2]\}}}$$

which if the process is covariance-stationary will be given by

$$r_k \equiv \frac{E[u_t u_{t-k}]}{E[u_t^2]}$$

and called the *autocorrelation with lag k*. The r_k as a function of k are called the *autocorrelation function* for u, and the graph of the autocorrelation function is called the correlogram. Throughout the discussion below, we will assume that the processes are covariance-stationary.

Consider first the AR(1) process. Repeated substitution gives

$$\begin{aligned} u_t &= \rho u_{t-1} + \varepsilon_t \\ &= \rho(\rho u_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\ &\vdots \\ &= \varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + \rho^3\varepsilon_{t-3} + \cdots \\ &= \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-j} \qquad (-1 < \rho < 1) \end{aligned}$$

This is called the moving average representation of u_t. Note that it is an infinite sum. From this we deduce that u_{t-1} depends on $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots$. Hence, ε_t is independent of u_{t-1}, and indeed all past u's. Thus

$$\begin{aligned} E[u_t u_{t-1}] = E[(\rho u_{t-1} + \varepsilon_t)u_{t-1}] &= E[(\rho u_{t-1}^2 + \varepsilon_t u_{t-1})] \\ &= \rho E[u_{t-1}^2] + E[\varepsilon_t u_{t-1}] \\ &= \rho E[u_{t-1}^2] \end{aligned}$$

and therefore

$$r_1 \equiv \frac{E[u_t u_{t-1}]}{E[u_t^2]} = \rho$$

In general,

$$\begin{aligned} E[u_t u_{t-k}] = E[(\rho u_{t-1} + \varepsilon_t)u_{t-k}] &= E[(\rho u_{t-1}u_{t-k} + \varepsilon_t u_{t-k})] \\ &= \rho E[u_{t-1}u_{t-k}] \end{aligned}$$

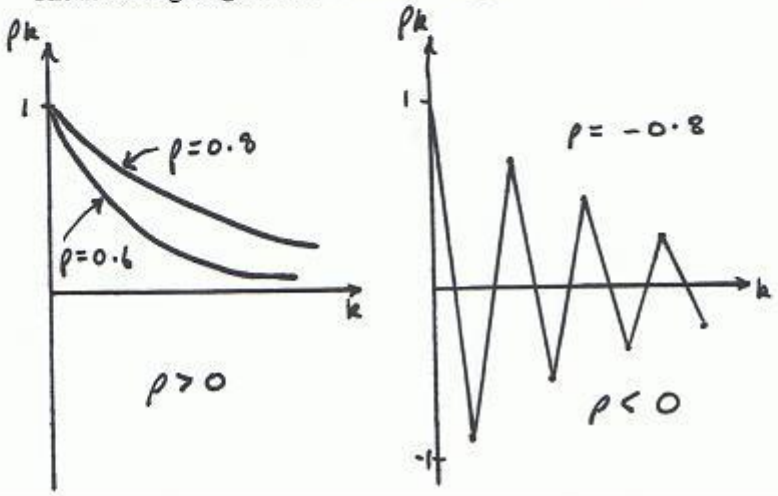
Therefore

$$r_k \equiv \frac{E[u_t u_{t-k}]}{E[u_t^2]} = \rho \frac{E[u_{t-1}u_{t-k}]}{E[u_{t-1}^2]} = \rho \cdot r_{k-1}$$

But $r_{k-1} = \rho r_{k-2}$, etc.., so we end up with the result

$$r_k = \rho^k \qquad k = 0, 1, 2, \ldots$$

The following diagram shows the correlogram for an AR(1) process $(-1 < \rho < 1)$



