HANDOUT 04

TIME SERIES ANALYSIS II

Outline of autocorrelation lecture:

I. Integrated Stochastic Processes

Recall the following pure random walk process:

$$Y_t = Y_{t-1} + u_t$$

where u_t is a white noise with mean 0 and variance σ^2.

We can also write as follows:

$$Y_t - Y_{t-1} = +u_t$$

or

$$\Delta Y_t = +u_t$$

where Δ is the first difference operator. Since u_t is a white noise, we know that:

$$E(u_t)=0$$
$$Var(u_t)=\sigma^2$$
$$Cov(u_t,u_{t-s})=0,\ t\neq s$$

Then ΔY_t is a stationary series. In other words, the first difference of a random walk series is stationary. Therefore we call random walk without drift *integrated of order 1*, denoted as I(1).

Likewise, if a time series has to be differenced twice (first difference of the first differences) to make it stationary, we call such a series *integrated of order 2*.

For example if Y_t is I(2) then

$$\underbrace{\Delta\Delta Y_t}_{\Delta^2 Y_t}=\Delta(Y_t-Y_{t-1})$$

will become stationary. Here, note that $\Delta\Delta Y_t=\Delta^2 Y_t\neq Y_t-Y_{t-2}$

In general, if a stationary time series Y_t has to be differenced *d* times to make it stationary, that time series is said to be integrated of order *d* and denoted as $Y_t\sim I(d)$.

If a time series Y_t is *stationary* (i.e. if it does not require any differencing to be stationary) it is said to be *integrated of order zero*, denoted by $Y_t\sim I(0)$.

Hence, we can use the terms "stationary time series" and "time series integrated of order zero" interchangeable.

Note that <u>*most*</u> economic time series generally *I(1)*; that is, they become stationary only after taking their first differences.

Properties of Integrated Series

Let Y_t and X_t are two time series.

(1) If $X_t \sim I(0)$ and $Y_t \sim I(1)$, then

$$Z_t = (X_t + Y_t) \sim I(1)$$

That is, a linear combination of sum of stationary and nonstationary time series is nonstationary.

(2) If $X_t \sim I(d)$, then;

$$Z_t = a + bX_t \sim I(d)$$

That is, a linear combination of an $I(d)$ series is also called $I(d)$. Thus if $X_t \sim I(0)$, then;

$$Z_t = a + bX_t \sim I(0)$$

(3) If $X_t \sim I(d_1)$ and $Y_t \sim I(d_2)$ $d_2 > d_1$where

Then;

$$Z_t = aX_t + bY_t \sim I(d_2)$$

(4) If $X_t \sim I(d)$ and $Y_t \sim I(d)$,

then;

$$Z_t = aX_t + bY_t \sim I(d^*)$$

where d^* is generally equal to d, but in some cases $Z_t \sim I(0)$. This is the situation of *cointegration!*.

The integration of series is quite important. To understand, consider $Y_t = \beta_0 + \beta_1 X_t + u_t$. We know that $\hat{\beta}_1 = \frac{\sum x_t y_t}{\sum x_t^2}$. Suppose that $Y_t \sim I(0)$ but $X_t \sim I(1)$. Hence $\hat{\beta}_1 \sim I(1)$ which means the distribution of $\hat{\beta}_1$ will not have a constant variance and/or mean. Thus the $\hat{\beta}_1$ will not even have an asymptotic distribution.

The researcher has to be very careful in working nonstationary series. Consider $Y_t = \beta_0 + \beta_1 X_t + u_t$, there are 4 possible cases:

	Case 1 Both X_t and Y_t are stationary. ✓ In this case the classical regression model is appropriate.
Meaningless regression	**Case 2** X_t and Y_t are integrated of different orders. ✓ Regression equation using such variables are meaningless
Meaningless regression	**Case 3** Nonstationary X_t and Y_t are integrated of the same order and the residuals $(\hat{u}_t)$ is nonstationary. ✓ This is the spurious regression situation and hence the results from such spurious regressions are meaningless.
	Case 4 Nonstationary X_t and Y_t are integrated of the same order and the residual is also $(\hat{u}_t)$ stationary. ✓ In this case, X_t and Y_t are *cointegrated*. We will see the details of this situation later on.

