ADVANCE ECONOMETRICS

Online Lecture Notes

Prepared By:

Dr. Manoj Bhatt

CONTENTS

UNIT-1

DYNAMIC ECONOMETRIC MODELS: AUTOREGRESSIVE AND DISTRIBUTED –LAG MODEL

LESSON-1 AUTOREGRESSIVE AND DISTRIBUTED –LAG MODEL

STRUCTURE

1.1 INTRODUCTION:

In regression analysis involving time series data, if the regression model includes not only the current but also the lagged (past) values of the explanatory variables (the *X*'s), it is called a **distributed-lag model.** If the model includes one or more lagged values of the dependent variable among its explanatory variables, it is called an **autoregressive model.** Thus,

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + u_t$$

represents a distributed-lag model, whereas

$$Y_t = \alpha + \beta X_t + \gamma Y_{t-1} + u_t$$

is an example of an autoregressive model. The latter are also known as **dynamic models** since they portray the time path of the dependent variable in relation to its past value(s).

1.2 OBJECTIVES:

Autoregressive and distributed-lag models are used extensively in econometric analysis, and in this lesson we take a close look at such models with a view to finding out the following:

1. What is the role of lags in economics?

2. What are the reasons for the lags?

3. Is there any theoretical justification for the commonly used lagged models in empirical econometrics?

4. What is the relationship, if any, between autoregressive and

distributed-lag models? Can one be derived from the other?

1.3 THE ROLE OF "TIME", "LAG", IN ECONOMICS:

In economics the dependence of a variable Y (the dependent variable) on another variable(s) X (the explanatory variable) is rarely instantaneous. Very often, Y responds to X with a lapse of time is called a lag**.**

1.3.1 AUTOREGRESSIVE AND DISTRIBUTION-LAG

- In regression analysis involving time series data, if the regression model includes not only the correct but also the lagged (past) value of the explanatory variable (the X's) it is called a distribution lag Model.
- If the model one or more lagged value of the dependent variable among its explanatory variable, it is called an autoregressive model.

  $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + u_t$

  $Y_t = \alpha + \beta X_t + \gamma Y_{t-1} + u_t$ (Autoregressive Model)
- These are used extensively in econometric analysis.
- The partial sums i.e. change in X means value of Y following a unit change in X in the same period, if the change in X is maintained at the same level thereafter, then,(β_0 + β_1X) gives the change in y in the next period so on are called interim, or intermediate, multiplies

  $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots\ldots\ldots\ldots\ldots \beta_k X_{t-k} + u_t$

  (Distribution Log Model with the finite log k time period.)

  [called finite coz the length of the log K is specified]

$$\sum_{i=0}^{k}\beta_i = \beta_0 + \beta_1 + \beta_2 + \ldots\ldots\beta_k = \beta$$

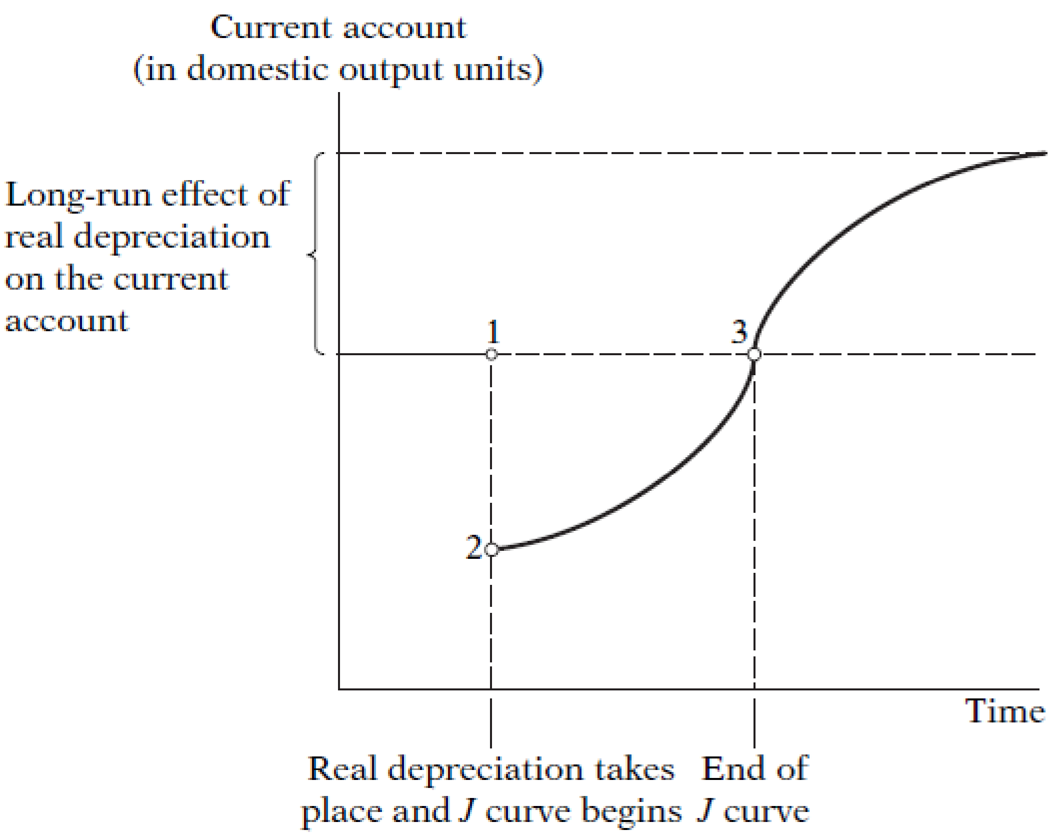
(after k period we obtain long run, or ,total distribution lag multiplies)

$$\therefore \beta_i^* = \frac{\beta_1}{\Sigma\beta_i} = -\frac{\beta_i}{\beta}$$

$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \ldots\ldots\ldots\ldots\ldots \beta_k X_{t-k} + u_t$

(Where length of the lag is not defined, i.e. infinite lag model)

➢ J curve shows relation between trade balance and deprication of currency.



The J curve.

➢ The acceleration principle of involvement theory states that investment is proportional to change in output

$u = \beta (X_t - X_{t-1}) \quad \beta > 0$

↓ ↓ ↓

Involved output output @ time (T-1)

1.3.2 THE REASONS FOR LAGS:

1.3.2.1. Psychological reasons. As a result of the force of habit (inertia), people do not change their consumption habits immediately following a price decrease or an income increase perhaps because the process of change may involve some immediate disutility. Thus, those who become instant millionaires by winning lotteries may not change the lifestyles to which they were accustomed for a long time because they may not know how to react to such a windfall gain immediately. Of course, given reasonable time, they may learn to live with their newly acquired fortune. Also, people may not know whether a change is “permanent’’ or “transitory.’’ Thus, my reaction to an increase in my income will depend on whether or not the increase is permanent. If it is only a nonrecurring increase and in succeeding periods my income returns to its previous level, I may save the entire increase, whereas someone else in my position might decide to “live it up.’’

1.3.2.2 Technological reasons. Suppose the price of capital relative to labor declines, making substitution of capital for labor economically feasible. Of course, addition of capital takes time (the gestation period). Moreover, if the drop in price is expected to be temporary, firms may not rush to substitute capital for labor, especially if they expect that after the temporary drop the price of capital may increase beyond its previous level. Sometimes, imperfect knowledge also accounts for lags. At present the market for personal computers is glutted with all kinds of computers with varying features and prices. Moreover, since their introduction in the late 1970s, the prices of most personal computers have dropped dramatically. As a result, prospective consumers for the personal computer may hesitate to buy until they have Moreover, they may hesitate to buy in the expectation of further decline in price or innovations.

1.3.2.3. Institutional reasons. These reasons also contribute to lags. For example, contractual obligations may prevent firms from switching from one source of labor or raw material to another. As another example, those who have placed funds in long-term savings accounts for fixed durations such as 1 year, 3 years, or 7 years, are essentially “locked in’’ even though money market conditions may be such that higher yields are available elsewhere. Similarly, employers often give their employees a choice among several health insurance plans, but once a choice is made, an employee may not switch to another plan for at least 1 year. Although this may be done for administrative convenience, the employee is locked in for 1 year. For the reasons just discussed, lag occupies a central role in economics. This is clearly reflected in the short-run–long-run methodology of economics. It is for this reason we say that short-run price or income elasticities are generally smaller (in absolute value) than the corresponding long-run elasticities or that short-run marginal propensity to consume is generally smaller than long-run marginal propensity to consume.

1.3.3 AD HOC ESTIMATION OF DISTRIBUTION LAG MODEL:

Since the explanatory variable X_t is assumed to be nonstochastic (or at least uncorrelated with the disturbance u_t);

Y_{t-1}, X_{t-2} and so an, are non-stochastic

∴ The principle of OLS be applied to Where length of the lag is not defined, i.e. infinite lag model.