Next, we consider a first-order moving-average process. For an MA(1) we have

$$\begin{aligned} E[u_t u_{t-1}] &= E[(\varepsilon_t - \theta_1\varepsilon_{t-1})(\varepsilon_{t-1} - \theta_1\varepsilon_{t-2})] \\ &= E[\varepsilon_t\varepsilon_{t-1} - \theta_1\varepsilon_{t-1}^2 - \theta_1\varepsilon_t\varepsilon_{t-2} + \theta_1^2\varepsilon_{t-1}\varepsilon_{t-2}] \\ &= -\theta_1\sigma_\varepsilon^2 \end{aligned}$$

since the ε_t are serially independent, and

$$\begin{aligned} E[u_t^2] = E[(\varepsilon_t - \theta_1\varepsilon_{t-1})^2] &= E[\varepsilon_t^2 - 2\theta_1\varepsilon_t\varepsilon_{t-1} + \theta_1^2\varepsilon_{t-1}^2] \\ &= \sigma_\varepsilon^2 + \theta_1^2\sigma_\varepsilon^2 \\ &= (1+\theta_1^2)\sigma_\varepsilon^2 \end{aligned}$$

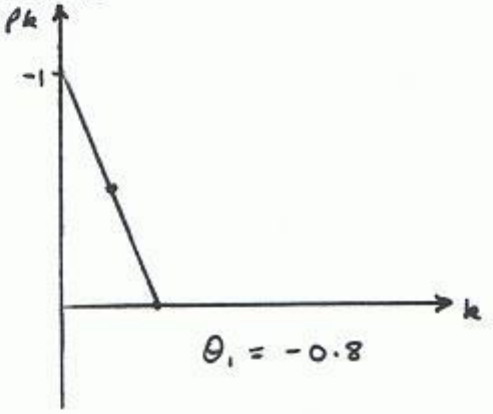
assuming covariance-stationarity. Thus

$$r_1 \equiv \frac{E[u_t u_{t-1}]}{E[u_t^2]} = -\theta_1/(1+\theta_1^2)$$

The other values of the autocorrelation function (ie. r_2, r_3, . . .) are all zero. For example,

$$\begin{aligned} E[u_t u_{t-2}] &= E[(\varepsilon_t - \theta_1\varepsilon_{t-1})(\varepsilon_{t-2} - \theta_1\varepsilon_{t-3})] \\ &= E[\varepsilon_t\varepsilon_{t-2} - \theta_1\varepsilon_t\varepsilon_{t-3} - \theta_1\varepsilon_{t-1}\varepsilon_{t-2} + \theta_1^2\varepsilon_{t-1}\varepsilon_{t-3}] \\ &= E[\varepsilon_t\varepsilon_{t-2}] - \theta_1 E[\varepsilon_t\varepsilon_{t-3}] - \theta_1 E[\varepsilon_{t-1}\varepsilon_{t-2}] + \theta_1^2 E[\varepsilon_{t-1}\varepsilon_{t-3}] \end{aligned}$$

and each of these terms is zero, since the ε_t are serially independent. So the correlogram for an MA(1) process looks as follows:



Theoretical correlograms generated by various error processes are useful for comparison when examining estimated residual correlograms for evidence of certain forms of autocorrelation.

16.5. Testing for autocorrelated disturbances

Having seen earlier the consequences of autocorrelation for OLS estimation, it is clearly desirable to be able to test for its presence. ***A basic difficulty is that assumptions about the form of any autocorrelation are stated in terms of the disturbances, which are obviously unobservable. We therefore have to work with our estimates of the errors: the regression residuals, the e's.*** The most commonly used test employs what is known as the *Durbin-Watson statistic*. This is defined as the ratio of the sum of squares of the first differences of the residuals to the sum of squares of the residuals themselves:

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

The Durbin-Watson test is designed for use with regression equations in which all the explanatory variables are exogenous. To understand why we use this statistic, suppose that the u's are generated by a first-order autoregressive process

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

The OLS estimator of ρ that we obtain if we use the residuals e_t in place of the unobservable error terms u_t is given by

$$\hat{\rho} = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=2}^{n} e_{t-1}^2}$$

Clearly d and $\hat{\rho}$ are related. To see how, expand the numerator of d to give

$$d = \frac{\sum_{t=2}^{n} e_t^2 + \sum_{t=2}^{n} e_{t-1}^2 - 2\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_{t-1}^2}$$