II. Stochastic Trends, AR Models, and Unit Root

Consider the following autoregressive model for Y_t:

$Y_t = a_1 Y_{t-1} + u_t$ [First-order autoregressive model, AR(1)]

where u_t is a *white noise*.

As known an autoregressive model is a regression model that relates a time series variable to its past values.

Hence the random walk model $Y_t = a_1 Y_{t-1} + u_t$ is a special case of an AR(1) model in which $a_1 = 1$.

In other words, if Y_t follows an AR(1) with $a_1 = 1$, then Y_t contains a stochastic trend and it is nonstationary.

If Y_t, we face what is known as the unit root problem, that is, a situation of nonstationarity. The name unit root is due to the fact that $a_1 = 1$.Thus the terms *stochastic trend*, *unit root* and *random walk* can be used interchangeably.

If however $|a_1| < 1$, then it can be shown that the time series Y_t is stationary. This finding is important to test for unit root. The hypothesis that Y_t has a stochastic trend can be tested by testing

$$H_0 : a_1 = 1$$
$$H_A : a_1 < 1$$

in $Y_t = a_0 + a_1 Y_{t-1} + u_t$. If $a_1 = 1$, the AR(1) has a unit root, and Y_t is nonstationary.

III. The Unit Root Test

Consider the following AR(1) process for Y_t:

(1) $Y_t = \rho Y_{t-1} + u_t$ where $-1 \le \rho \le 1$ and u_t is a white noise error.

We know that if $\rho = 1$ (the case of unit root) Equation (1) becomes a pure random walk model which we know is a nonstationary process.

<u>Idea:</u> Simply regress Y_t on Y_{t-1} and test if ρ is equal to 1. If it is, then Y_t is nonstationary.

However we cannot estimate (1) by OLS and test the hypothesis that $\rho = 1$ by a usual t test since that test is *<u>severely biased</u>* in the case of unit root.

Therefore, we manipulate equation (1) as follows:

$$
\begin{aligned}
&Y_t = \rho Y_{t-1} + u_t \\
&Y_t - Y_{t-1} = \rho Y_{t-1} - Y_{t-1} + u_t \\
&\Delta Y_t = (\rho - 1) Y_{t-1} + u_t
\end{aligned}
$$

or

(2) $\Delta Y_t = \delta Y_{t-1} + u_t$ where $\delta = \rho - 1$

In practice, therefore, we estimate equation (2) and test the (null) hypothesis that $\delta = 0$, versus the alternative hypothesis that $\delta < 0$ (stationary, since $\delta < 0 \Rightarrow \rho - 1 < 0 \Rightarrow \rho < 1$)

So

$H_0 : \delta = 0$ [nonstationarity]
$H_A : \delta < 0$ [stationarity]

Note that since the process is nonstationary under H_0; the usual t distribution cannot be used. since when Y_t is nonstationary the estimated coefficient of Y_{t-1} does not follow the t distribution even in large samples; in other words, it does not have an asymptotic distribution.

What is the alternative? Dickey and Fuller have shown that under null hypothesis that $\delta = 0$, the estimated t value of the coefficient of Y_{t-1} in equation (2) follows the τ *(tau) statistic*. They computed the critical values of the tau statistic by *Monte Carlo simulations*.

In the literature, the *tau statistic* or test is known as the *Dickey-Fuller (DF) test*. Keep in mind that DF test is a one sided test.