This approach is taken by Alt & Tinbergen.

They said that the regression should be sequence wise, fast regress Y_t and X_t then Y_t on X_t & X_{t-1}, then Y_t on Y_t, X_{t-1} & $X_{t-2}$and so on.

This will stop until lagged variable start becoming statistically insignificant or the coefficient of at least one of the variable changes its sign from positive to negative.

1.3.3.1 Drawback of ad HOC.

1. There is no a prior guide as to what is the maximum length of the lag.

2. Multicollinearity rears its ugly head.

1.4 SUMMARY AND CONCLUSIONS:

For psychological, technological, and institutional reasons, a regressand may respond to a regressor(s) with a time lag. Regression models that take into account time lags are known as dynamic or lagged regression models. There are two types of lagged models: distributed-lag and autoregressive. In the former, the current and lagged values of regressors are explanatory variables. In the latter, the lagged value(s) of the regressand appear as explanatory variables. A purely distributed-lag model can be estimated by OLS, but in that case there is the problem of multicollinearity since successive lagged values of a regressor tend to be correlated.

1.5 LETS SUM IT UP:

This lesson has surveyed a particular type of regression model, the dynamic regression. The signature feature of the dynamic model is effects that are delayed or that persist through time. In a static regression setting, effects embodied in coefficients are assumed to take place all at once. In the dynamic model, the response to an innovation is distributed through several periods. This lesson examined several different forms of single equation models that contained lagged effects.

1.6 EXCERCISES:

Q.1 What do you mean by dynamic models?

Q.2 Describe auto regressive model ?

Q.3 Describe the significance of lag in Economics ?

Q.4 Elaborate th various reasons for lags?

Q.5 Explain the AD HOC estimation of distribution lag model?

1.7 Suggested Reading / References:

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2nd Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

LESSON-2

THE KOYCK APPROACH TO DISTRIBUTED LAG MODELS

STRUCTURE

2.1 INTRODUCTION:

This lesson begins our introduction to the analysis of economic time series. By most views, this field has become synonymous with empirical macroeconomics and the analysis of financial markets.1 In this and the next lesson, we will consider a number of models and topics in which time and relationships through time play an explicit part in the formulation. Consider the **dynamic regression model**

$$y_t = \beta_1 + \beta_2 x_t + \beta_3 x_{t-1} + \gamma y_{t-1} + \varepsilon_t . \qquad \textbf{(1)}$$

Models of this form specifically include as right-hand side variables earlier as well as contemporaneous values of the regressors. It is also in this context that lagged values of the dependent variable appear as a consequence of the theoretical basis of the model rather than as a computational means of removing autocorrelation. There are several reasons why lagged effects might appear in an empirical model.

- In modeling the response of economic variables to policy stimuli, it is expected that there will be possibly long lags between policy changes and their impacts.The length of lag between changes in monetary policy and its impact on important economic variables such as output and investment has been a subject of analysis for several decades.

- Either the dependent variable or one of the independent variables is based on expectations. **Expectations** about economic events are usually formed by aggregating new information and past experience. Thus, we might write the expectation of a future value of variable *x*, formed this period, as

$$x_t = E_t [x_{t+1} | z_t , x_{t-1}, x_{t-2}, \ldots] = g(z_t , x_{t-1}, x_{t-2}, \ldots).$$

For example, forecasts of prices and income enter demand equations and consumption equations.

• Certain economic decisions are explicitly driven by a history of related activities. For example, energy demand by individuals is clearly a function not only of current prices and income, but also the accumulated stocks of energy using capital. Even energy demand in the macroeconomy behaves in this fashion—the stock of automobiles and its attendant demand for gasoline is clearly driven by past prices of gasoline and automobiles. Other classic examples are the dynamic relationship between investment decisions and past appropriation decisions and the consumption of addictive goods such as cigarettes and theater performances.

2.2 OBJECTIVES :

1. To understand the KOYCK approach to distributed lag models
2. To understand the rationalization of KOYCK approach
   A. The Adaptive Expectation Model (AEM)
   B. The Stock Adjustment or Partial Adj. Model
3. Understand the concept of Instrumental Variable

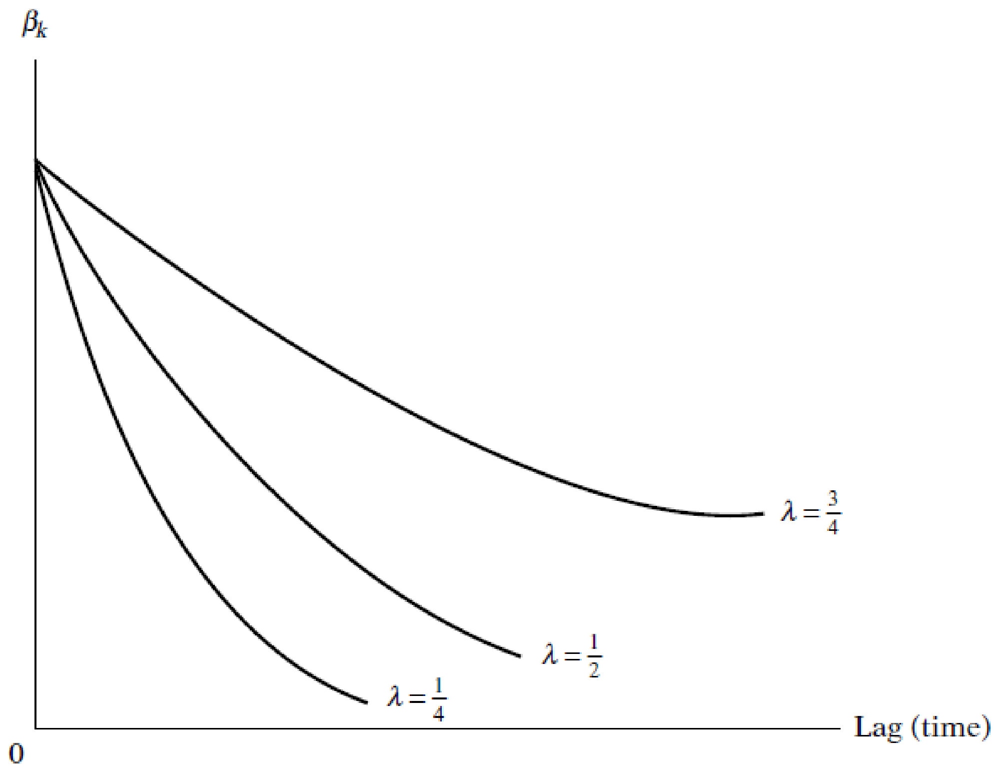
2.3 THE KOYCK APPROACH TO DISTRIBUTED LAG MODELS

Koyck has proposed an ingenious method of estimating distributed – lag models. Suppose we start with the infinite lag distributed- lag model … Assuming that the β's are all of the same sign, koyck assumes that they decline geometrically as follows.

$$\beta_k = \beta_0 \lambda^k \qquad k= 0, 1, \ldots \qquad (1)$$

Where λ such that $0 < \lambda < 1$, is known as the rate of decline or decay of the distributed lag and where $1 - \lambda$ is known as the speed of adjustment.

What postulates is that each successive β coefficient is numerically less than each preceding β, implying that as one goes back into the distant past, the effect of that lag on Y_t becomes progressively smaller, a quite plausible assumption. After all, current and recent past incomes are expected to affect current consumption expenditure more heavily than income in the distant past. Geometrically, the koyck scheme is depicted in figure.



Koyck scheme (declining geometric distribution).

As this figure shows, the value of the lag coefficient β_k depends, apart from the common β_0; on the value of λ .The closer λ is to 1 the slower the rate of decline in β_k, where as the closer it is to zero, the more rapid the decline in β_k in the former case, distant past values of X will exert sizable impact on Y_t whereas in the latter case their influence on Y_t Will peter out quickly. The pattern can be seen clearly from the following illustration

λ	β_0	β_1	β_2	β_3	β_4	β_5	...	β_{10}
0.75	β_0	$0.75\beta_0$	$0.56\beta_0$	$0.42\beta_0$	$0.32\beta_0$	$0.24\beta_0$	...	$0.06\beta_0$
0.25	β_0	$0.25\beta_0$	$0.06\beta_0$	$0.02\beta_0$	$0.004\beta_0$	$0.001\beta_0$	...	$0.0\beta_0$

Note these features of the koyck scheme: (1) By assuming nonnegative values for λ Koyck rules out the β'_s From changing sign; (2) by assuming $\lambda<1$ he gives lesser weight to the distant β'_s than the current ones; and (3) he ensures that the sum of the β'_s which gives the long-run multiplier, is finite; namely.

$$\sum_{k=0}^{\infty} \beta_k = \beta_0 \left(\frac{1}{1-\lambda} \right) \qquad (2)$$