The first two terms in the numerator differ from the denominator only in that they contain n-1 squared residuals, whereas the denominator is the sum of squares of all n residuals. If n is reasonably large, we can neglect this difference, and write

$$d \cong 2\left(1 - \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2}\right)$$

(the symbol $\cong$ means 'approximately equals'). The ratio term in this expression differs from the expression for $\hat{\rho}$ given above only in the number of squared residuals summed in the denominator. Thus, we have the famous result

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

If we have zero autocorrelation ($\rho = 0$) we would expect $d \cong 2$.
If we have positive autocorrelation ($0 < \rho < 1$) we would expect $0 < d < 2$.
If we have negative autocorrelation ($-1 < \rho < 0$) we would expect $2 < d < 4$.

So once we have calculated a value of d, we have to decide whether it is sufficiently far away from 2 to force us to reject the null hypothesis that the errors are serially independent. It is possible to compute significance points with which d can be compared, but a great disadvantage of the Durbin-Watson statistic is that it does not have a single critical point which is valid for all regression problems. ***This is because the distribution of d depends on the particular explanatory variables being used in the regression, so <u>general</u> statistical tables which apply to all regressions cannot be drawn up***. In addition, the d statistic is biased towards 2 in the presence of a lagged dependent variable, and hence may indicate serial independence when autocorrelation is in fact present. Most researchers use the Durbin-Watson statistic as a 'quick-and-dirty' test for serial correlation, before employing more generally applicable and convenient statistics to investigate further.

One such alternative, which can be employed even when there are lagged endogenous variables in the equation, is the so-called *Box-Pierce portmanteau statistic* which you can read about for yourself if you want to. It is not often used, so we will not consider it here. A much more commonly used alternative test, which is asymptotically equivalent to the Box-Pierce portmanteau, is the following. Consider the equation

$$Y_t = \alpha_1 Y_{t-1} + \cdots + \alpha_r Y_{t-r} + \beta_1 X_{1t} + \cdots + \beta_k X_{kt} + u_t \qquad t = 1, \ldots, n$$

which has lagged endogenous variables as well as exogenous variables on the right hand side. Suppose that the errors are generated by a pth-order autoregressive process

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \ldots + \phi_p u_{t-p} + \varepsilon_t$$

Then we wish to test the null hypothesis that $\phi_1 = \phi_2 = \ldots = \phi_p = 0$. ***This can be done by regressing the OLS residual e_t on its first p lags ($e_{t-1}, \ldots, e_{t-p}$) and the original explanatory variables $Y_{t-1}, \ldots, Y_{t-p}, X_{1t}, \ldots, X_{kt}$.*** If the null hypothesis of serially independent errors is correct, then nR^2 (ie. the product of the size of the sample used in the original regression and the R^2 calculated from this regression using the residuals) has asymptotically a χ^2 distribution with p degrees of freedom. ***Hence, we will reject the null hypothesis of serial independence if nR^2 is greater than the selected critical point of the $\chi^2(p)$ distribution.*** This test is known as the *Lagrange Multiplier test* for pth-order autocorrelation.

16.6. The Cochrane-Orcutt iterative method of estimation in the presence of autocorrelation

We now consider what to do if our testing indicates the presence of autocorrelation. Various techniques have been developed, but we will only look at the most famous one. This is called the *Cochrane-Orcutt iterative method* for the estimation of a simple bivariate regression equation

$$Y_t = \alpha + \beta X_t + u_t$$

where the u_t are generated by a first-order autoregressive process

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

The technique can easily be extended to deal with equations containing more than one explanatory variable. It is based on transforming the equation into one with serially independent errors. This is achieved by first lagging the equation and multiplying through by ρ to give

$$\rho Y_{t-1} = \alpha\rho + \beta\rho X_{t-1} + \rho u_{t-1}$$

and then subtracting this from the original to give

$$Y_t - \rho Y_{t-1} = \alpha(1 - \rho) + \beta(X_t - \rho X_{t-1}) + \varepsilon_t$$