Recall that a random walk process may have different forms:

(1) Pure random walk	$Y_t = \rho Y_{t-1} + u_t$ $Y_t - Y_{t-1} = (\rho - 1) Y_{t-1} + u$ $\Rightarrow \Delta Y_t = \delta Y_{t-1} + u_t$
(2) Random walk with drift	$Y_t = a_0 + \rho Y_{t-1} + u_t$ $Y_t - Y_{t-1} = a_0 + (\rho - 1) Y_{t-1} + u_t$ $\Rightarrow \Delta Y_t = a_0 + \delta Y_{t-1} + u_t$
(3) Random walk with drift and a deterministic trend	$Y_t = a_0 + \rho Y_{t-1} + a_2 t + u_t$ $Y_t - Y_{t-1} = a_0 + (\rho - 1) Y_{t-1} + a_2 t + u_t$ $\Rightarrow \Delta Y_t = a_0 + a_2 t + \delta Y_{t-1} + u_t$

The methodology is precisely the same regardless of which of the three forms of the equations is estimated. However, be aware that the *critical values* of the *tau statistic* <u>*do depend*</u> on *whether the process has an intercept and/or a deterministic trend term*.

Consequently;

(1) if we estimate (1) we will look at τ_{nc} where *nc* refers to "*no constant*",
(2) if we estimate (2) we will look at τ_c where *c* refers to "*constant*",
(3) if we estimate (3) we will look at τ_{ct} where *ct* refers to "*constant and trend*".

Note that in this DF test, we have assumed that the time series Y_t can be modeled as s first order AR process with a disturbance that is white noise.

But what happens if this not the case? Suppose that the disturbance is not white noise or suppose the time series is, in fact, the result of a higher order process: AR(p)!

Either of the possibilities is likely to show up as an autocorrelation problem in the residuals of the OLS estimated versions of equations (1), (2) and (3). Unfortunately, the above Dickey-Fuller test is *invalid* in these circumstances.

Two approaches have been suggested for tackling this problem. First; we can modify the actual testing procedure by generalizing DF test. Secondly, it is possible to retain equation (1), (2) and (3) but adjust the DF statistic to allow for autocorrelated residuals. The first approach leads to the Augmented Dickey-Fuller (ADF) test which we now consider. The second option refers to the Phillips and Perron (PP) unit root test.

IV. Augmented Dickey-Fuller Test

As stated above if there is an autocorrelation problem in the residuals of the OLS estimated versions of equation (1), (2) and (3), then the DF test, is invalid.

(1) $Y_t = \rho Y_{t-1} + u_t \Rightarrow \Delta Y_t = \delta Y_{t-1} + u_t$

(2) $Y_t = a_0 + \rho Y_{t-1} + u_t \Rightarrow \Delta Y_t = a_0 + \delta Y_{t-1} + u_t$

(3) $Y_t = a_0 + \rho Y_{t-1} + a_2 t + u_t \Rightarrow \Delta Y_t = a_0 + a_2 t + \delta Y_{t-1} + u_t$

This autocorrelation is likely to occur if our models [equation (1), (2) and (3)] did not have *sufficient lag terms* to capture the full dynamic nature of the process. Hence for models (1), (2) and (3); the extended test are as follows:

(1') $Y_t = \rho Y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta Y_{t-i} + u_t$

$\Rightarrow \Delta Y_t = \delta Y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta Y_{t-i} + u_t$

(2') $Y_t = a_0 + \rho Y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta Y_{t-i} + u_t$

$\Rightarrow \Delta Y_t = a_0 + \delta Y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta Y_{t-i} + u_t$

(3') $Y_t = a_0 + \rho Y_{t-1} + a_2 t + \sum_{i=1}^{p} \beta_i \Delta Y_{t-i} + u_t$

$\Rightarrow \Delta Y_t = a_0 + a_2 t + \delta Y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta Y_{t-i} + u_t$

A problem with the ADF test is that, we do not know beforehand what order of AR process must be used, i.e., we do not know p!

In practice, the usual approach followed is to include as many lags which are necessary to produce non-autocorrelated OLS residuals.

The LM tests for autocorrelation are usually used for this purpose. That is, a poor value for an LM statistic is regarded as indicating a need for extra lagged terms to be included. Restating the need for extra lags is as follows:

(1) The order of original equations (1), (2) and (3) may be misspecified.

(2) There can be a disturbance term in equations (1), (2) and (3) which is not white noise but is genuinely autocorrelated.