As a result of the infinite lag model may be written as

$$Y_t = \alpha + \beta_0 X_t + \beta_0 \lambda X_{t-1} + \beta_0 \lambda X_{t-2} + \ldots\ldots\ldots\ldots\ldots + u_i \qquad (3)$$

As it stands, the model is still not amenable to easy estimation since a large (literally infinite) number of parameters remain to be estimated and the parameter λ Enters in a highly nonlinear form: Strictly speaking, the method of linear (in the parameters) regression analysis cannot be applied to such a model. But now Koyck suggests an ingenious way out, he lags (3.) by one period to obtain.

$$Y_{t-1} = \alpha + \beta_0 X_t + \beta_0 \lambda X_{t-2} + \beta_0 \lambda^2 X_{t-3} + \ldots\ldots\ldots\ldots u_{t-1} \qquad (4)$$

Multiplying equation (4) by λ

$$\lambda Y_{t-1} = \lambda\alpha + \lambda\beta_0 X_{t-1} + \beta_0 \lambda^2 X_{t-2} + \beta_0 \lambda^3 X_{t-3} + \cdots + \lambda u_{t-1} \quad ..5$$

Subtracting (5) from (3.), he gets

$$Y_t - \lambda Y_{t-1} = \alpha(1-\lambda) + \beta_0 X_t + (u_t - \lambda u_{t-1}) \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

or, rearranging

$Y_t = \alpha(1 - \lambda) + \beta_0 X_t + \lambda Y_{t-1} + v_t$ ……………..(7)

Where ($v_t = u_t - \lambda u_{t-1}$) a moving average of u_t and u_{t-1}.

The procedure just described is known as the Koyck transformation. Comparing (7.) with (1.) We see the tremendous simplification accomplished by Koyck. Whereas before we had to estimate α and an infinite number of β's, we now have to estimate only three unknowns α, β_0, and λ. Now there is no reason to expect multicollinearlity. In a sense multicollinearity is resolved by replacing X_{t-1}, X_{t-2}……….. by a single variable, namely Y_{t-1}. But note the following features of the Koyck transformation;

1. We started with a distributed- lag model but ended up with an autoregressive model because Y_{t-1} appears as one of the explanatory variables. This transformation shows how one can "convert" a distributed-lag model into an autoregressive model.

2. The appearance of Y_{t-1} is likely to create some statistical problems. Y_{t-1} like Y_t is stochastic, which means that we have a stochastic explanatory variable in the model. Recall that the classical least – squares theory is predicated on the assumption that the explanatory variables either are non-stochastic or, if stochastic, are distributed independently of the stochastic disturbance term. Hence we must find out if Y_{t-1} satisfied this assumption.

3. In the original model (1) the disturbance term was μ_t whereas in the transformed model it $v_t = (u_t - \lambda u_{t-1})$. The statistical properties of u_t Depend on what is assumed about the statistical properties of u_t. For, as shown later, if the original u_t's are serially uncorrelated, the v_t are serially correlated. Therefore, we may have to face up to the serial correlation problem in addition to the stochastic explanatory variable Y_{t-1}.

4. The presence of lagged Y violates one of the assumptions underlying the Durbin-Watson d test. Therefore we will have to develop an alternative to test for serial correlation in the presence of lagged Y. One alternative is the Durbin h test.

The partial sums of the standardized βi tell us the proportion of the long-run, or total, impact felt by a certain time period. In practice, though, the mean or median lag is often used to characterize the nature of the lag structure of a distributed lag model.

The Median Lag:

The median lag is the time required for the first half, or 50 percent, of the total change in *Y* following a unit sustained change in *X*. For the Koyck model, the median lag is as follows

$$\text{Koyck model: Median lag} = -\frac{\log 2}{\log \lambda}$$

Thus, if $\lambda = 0.2$ the median lag is 0.4306, but if $\lambda = 0.8$ the median lag is 3.1067. Verbally, in the former case 50 percent of the total change in *Y* is accomplished in less than half a period, whereas in the latter case it takes more than 3 periods to accomplish the 50 percent change. But this contrast should not be surprising, for as we know, the higher the value of λ the lower the speed of adjustment, and the lower the value of λ the greater the speed of adjustment.

The Mean Lag:

Provided all βk are positive, the mean, or average, lag is defined as

$$\text{Mean lag} = \frac{\sum_0^\infty k\beta_k}{\sum_0^\infty \beta_k}$$

which is simply the weighted average of all the lags involved, with the respective β coefficients serving as weights. In short, it is a lag-weighted average of time. For the Koyck model the mean lag is:

$$\text{Koyck model: Mean lag} = \frac{\lambda}{1-\lambda}$$

Thus, if $\lambda = 12$, the mean lag is 1.

From the preceding discussion it is clear that the median and mean lags serve as a summary measure of the speed with which *Y* responds to *X*. In the example given in Table 17.1 the mean lag is about 11 quarters, showing that it takes quite some time, on the average, for the effect of changes in the money supply to be felt on price changes.

2.4 RATIONALIZATION OF THE KOYCK MODEL

2.4.1 The Adaptive Expectation Model (AEM)

- AEM is a purely algebraic process.

Suppose we postulate the following Model.

$$Y_t = \alpha + \beta_0 + \beta_1 X_t^* + u_t \qquad \rightarrow (1)$$

e.g. (1) postulates that the element for money is a function of expected rate of interest.

Y = demand for money (real cash bal.)

X* = equilibrium optimum, normal rate of unit.

u = error term.

Now next eq. will hypothesis about how expectation are formed;

$$(X_t^* - X_{t-i}^*) = \gamma(X_t + X_{t-1}^*) \qquad \rightarrow (2)$$

$0 \neq \gamma \neq 1$ coeff. of expectation.

This hypothesis (2) is know as adaptive expectation, progressive expectation or error bearing, popularized by Cagan & Freidman.

So, eq.2 implies that, "economic agents will adapt their expectation in the light of past experience and that in particular they will Bearn from their mistakes.

$$X_t^* - X_{t-i}^* = \gamma X_t - \gamma X_{t-1}^*$$
$$X_t^* = \gamma X_t - \gamma X_{t-1}^* + X_{t-1}^* \rightarrow (3)$$
$$X_t^* = \gamma X_t - (1-\gamma) X_{t-1}^*$$

of γ=1 $X_t^* = X_t$ & expectation are realized immediately and fully in the same.

γ= 0, $X_t^* = X_{t-1}^*$ expectation are state.

Now put the value of (3) in (1) we get

$$Y_t = \beta_0 + \beta_1(\gamma X_t - (1-\gamma)\, X^*_{t-2}) + u_t$$

$$Y_t = \beta_0 + \beta_1\gamma X_t - \beta_1(1-\gamma)\, X^*_{t-1}) + u_t \rightarrow (4)$$

Taking lag in (1) we get.

$$Y_t = \beta_0 + \beta_1 X^*_{t-1} + u_{t-1}$$

$$Y_{t-1} - \beta_0 - u_{t-1} = \beta_1 X^*_{t-1}$$

Multiply by (1-γ) we get

$$Y_{t-1}(1-\gamma) - \beta_0(1-\gamma) + u_t(1-\gamma) = \beta_0(1-\gamma)\, X^*_{t-2} \rightarrow (5)$$

Put the value of (5) in (4) we get

$$Y_t = \beta_0 + \beta_1\gamma X_t + (1-\gamma)Y_{t-1} - \beta_0(1-\gamma) - U_1(1-\gamma) + U_t$$

$$Y_t = \cancel{\beta_0} + \beta_1\gamma X_t + (1-\gamma)Y_{t-1} - \beta_0 + \cancel{\beta_0}\gamma - \underbrace{U1(1-\gamma) + Ut}_{V_t}$$

$$Y_t = \beta_0\gamma + \beta_1\gamma X_t + (1-\gamma)Y_{t-1} + V_t \rightarrow (6)$$

- β_1 measures the average response of Y to a unit change in X, the equal or long run value of X.
- Model is autoregressive and its error learn are similar are Koyck error term.

<u>2.4.2 The Stock Adjustment or Partial Adjustment Model</u>:-

Marc Nerlove gave this model in economics

(The way to adjustment the stock)

$$Y^*_t = \beta_0 + \beta_1 X_t + U_t \rightarrow (1)$$

Since the desired level of capital is not directly observable, Nerlove postulated this hypothesis, known as the partial adjustment.

$$(Y_t\ Y_{t-1}) = \delta(Y^*_t - Y_{t-1}) \rightarrow (2)$$

$0 \le \delta \le 1$= coeff. of adjustment….