(note that the disturbance term is $u_t - \rho u_{t-1} = \varepsilon_t$). ***If ρ were known, since the transformed equation satisfies the Classical Assumptions, we could simply regress $(Y_t - \rho Y_{t-1})$ on $(X_t - \rho X_{t-1})$ using OLS, and this would give us 'best' (ie. minimum variance) linear unbiased estimates. However, ρ is <u>not</u> usually known, and must therefore be estimated.***

The Cochrane-Orcutt iterative method involves using the OLS residuals to provide an estimate of ρ. The estimate $\hat{\rho}$ is then used to generate 'transformed' variables of the form $(Y_t - \hat{\rho} Y_{t-1})$ and $(X_t - \hat{\rho} X_{t-1})$, and the equation is re-estimated using these transformed variables. The residuals from this re-estimation can then be used to get an improved estimate of ρ, and so on. The full procedure is as follows:

(1). Estimate the equation $Y_t = \alpha + \beta X_t + u_t$ by OLS, and and use the estimates of α and β to calculate the residuals $e_1, e_2, \ldots, e_n$. Use these residuals to estimate, by OLS, the first-order autoregressive process $u_t = \rho u_{t-1} + \varepsilon_t$ with the unobservable u_t's replaced by their estimates, the e_t's. This gives

$$\hat{\rho} = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=2}^{n} e_{t-1}^2}$$

(2). Construct transformed variables Y_t^* and X_t^* for $t = 1, \ldots, n$, where

$$Y_t^* = Y_t - \hat{\rho} Y_{t-1}$$

$$X_t^* = X_t - \hat{\rho} X_{t-1}$$

and estimate, by OLS, the equation $Y_t^* = \alpha^* + \beta X_t^* + u_t^*$, where $\alpha^* \equiv \alpha(1 - \hat{\rho})$. We can obtain an estimate of α from the resulting estimate of α^* using the formula

$$\hat{\alpha} = \frac{\hat{\alpha}^*}{(1-\hat{\rho})}$$

Using the revised estimates of α and β, calculate the new residuals $e_1^*, e_2^*, \ldots, e_n^*$ by substitution into the original equation:

$$e_t^* = Y_t - \hat{\alpha} - \hat{\beta} X_t$$

Use these in turn to obtain a new estimate of ρ:

$$\hat{\rho} = \frac{\sum_{t=2}^{n} e_t^* e_{t-1}^*}{\sum_{t=2}^{n} e_{t-1}^{*2}}$$

(3). Construct new transformed variables $(Y_t - \hat{\rho} Y_{t-1})$ and $(X_t - \hat{\rho} X_{t-1})$ and repeat step (2). Continue the iterations until the estimate of ρ converges (ie. until the estimate of ρ changes by less than some specified small number in successive iterations). When the estimate of ρ has converged, it can be used to estimate the equation $Y_t - \rho Y_{t-1} = \alpha(1 - \rho) + \beta(X_t - \rho X_{t-1}) + \varepsilon_t$ by OLS. Since ε_t satisfies the Classical Assumptions, the resulting estimates of α and β will be BLUE.

Variants of this method which you might like to read about for yourself are the Cochrane-Orcutt Two-Step Method, Durbin's Two-Step Method, The Prais-Winsten Modification, The Hildreth-Lu Search Method, and direct estimation of the transformed equation by non-linear least squares or maximum likelihood.