In practice, the choice of augmentation of the autoregressive component $\sum_{i=1}^{p} \beta_i \Delta Y_{t-i}$ is of the utmost importance and is often neglected in the literature. Since the primary goal of the inclusion of the (lag terms) augmentation terms is to save a *white noise* property for u_t, the emphasis should be put into applying a series of test for ensuring that the series u_t is indeed independently and identically distributed, iid (which implies non-autocorrelated) (Cheremza).

A simple guide for choosing the appropriate lag length (p) is to apply the *general to specific* procedure: start by selecting a reasonably large value for p (it may be 12) and then systematically reduce the number of augmentations by imposing and testing zero restrictions and testing for autocorrelation in the restricted model. In most cases the deletion of insignificant augmentations does not affect the property of lack of autocorrelation of the residuals

Hence, it is normally possible to finish up with a model where only significant augmentations (lags) remain and, at the same time, no autocorrelation is present in the residuals. It is therefore important to remember that, if the total length of augmentation (lag length) is *p*, the actual number of augmentation terms which appear might be smaller than p, since some of the coefficients β_i can be zeros.

In other words, we can start with a large value of p and then sequentially reducing the lag order by testing for the significance of the coefficient of the largest lag and testing for autocorrelation.

Another method to select the lag order *p* is using an information criterion for various *p* and select the one with lowest information criteria. However, the model selected for this *p* must be checked for AC by an LM test. The selected model must not have autocorrelation. If there is AC, the model with second lowest information criteria can be checked for AC and so on. Eviews has an option that automatically selected the lag length based on AIC, SBC and other information criteria.

Studies of the ADF statistic suggest that it is better to have too many lags then too few, so it is recommended to use AIC instead of SBC. Note that in ADF test we still test whether $\delta = 0$ and the ADF test follows the same asymptotic distribution as the DF statistic, so the same critical values can be used.

V. Testing for the Order of Integration

A test for the order of integration is a test for the number of unit roots, and follows these steps:

Step 1 Test Y_t to see if stationary. If yes, then $Y_t \sim I(0)$. If no, then go to step 2.

Step 2 Take first differences of Y_t as $\Delta Y_t = Y_t - Y_{t-1}$ and test to see if ΔY_t is stationary. If yes, then $\Delta Y_t \sim I(0)$, which implies $Y_t \sim I(1)$. If no, then go to step 3.

Step 3 Take second differences of Y_t as $\Delta^2 Y_t = \Delta Y_t - \Delta Y_{t-1}$ and test $\Delta^2 Y_t$ to see if it is stationary. If yes, then $\Delta^2 Y_t \sim I(0)$, which implies that $Y_t \sim I(2)$. If no then go to next step and proceed so on until it is found to be stationary and then stop. For example if $\Delta^3 Y_t \sim I(0)$ then $Y_t \sim I(3)$, which means that Y_t needs to be differenced three times to become stationary.

Example

T	Y_t	ΔY_t	$\Delta^2 Y_t$	$\Delta^3 Y_t$
1	100	-	-	-

2	250	150	-	-
3	320	70	-80	-
4	410	90	20	100
5	600	190	100	80

VI. Numerical Example for DF and ADF Tests

To allow for the various possibilities, the DF test is estimated in three different forms, that is, under 3 different null hypothesis:

Y_t is a (pure) random walk • This is a test for a random walk against a stationary autoregressive process of order one (AR(1))	$\Delta Y_t = \delta Y_{t-1} + u_t$	(1)
Y_t is a random walk with drift • This is a test for a random walk against a stationary AR(1) with drift	$\Delta Y_t = a_0 + \delta Y_{t-1} + u_t$	(2)
Y_t is a random walk with drift and a deterministic trend • This is a test for a random walk against a stationary AR(1) with drift and a time trend.	$\Delta Y_t = a_0 + a_2 t + \delta Y_{t-1} + u_t$	(3)

For each case the hypotheses are:

$H_0 : \delta = 0$ [Y_t is nonstationarity[1]]
$H_A : \delta < 0$ [Y_t is stationarity[2]]

The test statistic does not follow the usual *t*-distribution under the null, since the null is one of non-stationarity, but rather follows a non-standard distribution. Critical values are derived from Monte Carlo experiments in, for example, Fuller (1976).