$(Y_t\ Y_{t-1})$ = actual change.

$$(Y^*_t - Y_{t-1}) = \text{desired change}$$

$$Y_t = \delta Y^*_t - \delta Y_{t-1} + Y_{t-1}$$

$$Y_t = \delta Y^*_t - (\delta\text{-}1)Y_{t-1}$$

$$\delta Y^*_t = Y_t - (\delta\text{-}1)Y_{t-1} \quad \rightarrow (3)$$

Multiply eq (1) by δ we get

$$\delta Y^*_t = \delta\beta_0 + \delta\beta_1 X_t + \delta\mu_t \quad \rightarrow (4)$$

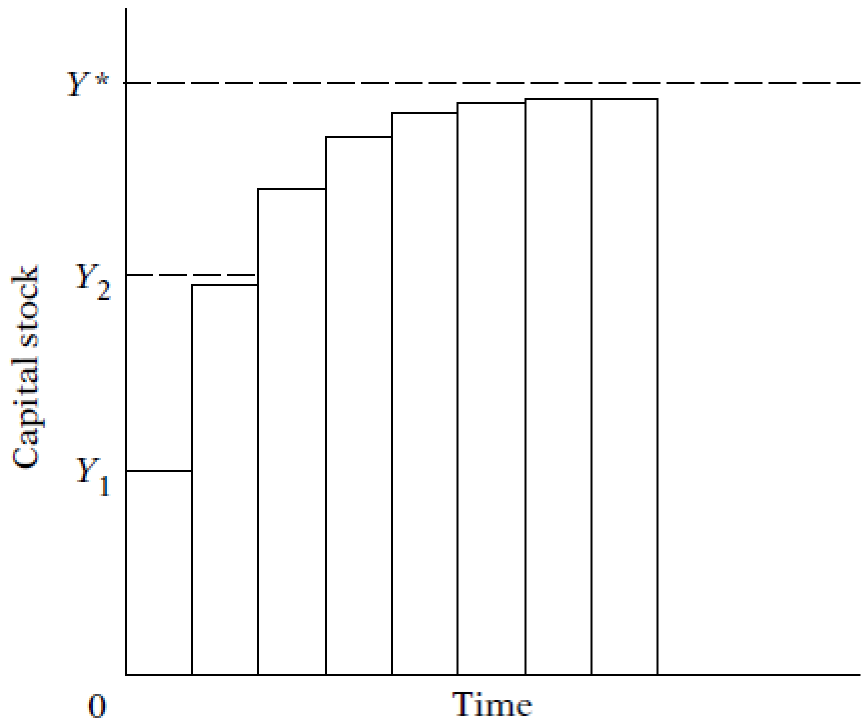
Put the value of (3) of in (4)

$$Y_t - (\delta\text{-}1)Y_{t-1} = \delta\beta_0 + \delta\beta_1 X_t + \delta u_t$$

$$Y_t = \delta\beta_0 + \delta\beta_1 X_t + (\delta\text{-}1)Y_{t-1} + \delta u_t \quad \rightarrow (t)$$

- If $\delta = 1$ it means the actual stock of capital = to the desired stock.
- $\delta = 0$, means nothing change since actual stock at time t.
- δ is expected to be between these extremes sine adjustment to the desired stock of capital is likely to be incomplete because of rigidity etc. hence the name partial adjustment model.
- Eq (1) represents the long run demand for capital stock.
- Eq (4) represents the short run demand function for capital stock.

Partial Adjustment Model



The gradual adjustment of capital stock

Y^* = desired capital stock.

Y_1 = current actual capital stock.

- ➢ The firm plans to close half the gap between the actual & the desired stock of capital each period.
- ➢ Ist period t moves to Y_2, within involved $(Y_2\text{-}Y_1) = ½(Y^*\text{-}Y_1)$.
- ➢ Each subsequent period it closes half the gap between the capital stock at the beginning of the period & the desired capital stock Y^*.

<u>SOME IMPORTANT POINTS:-</u>

- Both the models are different from one & another.
- Models in autoregressive
- AEM is based on uncertainty. (about the course of prices)
- PAM is due to test or institutional rigidities, cost of change etc.

2.5 INSTRUMENTAL VARIABLE:

Under Koyck or AEM is the explanatory variable Y_{t-1} teach to be correlated with the error term V_t, this is the reason why we can't apply OLS to this model.

If the correlation will be removed then the OLS can be applied.

A proxy for Y_{t-1} that is highly correlated with Y_{t-1} but is uncorrelated with V_t, where V_t is the error term appearing in the Koyck or AEM, such the proxy is called Instruments variable.

Liviatan suggests X_{t-1} as the Instrumental variable for Y_{t-1}.

Koyck = $Y_t = \alpha(1-\lambda) + \beta_0 X_t + \lambda Y_t + (U_1 - \lambda U_{t-1})$

ARM = $Y_t = \gamma\beta_0 + \gamma\beta_1 X_t + (1-\gamma)Y_{t-1} + [U_1 - (1-\gamma)U_{t-1}]$

PAM= $Y_t = \delta\beta_0 + \delta\beta_1 X_t + (1-\delta)Y_{t-1} + \delta U_1$

All these models have the common form

$Y_t = \alpha_0 + \alpha_1 X_t + \alpha_2 Y_{t-1} + V_t \qquad \rightarrow(A)$

Now we will obtain the normal equation of (A)

1)
$$\begin{aligned}
\sum Y_t &= n\hat{\alpha}_0 + \hat{\alpha}_1 \sum X_t + \hat{\alpha}_2 \sum Y_{1-t} \\
\sum Y_t X_t &= \hat{\alpha}_0 \sum X_t + \hat{\alpha}_1 \sum X_t^2 + \hat{\alpha}_2 \sum Y_{t-1} X_t \\
\sum Y_t X_t &= \hat{\alpha}_0 \sum X_{t-1} + \hat{\alpha}_1 \sum X_t X_{t-1} + \hat{\alpha}_2 \sum Y_{t-1} X_{t-1}
\end{aligned}$$

But if we will apply OLS directly to eq. (A) normal eq will be

2)
$$\begin{aligned}
\sum Y_t &= n\hat{\alpha}_0 + \hat{\alpha}_1 \sum X_t + \hat{\alpha}_2 \sum Y_{1-t} \\
\sum Y_t X_t &= \hat{\alpha}_0 \sum X_t + \hat{\alpha}_1 \sum X_t^2 + \hat{\alpha}_2 \sum Y_{t-1} X_t \\
\sum Y_t X_{t-1} &= \hat{\alpha}_0 \sum Y_{t-1} + \hat{\alpha}_1 \sum X_t Y_{t-1} + \hat{\alpha}_2 \sum Y_{t-1}^2
\end{aligned}$$

Difference between these two normal equal is in eq (1) α's estimated are consistent where as estimation for (2) may not be consistent because Y_{t-1} and X_1V_t may be correlated where as X_t and X_{t-1} are uncorrelated with v_t.

Dennis Sargan has to developed a test dubbed the SARG Test to find out the validity of IV SARG = (n-k)R^2.

no. of observe   no. of coeff. in original regression.

2.6 SUMMARY AND CONCLUSIONS:

A purely distributed-lag model can be estimated by OLS, but in thatcase there is the problem of multicollinearity since successive lagged values of a regressor tend to be correlated. As a result, some shortcut methods have been devised. These include the Koyck, the adaptive expectations, and partial adjustment mechanisms, the first being a purely algebraic approach and the other two being based on economic principles. But a unique feature of the **Koyck, adaptive expectations,** and **partial adjustment models** is that they all are autoregressive in nature in that the lagged value(s) of the regressand appear as one of the explanatory variables. Autoregressiveness poses estimation challenges; if the lagged regressand is correlated with the error term, OLS estimators of such models are not only biased but also are inconsistent. Bias and inconsistency are the case with the Koyck and the adaptive expectations models; the partial adjustment model is different in that it can be consistently estimated by OLS despite the presence of the lagged regressand. To estimate the Koyck and adaptive expectations models consistently, the most popular method is the **method of instrumental variable.** The instrumental variable is a proxy variable for the lagged regressand but with the property that it is uncorrelated with the error term.

2.7 LETS SUM IT UP:

In the end we can say that Koyck, adaptive expectations, and partial adjustment models are all autoregressive in nature and in all these models the lagged values of the regressand appear as one of the explanatory variables.

2.8 EXCERCISES:

Q1. Discuss partial adjustment model

Q2. Explain koyck approach to estimate distributed lag model.

Q3. Describe adaptive expectations model.

Q4. Elaborate the method of instrumental variable.

Q5. Distinguish between Mean Lag and Median Lag?

2.9 Suggested Reading / References:

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2nd Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

LESSON-3

THE ALMON APPROACH TO DISTRIBUTION LAG MODEL

STRUCTURE

3.1 INTRODUCTION:

An alternative to the lagged regression models just discussed is the **Almon polynomial distributed-lag model,** which avoids the estimation problems associated with the autoregressive models. The major problem with the Almon approach, however, is that one must *prespecify* both the lag length and the degree of the polynomial. There are both formal and informal methods of resolving the choice of the lag length and the degree of the polynomial.