16.7. An interpretation of autocorrelation as 'dynamic misspecification'

So far, we have viewed autocorrelation as a 'nuisance', since all the t-ratios will be miscalculated if we do not take it into account by using one of the estimation methods in Section 16.6. In this section, we consider a slightly different interpretation of autocorrelation. It will be presented in the context of the simple bivariate model with an error generated by a first-order autoregressive process:

$$Y_t = \alpha + \beta X_t + u_t$$
$$u_t = \rho u_{t-1} + \varepsilon_t$$

Recall from the last section that we can transform the equation into one with a serially independent error term, ε_t:

$$Y_t = \alpha(1 - \rho) + \rho Y_{t-1} + \beta X_t - \beta\rho X_{t-1} + \varepsilon_t$$

This can be stated as a dynamic linear regression equation with a non-linear restriction on its parameters:

$$Y_t = \gamma_1 + \gamma_2 Y_{t-1} + \gamma_3 X_t + \gamma_4 X_{t-1} + \varepsilon_t$$

subject to

$$\gamma_2\gamma_3 + \gamma_4 = 0.$$

To begin with, suppose that this restriction is true. ***Then the bivariate model with an autoregressive error can be regarded as a simplification of the more general dynamic model. It reduces the number of parameters to be estimated from four to three, thus yielding more efficient estimates.*** (The fewer parameters you estimate, the more 'efficient' your estimates will be, because there will be more 'degrees of freedom'. Recall from Lecture 15, page 5, that the

degrees of freedom are computed as (n-k), where k is the number of parameters being estimated. The term (n-k) appears in the denominator of formula for the standard error of an estimate. Thus, the smaller is k, the smaller are the standard errors). ***Thus, an autoregressive error process can sometimes be used to simplify the dynamic specification of an equation.***

However, suppose now that the above restriction is not true. In other words, suppose that the correct specification of the equation is

$$Y_t = \gamma_1 + \gamma_2 Y_{t-1} + \gamma_3 X_t + \gamma_4 X_{t-1} + \varepsilon_t$$

without any restriction on the parameters. Then if we estimate the simple bivariate model without any lags

$$Y_t = \alpha + \beta X_t + u_t$$

we will probably obtain a significant value of the Durbin-Watson statistic, since u will 'pick up' the effect of the omitted lagged Y and X. ***Y is obviously autocorrelated if the true model is*** $Y_t = \gamma_1 + \gamma_2 Y_{t-1} + \gamma_3 X_t + \gamma_4 X_{t-1} + \varepsilon_t$ ***(and X is probably autocorrelated too, as are most economic time series), so*** u_t ***will also be autocorrelated if we estimate the 'incorrect' model*** $Y_t = \alpha + \beta X_t + u_t$. The equation $Y_t = \alpha + \beta X_t + u_t$ plus the assumption that the u's are generated by an AR(1) process is a better approximation to the 'correct' specification than the equation $Y_t = \alpha + \beta X_t + u_t$ with the assumption that the u's are purely random. Thus, the Durbin-Watson statistic will be significantly different from 2. ***The actual problem with this equation, however, is not that the errors follow an AR(1) process. The problem is one of misspecification, in particular the omission of*** X_{t-1} ***and*** Y_{t-1}. ***Thus, in this instance, a significant Durbin-Watson statistic is in fact indicating 'dynamic misspecification'.***

This explains why the addition of lagged dependent and lagged explanatory variables on the right hand side of a time series model (eg. the linear partial adjustment model in the applied econometrics exercise) often significantly improves the performance of the model.

(End of Lecture 16)

Assignment for Lecture 16. Autocorrelation

Please attempt the following problem. You will get very similar ones in your tests and exams. The solution is attached, but you should not look at it until you have tried your best to solve the problem on your own.

Consider the following dynamic model:

$$Y_t = \beta Y_{t-1} + u_t \qquad -1 < \beta < 1$$

Suppose the disturbance term follows a first-order autoregressive process of the form

$$u_t = \rho u_{t-1} + \varepsilon_t \qquad -1 < \rho < 1$$

Show that $Cov[u_t, Y_{t-1}] = \rho\sigma_u^2/(1-\rho\beta)$. *(Hint: you need to use the formula for the sum of an infinite geometric series. An infinite geometric series can be written in the form* $a + ar + ar^2 + ar^3 + + ar^4 + \ldots$ *Provided that* $|r| < 1$, *this sum is equal to* $a/(1-r)$). ***Briefly describe the consequences of this for the Ordinary Least Squares estimate of*** β.