Note that DF table values are always negative, so to exceed to critical value, the $t^*_{\hat{\delta}}$ (or $\tau_{\hat{\delta}}$) value <u>must be more negative</u> than the DF table

[1] Or: there is unit root, Y_t has a stochastic trend.
[2] Y_t is stationary possibly around a deterministic trend.

value. This test is non symmetrical so we do not express the test decision rule using absolute value.

If $t^*_{\hat{\delta}} < DF_{table} \Rightarrow$ Reject H_0 that $\delta = 0 \rightarrow Y_t$ is stationary

or $\tau_{\hat{\delta}}$

where $t^*_{\hat{\delta}} = \dfrac{\hat{\delta}}{se(\hat{\delta})}$

Table K 1% and 5% Critical Dickey–Fuller t (=τ) values for unit root tests

Sample size	t_{nc}^* 1%	t_{nc}^* 5%	t_c^* 1%	t_c^* 5%	t_{ct}^* 1%	t_{ct}^* 5%
25	−2.66	−1.95	−3.75	−3.00	−4.38	−3.60
50	−2.62	−1.95	−3.58	−2.93	−4.15	−3.50
100	−2.60	−1.95	−3.51	−2.89	−4.04	−3.45
250	−2.58	−1.95	−3.46	−2.88	−3.99	−3.43
500	−2.58	−1.95	−3.44	−2.87	−3.98	−3.42
∞	−2.58	−1.95	−3.43	−2.86	−3.96	−3.41

**Subscripts nc, c, and ct denote, respectively, that there is no constant, a constant, and a constant and trend term.*
Source: Gujarati

Suppose that you are given the following results:

$$\Delta\hat{Y}_t = \underset{(-1.05)}{-0.0066} \underset{(-1.49)}{-0.190} Y_{t-1} \qquad t = 1966...1989$$

$$R^2 = 0.09 \qquad LM[AR(1)] = 4.30 \qquad LM[AR(2)] = 5.41$$

We see that $t^*_{\hat{\delta}} = -1.49$ T=24.

This is a random walk with drift process. So we will look at DF table values for t_c^* at 0.05 level of significance with T=24.

The DF critical value for T=25 is -3.00.

$t^*_{\hat{\delta}} > t^*_c$ since $-1.49 > -3.00$, hence we do not reject null hypothesis that δ=0. In other words, Y_t is nonstationary.

However, this DF test is invalid since LM statistics given for the model do suggest an autocorrelation problem for the residual. $\chi^2_{0.05}$ critical values for $LM[AR(1)]$ and $LM[AR(2)]$ are (with 1 and 2 degrees of freedom) 3.84 and 5.99, respectively.

To remove AC problem, we need to use ADF test. You are given the following regression results:

$$\Delta\hat{Y}_t = \underset{(-1.54)}{-0.0091} \underset{(-2.28)}{-0.288} Y_{t-1} + \underset{(2.19)}{0.445} \Delta Y_{t-1}$$

$$R^2 = 0.26 \qquad LM[AR(1)] = 1.54 \qquad LM[AR(2)] = 1.55$$

As can be seen the LM statistics are now both well below their critical values, so now there is no autocorrelation in residuals. Consequently now the ADF test is valid.

Still $t^*_{\hat{\delta}} > t^*_c$ since $-2.28 > -3.00$, hence we do not reject null hypothesis that δ=0. In other words, we do not reject the nonstationarity hypothesis: Y_t is nonstationary.

However as we have repeatedly noted, the first difference of an economics series is frequently stationary.

We therefore now test ΔY_t for stationarity. We can again start by DF test.

$$\Delta^2\hat{Y}_t = \underset{(-0.07)}{-0.0003} \underset{(-3.46)}{-0.7181} \Delta Y_{t-1} \qquad t = 1966...1989$$

$$R^2 = 0.35 \qquad LM[AR(1)] = 3.26 \qquad LM[AR(2)] = 3.32$$

$t^*_{\hat{\delta}} > t^*_c$ since $-3.46 < -3.00$ ($t^*_{\hat{\delta}}$ is more negative than DF table value), hence we do reject null hypothesis that δ=0. In other words, ΔY_t is stationary.