3.2 OBJECTIVES:

1. Understand the lagged/ autoregressive models.

2. Undestand the Almon Approach to distribution lag model.

3. Undestand the Almon Approach through grphic representation.

3.3 THE ALMON APPROACH TO DISTRIBUTION LAG MODEL:

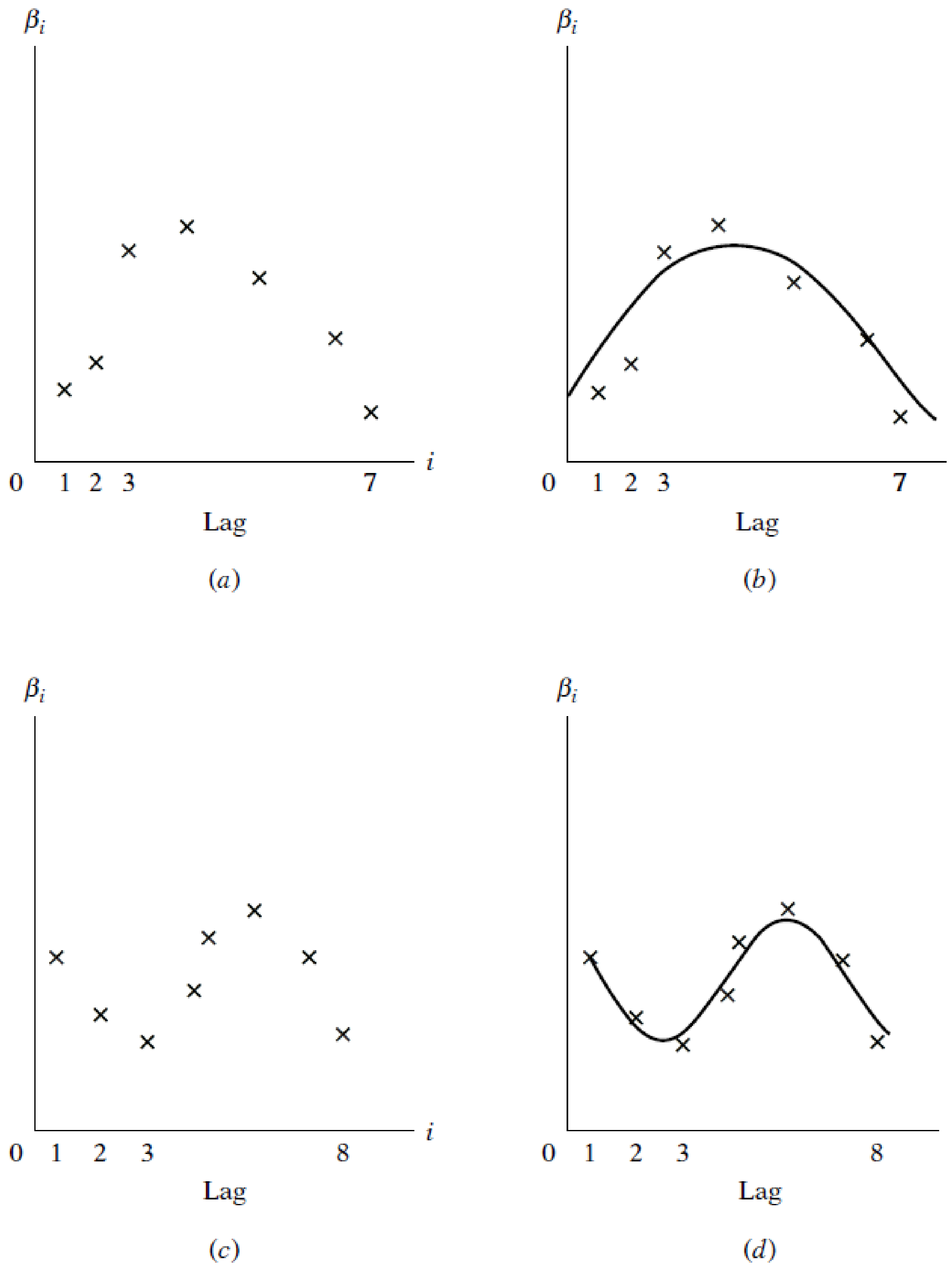
Although used extensively in practice, the Koyck distribution-lag model is based on the assumption that the β coefficients. This approach is precisely the one suggestion by Shirley Almon. To illustrate her technique, let us revert to the finite distribution-lag model. β'_0 increasing at first and then decreasing in 'C' β follow a cyclical pattern.

In the figure it is assumed that the *β*'s increase at first and then decrease, whereas in figure it is assumed that they follow a cyclical pattern. Obviously, the Koyck scheme of distributed-lag models will not work in these cases. However, after looking at Figure 17.7*a* and *c*, it seems that one can express βi as a function of *i*, the length of the lag (time), and fit suitable curves to reflect the functional relationship between the two, as indicated in Figure 17.7*b* and *d*. This approach is precisely the one suggested by Shirley Almon. To illustrate her technique, let us revert to the finite distributed-lag model considered previously, namely,

$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \cdots + \beta_k X_{t-k} + u_t$

which may be written more compactly as

$Y_t = \alpha + \sum_{i=0}^{k} \beta_i X_{t-i} + u_i$



Almon polynomial-lag scheme.

Figure.

And hence the equation follows:

$$Yt = \alpha + \sum_{i=0}^{k} \beta_i X_{t-i} + u_i$$

Following a theorem in mathematics known as Weierstrass' theorem, Almon assumes that βi can be approximated by a suitable-degree polynomial in *i*, the length of the lag. For instance, if the lag scheme shown in

Figure a, applies, we can write

$$\beta_i = a_0 + a_1 i + a_2 i^2 \qquad (2)$$

which is a quadratic, or second-degree, polynomial in *i* (see Figure b).

However, if the β's follow the pattern of Figure *c*, we can write

$$\beta_i = a_0 + a_1 i + a_2 i^2 + a_3 i^3 \qquad (3)$$

which is a third-degree polynomial in *i* (see Figure *d*). More generally, we may write

$$\beta_i = a_0 + a_1 i + a_2 i^2 + \cdot\cdot\cdot + a_m i^m \qquad (4)$$

which is an *m*th-degree polynomial in *i*. It is assumed that *m* (the degree of the polynomial) is less than *k* (the maximum length of the lag).

To explain how the Almon scheme works, let us assume that the β's follow the pattern shown in Figure a and, therefore, the second-degree polynomial approximation is appropriate. Substituting (2) into (1), we obtain

$$Y_1 = \alpha \sum_{i=0}^{k} (a_0 + a_1 i + a_2 i^2) X_{t-i} + u_1$$

(5)

$$= \alpha + a_0 \sum_{i=0}^{k} X_{t-i} + a_1 \sum_{i=0}^{k} iX_{t-i} + a_2 \sum_{i=0}^{k} i^2 X_{t-i} + u_t$$

Defining

$$Z_{0t} \sum_{i=0}^{k} X_{t-i}$$

$$Z_{1t} \sum_{i=0}^{k} iX_{t-i}$$

$$Z_{2t} \sum_{i=0}^{k} t^2 X_{t-i}$$

(7)

we may write (.5) as

$Y_t = \alpha + a_0 Z_0 t + a_1 Z_{1t} + a_2 Z_{2t} + u_t$ **(8)**

In the Almon scheme *Y* is regressed on the constructed variables *Z*, not the original *X* variables. Note that (7) can be estimated by the usual OLS procedure. The estimates of α and *ai* thus obtained will have all the desirable statistical properties provided the stochastic disturbance term *u* satisfies the assumptions of the classical linear regression model. In this respect, the Almon technique has a distinct advantage over the Koyck method because, as we have seen, the latter has some serious estimation problems that result from the presence of the stochastic explanatory variable *Yt*−1 and X its likely correlation with the disturbance term.

3.4 SUMMARY AND CONCLUSIONS:

Despite the estimation problems, which can be surmounted, the distributed and autoregressive models have proved extremely useful in empirical economics because they make the otherwise static economic theory a dynamic one by taking into account explicitly the role of time. Such models help us to distinguish between short- and long-run response of the dependent variable to a unit change in the value of the explanatory variable(s). Thus, for estimating short- and long-run price, income, substitution, and other elasticities these models have proved to be highly useful.

3.5 LETS SUM IT UP:

This lesson has surveyed a particular type of regression model, the dynamic regression. In the dynamic model, the response to an innovation is distributed through several periods. The progression, which mirrors the current literature is from tightly structured lag "models" (which were sometimes formulated to respond to a shortage of data rather than to correspond to an underlying theory) to unrestricted models with multiple period lag structures. **Almon polynomial distributed-lag model,** which avoids the estimation problems associated with the autoregressive models. The major problem with the Almon approach, however, is that one must *prespecify* both the lag length and the degree of the polynomial.

EXCERCISES:

Q1. Show how to estimate a polynomial distributed lag model with lags of six periods and a third-order polynomial.

Q2. Expand the rational lag model $y_t = [(0.6 + 2L)/(1 - 0.6L + 0.5L^2)]x_t + e_t$. What are the coefficients on x_t , x_{t-1}, x_{t-2}, x_{t-3}, and x_{t-4}?

Q3. Whenever the lagged dependent variable appears as an explanatory variable, the R^2 is usually much higher than when it is not included. What are the reasons for this observation?

Q4. Consider the lag patterns in Figure 1.5 What degree polynomials would you fit to the lag structures and why?

Q5. Consider the following distributed-lag model:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \beta_3 X_{t-3} + \beta_4 X_{t-4} + u_t$$

Assume that β_i can be adequately expressed by the second-degree polynomial as follows:

$$\beta_i = a_0 + a_1 i + a_2 i^2$$

How would you estimate the β's if we want to impose the restriction that $\beta_0 = \beta_4 = 0$?

Q6. Explain the Almon Approach to the distributed lag model?

3.7 Suggested Reading / References:

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2nd Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

LESSON-4

COINTEGRATION AND ERROR CORRECTION MECHANISM (ECM) AND THE CAUSALITY IN ECONOMICS THE GRANGER CAUSALITY TEST

STRUCTURE

4.1 INTRODUCTION:

Consider the Auto Regressive Distributed Model, which has become a workhorse of the modern literature on time-series analysis. By defining the first differences $\_y_t = y_t - y_{t-1}$ and $\_x_t = x_t - x_{t-1}$ we can rearrange

$y_t = \mu + \gamma_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$ to obtain

$$\_y_t = \mu + \beta_0 \_x_t + (\gamma_1 - 1)(y_{t-1} - \theta x_{t-1}) + \varepsilon_t , \qquad \textbf{(1)}$$

where $\theta = -(\beta_0 + \beta_1)/(\gamma_1 - 1)$. This form of the model is in the **error correction** form. In this form, we have an **equilibrium relationship,** $\_y_t = \mu + \beta_0 \_x_t + \varepsilon_t$, and the **equilibrium error,** $(\gamma_1 - 1)(y_{t-1} - \theta x_{t-1})$, which account for the deviation of the pair of variables from that equilibrium. The model states that the change in y_t from the previous period consists of the change associated with movement with x_t along the long-run equilibrium path plus a part $(\gamma_1 - 1)$ of the deviation $(y_{t-1} - \theta x_{t-1})$ from the equilibrium.With a model in logs, this relationship would be in proportional terms.

4.2 OBJECTIVES:

1. Understand the concept of cointegration and error correction mechanism.

2. Understand the Granger Causality Test.

3. Understand the relationship between causality and exogeneity.

4.3 COINTEGRATION AND ERROR CORRECTION MECHANISM (ECM)

We have warned that the regression of a non-stationary time series on another non-stationary time series may produce a spurious regression. Suppose, then, that we regress PCE on PDI as follows:

Let us write this as:

$$\text{PCE}_t = \beta_1 + \beta_2\text{PDI}_t + u_t$$

$$u_t = \text{PCE}_t - \beta_1 - \beta_2\text{PDI}_t$$

PCE=Personal Disposable Income

PCC=Personal Consumption Expenditure

Suppose we now subject *ut* to unit root analysis and find that it is stationary; that is, it is $I(0)$. This is an interesting situation, for although PCE*t* and PDI*t* are individually $I(1)$, that is, they have stochastic trends, their linear combination is eq. 2 $I(0)$. So to speak, the linear combination cancels out the stochastic trends in the two series. If you take consumption and income as two $I(1)$ variables, savings defined as (income − consumption) could be $I(0)$. As a result, a regression of consumption on income as in.1) would be meaningful (i.e., not spurious). In this case we say that the two variables are cointegrated. Economically speaking, two variables will be co-integrated if they have a long-term, or equilibrium, relationship between them. Economic theory is often expressed in equilibrium terms, such as Fisher's quantity theory of money or the theory of purchasing parity (PPP), just to name a few.

We just showed that PCE and PDI are co-integrated; that is, there is a long term, or equilibrium, relationship between the two. Of course, in the short run there may be disequilibrium.

The error correction mechanism (ECM) first used by Sargan and later popularized by Engle and Granger corrects for disequilibrium. An important theorem, known as the Granger representation theorem, states that if two variables *Y* and *X* are cointegrated, then the relationship between the two can be expressed as ECM. Now consider the following model:

$$\Delta\text{PCE}_t = \alpha_0 + \alpha_1\Delta\text{PDI}_t + \alpha_2 u_{t-1} + \varepsilon_t \qquad \textbf{(3)}$$

where Δ as usual denotes the first difference operator, εt is a random error term, and u_{t-1} = $(\text{PCE}_{t-1} - \beta_1 - \beta_2\text{PDI}_{t-1})$, that is, the one-period lagged value of the error.

ECM equation (3) states that ΔPCE depends on ΔPDI and also on the equilibrium error term. If the latter is nonzero, then the model is out of equilibrium. Suppose ΔPDI is zero and u_{t-1} is positive. This means PCE_{t-1} is too high to be in equilibrium, that is, PCE_{t-1} is above its equilibrium value of (α_0 + $\alpha_1 PDI_{t-1}$). Since $\alpha 2$ is expected to be negative, the term $\alpha_2 u_{t-1}$ is negative and, therefore, ΔPCE_t will be negative to restore the equilibrium. That is, if PCE_t is above its equilibrium value, it will start falling in the next period to correct the equilibrium error; hence the name ECM. By the same token, if u_{t-1} is negative (i.e., PCE is below its equilibrium value), $\alpha_2 u_{t-1}$ will be positive, which will cause ΔCPE_t to be positive, leading PCE_t to rise in period t.

Thus, the absolute value of α_2 decides how quickly the equilibrium is restored. In practice, we estimate u_{t-1} by $\hat{u}_{t-1}$ = (PCE$t - \hat{\beta}_1 - \hat{\beta}_2$ PDI$_t$).

Statistically, the equilibrium error term is zero, suggesting that PCE adjust to changes in PDI in the same time relation.

4.4 THE GRANGER CAUSALITY TEST:

To explain the Granger test, we will see the GDP and many effect upon each other.

Is GDP that "Causes" the money supply M(GDP$\rightarrow$M).

or money supply M caused GDP (M$\rightarrow$GDP).

($\rightarrow$ cause)

Test involves the following of regression

$$GDP_t = \sum_{i=1}^{n} \lambda_i M_{t-i} + \sum_{i=1}^{n} \delta_i GDP_{t-i} + u_{2t} \qquad \rightarrow (1)$$

$$M_t = \sum_{i=1}^{n} \lambda_i M_{t-i} + \sum_{i=1}^{n} \delta_i GDP_{t-i} + u_{2t} \qquad \rightarrow (2)$$

- $u_t = u_{2t}$ are uncorrelated.
- Multivariable causality through the technique of vector autoregressive (VAR)

4.4.1 Now we will discuss 4 cases.

1. Unidirectional causally from M to GDP is indicated if the estimated coefficients on the lagged M in (1) are statistically different from zero in (2).

2. Conversely, unidirectional causality from GDP to M exists if the set of lagged M. Coefficient in (1) is not statically different from zero & the set of lagged GDP coefficient (2).

3. Feedback, or bilateral causality, when set of M & GDP coefficients due. Statically significant diff from zero in both regression.

4. Independent is suggest when the set of M & GDP coefficients due statically significant in both the regression.

Cox = MS $\rightarrow$GDP

GDP $\rightarrow$MS

GDP $\rightleftharpoons$ MS

Step involved in implementing the granger causality list are a under

1. RSS_R:- Residual sum of Square (restricted) i.e. regress current GDP on all lagged. GDP term & other variable, don't include M variable in the regressive.

2. Include lagged M terms.

   RSS_{UR}= Unrestructed

3. The Null hypothesis is H_0 $\Sigma\alpha i = 0$ i.e. lagged M terms do not belong in the regression.

4. Apply F Test

   $$F= \frac{(RSS_R - RSS_{UR})/m}{RSS_{UR}/(n-k)}$$

   m = no. of lagged M

   k = no. of parameter estimated in the unrestricted region

5. If F value exceeds the critical F value then M cause GDP.

6. If step 1 to 5 repacked to test the model i.e GDP cause M.

4.4.2 Assumptions

1. GDP and M are stationary.

2. No. of lagged terms to be introduced in the causality list is an imp-que.

3. Error are uncorrelated.

4. Vector Autoregressive.

4.5 A NOTE ON CAUSALITY AND EXOGENEITY

Economic variables are often classified into two broad categories, endogenous and exogenous. Loosely speaking, endogenous variables are the equivalent of the dependent variable in the single-equation regression model and exogenous variables are the equivalent of the *X* variables, or regressors, in such a model, provided the *X* variables are uncorrelated with the error term in that equation.