Thus, since the LM statistics suggest no autocorrelation problem in the residuals, we can conclude that the first difference of Y_t [i.e. ΔY_t] is stationary. This implies that $Y_t \sim I(1)$.

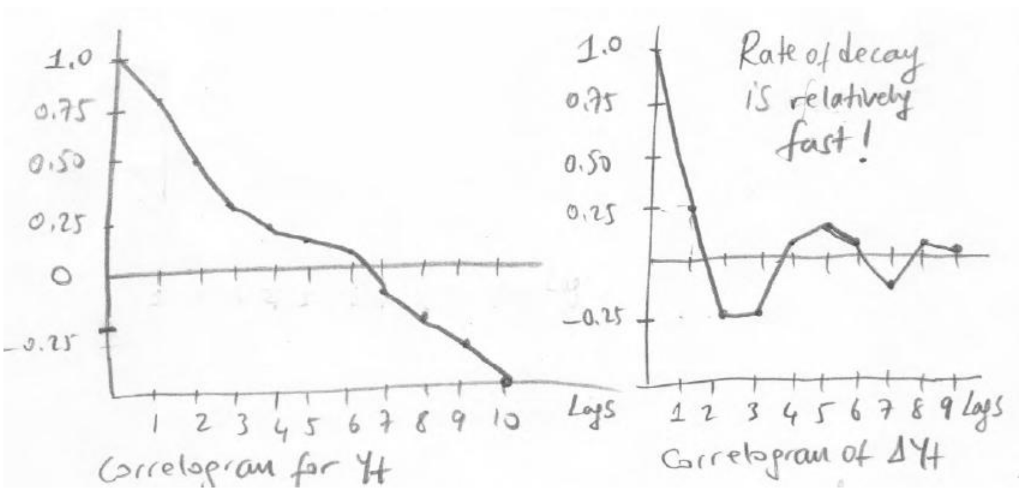
However, in order to confirm this finding we can also estimate an ADF regression:

$$\Delta^2 \hat{Y}_t = \underset{(-0.15)}{-0.0007} \underset{(-4.23)}{-1.022} \Delta Y_{t-1} \underset{(2.09)}{+0.420} \Delta^2 Y_{t-1}$$

$R^2 = 0.46$ $\quad LM[AR(1)] = 0.40$ $\quad LM[AR(2)] = 0.59$

The $t^*_{\hat{\delta}}$ value is now -4.23 and reinforces our conclusion that ΔY_t is stationary. Thus our ADF test confirms and reinforces the DF test.

The correlograms of Y_t and ΔY_t also suggest that $Y_t \sim I(1)$.



References

- Cameron, Samuel (2005) *Econometrics*, McGraw Hill, Berkshire.
- Chauhan, S.P.S., (2009) *Microeconomics: An Advanced Treatise*, Eastern Economy Edition, New Delhi.
- Dougherty, Christopher (2007) *Introduction to Econometrics*, Oxford, New York.
- Erlat, Haluk (1997) *Introduction to Econometrics*, *Chapter 6: Autocorrelation*, Draft (corrected for misprints), Ankara.
- Ezekiel, Mordecai (1938) The Cobweb Theorem, *The Quarterly Journal of Economics*, Vol. 52, No. 2, pp. 255-280.
- Gujarati, D., and Porter (2011) *Basic Econometrics*, McGraw Hill, New York.
- Hill, R. C., Griffiths, W. E., and Judge, G. G., (2001) *Undergraduate Econometrics*, Second Edition, Wiley, New York.
- Kennedy, Peter (1998) *A Guide to Econometrics*, Fourth Edition, Blackwell, New york.
- Stock, J., and Watson, M. M., (2012) *Introduction to Econometrics*, Third Edition, Pearson, Boston.