Now we raise an interesting question: Suppose in a Granger causality test we find that an *X* variable (Granger) causes a *Y* variable without being caused by the latter (i.e., no bilateral causality). Can we then treat the *X* variable as exogenous? In other words, can we use Granger causality (or noncausality) to establish exogeneity?

To answer this question, we need to distinguish three types of exogeneity: (1) weak, (2) strong, and (3) super. To keep the exposition simple, suppose we consider only two variables, Y_t and X_t, and further suppose we regress Y_t on X_t. We say that X_t is *weakly exogenous* if Y_t also does not explain X_t. In this case estimation and testing of the regression model can be done, conditional on the values of X_t. As a matter of fact, going back to previous Lesson , you will realize that our regression modeling was conditional on the values of the *X* variables. X_t is said to be *strongly* exogenous if current and lagged *Y* values do not explain it (i.e., no feedback relationship). And *Xt* is *superexogenous* if the parameters in the regression of *Y* and *X* do not change even if the *X* values change; that is, the parameter values are invariant to changes in the value(s) of *X*. If that is in fact the case, then, the famous "Lucas critique" may lose its force.

The reason for distinguishing the three types of exogeneity is that, "In general, weak exogeneity is all that is needed for estimating and testing, strong exogeneity is necessary for forecasting and super exogeneity for policy analysis."

Returning to Granger causality, if a variable, say *Y*, does not cause another variable, say *X*, can we then assume that the latter is exogenous? Unfortunately, the answer is not straightforward. If we are talking about weak exogeneity, it can be shown that *Granger causality is neither necessary nor sufficient to establish exogeneity*. On the other hand, Granger causality is necessary (but not sufficient) for strong exogeneity. The proofs of these statements are beyond the scope of this book. For our purpose, then, it is better to keep the concepts of Granger causality and exogeneity separate and treat the former as a useful descriptive tool for time series data. In subsequent Lesson we will discuss a test to find out if a variable can be treated as exogenous.

4.6 SUMMARY AND CONCLUSIONS

Despite the estimation problems, which can be surmounted, the distributed and autoregressive models have proved extremely useful in empirical economics because they make the otherwise static economic theory a dynamic one by taking into account explicitly the role of time. Such models help us to distinguish between short- and long-run response of the dependent variable to a unit change in the value of the explanatory variable(s). Thus, for estimating short- and long-run price, income, substitution, and other elasticities these models have proved to be highly useful.

Because of the lags involved, distributed and or autoregressive models raise the topic of causality in economic variables. In applied work, the Granger causality modeling has received considerable attention. But one has to exercise great caution in using the Granger methodology because it is very sensitive to the lag length used in the model.

Even if a variable (X) "Granger-causes" another variable (Y), it does not mean that X is exogenous. We distinguished three types of exogeneity weak, strong, and super and pointed out the importance of the distinction.

4.7 LETS SUM IT UP:

In last we conclude that , distributed and or autoregressive models raise the topic of causality in economic variables and causality is all about finding the variable which is affecting the dependent variable. The causality can be both ways one way or none at all.

4.8 EXCERCISES:

Q1.describe cointegration and error correction mechanism (ecm)?

Q2. What do you mean by equilibrium error?

Q3. Explain the Granger Causality test?

Q.4 Write a note on Causality and Exogeneity?

4.9 Suggested Reading / References:

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2^{nd} Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

•

UNIT- 2

SIMULTANEOUS EQUATION MODELS

LESSON-1

SIMULTANEOUS EQUATION. MODELS

STRUCTURE

1.1 INTRODUCTION:

In this lesson we are concerned exclusively with single equation models, i.e., models in which there was a single dependent variable Y and one or more explanatory variables, the X's. In such models the emphasis was on estimating and/or predicting the average value of Y conditional upon the fixed values of the X variables. The cause-and-effect relationship, if any, in such models therefore ran from the X's to the Y. But in many situations, such a one-way or unidirectional cause-and-effect relationship is not meaningful. This occurs if Y is determined by the X's, and some of the X's are, in turn, determined by Y. In short, there is a two way, or simultaneous, relationship between Y and (some of) the X's, which makes the distinction between dependent and explanatory variables of dubious value. It is better to lump together a set of variables that can be determined simultaneously by the remaining set of variables—precisely what is done in simultaneous-equation models. In such models there is more than one equation—one for each of the mutually, or jointly, dependent or endogenous variables.1 And unlike the single-equation models, in the simultaneous-equation models one may not estimate the parameters of a single equation without taking into account information provided by other equations in the system.

What happens if the parameters of each equation are estimated by applying, say, the method of OLS, disregarding other equations in the system? Recall that one of the crucial assumptions of the method of OLS is that the explanatory *X* variables are either nonstochastic or, if stochastic (random), are distributed independently of the stochastic disturbance term. If neither of these conditions is met, then, as shown later, the least-squares estimators are not only biased but also inconsistent; that is, as the sample size increases indefinitely, the estimators do not converge to their true (population) values. Thus, in the following hypothetical system of equations,

$$Y_{1i} = \beta_{10} + \beta_{12}Y_{2i} + \gamma_{11}X_{1i} + u_{1i} \qquad \textbf{(1)}$$

$$Y2i = \beta 20 + \beta 21Y1i + \gamma 21X1i + u2i \qquad \textbf{(2)}$$

where $Y1$ and $Y2$ are mutually dependent, or endogenous, variables and $X1$ is an exogenous variable and where $u1$ and $u2$ are the stochastic disturbance terms, the variables $Y1$ and $Y2$ are both stochastic. Therefore, unless it can be shown that the stochastic explanatory variable $Y2$ in (1) is distributed independently of $u1$ and the stochastic explanatory variable $Y1$ in (2) is distributed independently of $u2$, application of the classical OLS to these equations individually will lead to inconsistent estimates.

1.2 OBJECTIVES:

1. Understand the concept of simultaneous equations.

2. Understand the structural form, Reduced form, and Recursive form of simultaneous equations.

3.The another objective is to find the inconsistency of OLS estimators

1.3 SIMULTANEOUS EQUATIONS.:-

Simultaneous equations are the equation when there is two way relation i.e. Y is determined by X's and since X's are determined by Y.

In these models there is more than 2 equation one of each of the mutually, dependent or endogenous variable

ex. $Y_{1t} = \beta_{10} + \beta_{12}Y_{2i} + Y_{1i} X_{1i} + U_{1i}$

$Y_{2t} = \beta_{20} + \beta_{21}Y_{1i} + Y_{2i} X_{1i} + U_{2i}$

Where Y_1 & Y_2 are mutually dependent or endogenous, and X_1 is exogenous & U_2 & U_2 stochastic disturbance terms, the variable Y_1 and Y_2 are stochastic.

1.3.1 <u>STRUCTURAL MODEL</u>:- It is a complete system of equation which describe the structure of the relationships of the economic variable it express the endogenous variable as functions of others endogenous variables.

Structural systems

$Y_1 = \beta_{13} Y_3 + U_1$

$Y_2 = \beta_{23} Y_3 + \gamma_{21} X_1 + U_2$

$Y_3 = Y_1 + Y_2 + Y_2$

1.3.2 <u>REDUCED FORM MODEL</u>:- The reduced for of a structural model is the model in which the endogenous variable are expressed as a function of the predetermined variable.

These are two way to get reduced form a express the endogenous variable directly as function of predetermined variable obtaining the reduced form of model is to solve the structural system of endogenous variable in terms of the predetermined variable the structural parameters & the disturbance.

1.3.3 <u>RECURSIVE MODEL</u>:- A model is called recursive of its structural eq. can be ordered in such a very that the first includes only predetermined variable in the rights hand side, the seconds eq. contains predetermined variable and the first endogenous variable (of the first eq.) in the eight hand side an so on.

Cx - $Y_1 = f(X_1 X_2 \ldots\ldots\ldots\ldots\ldots\ldots X_k, U_1)$

$Y_2 = f(X_1 X_2 \ldots\ldots\ldots\ldots\ldots\ldots X, Y_1, U_2)$

$Y_3 = f(X_1 X_2 \ldots\ldots\ldots\ldots\ldots\ldots X, Y_1, Y_2, U_3)$

and so on

Also called triangular system coz the coefficient of the end (β's) for a triangular ray.

1.4 The Simultaneous Equation Bias:- Inconsistency of OLS Estimation

If one or more of the explanatory variable are correlated with the disturbance term in that equation because the estimator thus obtained are inconsistent.

Its prove that Y_t & U_t are correlated following procedure is:-

$C_t = \beta_0 + \beta_1 Y_t + U_t \qquad 0<\beta, <1 \qquad (1)$

(Consumption function)

$Y_t = C_t + I_t\ (\div S_t) \qquad (2)$

Income indentification

Subtracting (1) in (2) we get

$Y_t = \beta_0 + \beta_1 Y_t + U_t + I_t \qquad (3)$

$Y_t - \beta_1 Y_t = \beta_0 + U_t + I_t$

$Y_t\ (1-\beta_1) = \beta_0 + U_t + I_t \qquad (4)$

$Y_t = \frac{\beta_0}{(1-\beta_1)} + \frac{1}{(1-\beta_1)} I_t + \frac{1}{(1-\beta_1)} U_t \qquad (5)$

$E(Y_t) = \frac{\beta_0}{(1-\beta_1)} + \frac{1}{(1-\beta_1)} I_t \qquad (6)$

$E(U_t) = 0$

It = exogenous

Now subtracting (6) from (5) we get

Y_t - E(Y_t)= $\frac{\beta_0}{(1-\beta_1)} + \frac{1}{(1-\beta_1)} I_t + \frac{1}{(1-\beta_1)} U_t$ -

$$\frac{\beta_0}{(1-\beta_1)} - \frac{1}{(1-\beta_1)} I_t$$

Yt - E(Y_t) = $\frac{U_t}{(1-\beta_1)}$ (7)

Now

U_t - E(U_t) = U_t (8)

cov (Y_t, U_t) = E [Y_t -E(U_t)) (U_t - E(U_t)]

= $E\frac{(U_t)}{(1-\beta_1)}$ U_t from (7) & (8)

= $E\frac{(u_t^2)}{(1-\beta_1)}$

Cov (Y_t, u_t) = $\frac{\sigma^2}{(1-\beta_1)}$ (9)

1.4.1 ASSUMPTION:

1. σ^2 is positive, cov. between Y & U_t is bound to be different from zero.

2. OLS estimators in this situation are inconsistent.

Now OLS estimator $\hat{\beta}_1$ is an inconsistent estimator of β, bceause of the correlation between Y_t & u_t.

$\hat{\beta}_1$ = $\frac{\Sigma(C_t-\bar{C})(Y_t-\bar{Y})}{\Sigma(Y_t+\bar{Y})^2}$

= $\frac{\Sigma c_t y_t}{\Sigma y_t^2}$ (10)

$= \frac{\Sigma\, C_t Y_t}{\Sigma y_t^2}$ (Small letter we mean deviation of mean value)

Subtracting C_t from (1)

$\hat{\beta}_1 = \frac{\Sigma(\beta_0 + \beta_1 Y_t + U_t) y_t}{\Sigma y_t^2}$

$= \beta_1 \frac{\Sigma y_t + u_t}{\Sigma y_t^2}$ (11) $\Sigma Y_t = 0, \frac{\Sigma Y_t + U_t}{\Sigma y_t^2} = 1$

$E(\hat{\beta}_1) = \beta_1 + E\left\{\frac{\Sigma y_t + u_t}{\Sigma y_t^2}\right\}$ (12)

We can't evaluate $E\left\{\frac{\Sigma y_t + u_t}{\Sigma y_t^2}\right\}$

Since the expectation operation is a linear operator.

$\hat{\beta}_1$ is a biased estimator of β_1

Applying plan (pro ability limit) to (11)

$\text{plim}\,(\hat{\beta}_1) = \text{plim}\,(\beta_1) + \text{plim}\left\{\frac{\Sigma y_t + u_t}{\Sigma y_t^2}\right\}$

$= \text{plim}\,(\beta_1) + \text{plim}\,\frac{\Sigma y_t + u_t/n}{\Sigma y_t^2/n}$ (13)

n= Total no. of observations

We have about by n, now the sample covariance between Y & U and sample variance of Y.

If n increasing indefinitely, cov between Y & U to approximate the true population cov. by which eq (a) is equal to $(\sigma^2/(1 - \beta_1)$.

If n tends to infinity, the sample variance of Y will apporx. its population variance say σ^2_y.

$\text{Plim}\,(\hat{\beta}_1) = \beta_1 + \frac{\sigma^2/(1-\beta_1)}{\sigma_y^2}$

$$= \beta_1 + \frac{1}{(1-\beta_1)} \frac{\sigma^2}{\sigma_y^2} \quad (14)$$

- $0 < \beta_1 < 1$
- σ^2 & σ_y^2 both are +ve
- plim ($\hat{\beta}_1$) > β_1
- $\hat{\beta}_1$ is biased estimate and the bias will not disappear no matter how large the sample size is.
-

1.5 SUMMARY AND CONCLUSIONS:

In contrast to single-equation models, in simultaneous-equation models more than one dependent, or endogenous, variable is involved, necessitating as many equations as the number of endogenous variables. A unique feature of simultaneous-equation models is that the endogenous variable (i.e., regressand) in one equation may appear as an explanatory

variable (i.e., regressor) in another equation of the system. As a consequence, such an endogenous explanatory variable becomes stochastic and is usually correlated with the disturbance term of the equation in which it appears as an explanatory variable. In this situation the classical OLS method may not be applied because the estimators thus obtained are not consistent, that is, they do not converge to their true population values no matter how large the sample size.

1.6 LETS SUM IT UP:

Although most of our work thus far has been in the context of single-equation models, even a cursory look through almost any economics textbook shows that much of the theory is built on sets, or *systems*, of relationships. Familiar examples include market equilibrium, models of the macroeconomy, and sets of factor or commodity demand equations. Whether one's interest is only in a particular part of the system or in the system as a whole, the interaction of the variables in the model will have important implications for both interpretation and estimation of the model's parameters.

1.7 EXCERCISES:

Q1. Develop a simultaneous-equation model for the supply of and demand for dentists in the United States. Specify the endogenous and exogenous variables in the model.

Q2. To study the relationship between inflation and yield on common stock,Bruno Oudet‡ used the following model:

$R_{bt} = \alpha_1 + \alpha_2 R_{st} + \alpha_3 R_{bt-1} + \alpha_4 L_t + \alpha_5 Y_t + \alpha_6 \text{NIS}_t + \alpha_7 I_t + u_{1t}$

$R_{st} = \beta_1 + \beta_2 R_{bt} + \beta_3 R_{bt-1} + \beta_4 L_t + \beta_5 Y_t + \beta_6 \text{NIS}_t + \beta_7 E_t + u_{2t}$

where L = real per capita monetary base

Y = real per capita income

I = the expected rate of inflation

NIS = a new issue variable

E = expected end-of-period stock returns, proxied by lagged stock price ratios

R_{bt}= bond yield

R_{st}= common stock returns

a. Offer a theoretical justification for this model and see if your reasoning agrees with that of Oudet.

b. Which are the endogenous variables in the model? And the exogenous variables?

c. How would you treat the lagged R_{bt}—endogenous or exogenous?

Q3. G. Menges developed the following econometric model for the West German economy*:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 I_t + u_{1t}$$

$$I_t = \beta_3 + \beta_4 Y_t + \beta_5 Q_t + u_{2t}$$

$$C_t = \beta_6 + \beta_7 Y_t + \beta_8 C_{t-1} + \beta_9 P_t + u_{3t}$$

$$Q_t = \beta_{10} + \beta_{11} Q_{t-1} + \beta_{12} R_t + u_{4t}$$

Where

Y = national income

I = net capital formation

C = personal consumption

Q = profits

P = cost of living index

R = industrial productivity

t = time

u = stochastic disturbances

a. Which of the variables would you regard as endogenous and which as exogenous?

b. Is there any equation in the system that can be estimated by the single-equation least-squares method?

c. What is the reason behind including the variable P in the consumption function?

Q4. Describe the various forms of simultaneous equations?

Q5. Explain the simiultaneous equation bias?

1.8 Suggested Reading / References:

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2nd Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

LESSON-2
THE IDENTIFICATION PROBLEM

STRUCTURE

2.1 INTRODUCTION:

In this lesson we consider the nature and significance of the identification problem. The crux of the identification problem is as follows: Recall the demand-and-supply model. Suppose that we have time series data on Q and P only and no additional information (such as income of the consumer, price prevailing in the previous period, and weather condition). The identification problem then consists in seeking an answer to this question: Given only the data on P and Q, how do we know whether we are estimating the demand function or the supply function? Alternatively, if we *think* we are fitting a demand function, how do we guarantee that it is, in fact, the demand function that we are estimating and not something else?

A moment's reflection will reveal that an answer to the preceding question is necessary before one proceeds to estimate the parameters of our demand function. In this lesson we shall show how the identification problem is resolved. We first introduce a few notations and definitions and then illustrate the identification problem with several examples. This is followed by the rules that may be used to find out whether an equation in a simultaneous-equation model is identified, that is, whether it is the relationship that we are actually estimating, be it the demand or supply function or something else.

2.2 OBJECTIVES:

1. Our objective is to identify the parameters of a structural equation.

2. To identify the underidentified equations.

3. To identify the exactly identified equations.

4. To identify the overidentified equations.

2.3 IDENTIFICATION PROBLEM:

By the **identification problem** we mean whether numerical estimates of the parameters of a structural equation can be obtained from the estimated reduced-form coefficients. If this can be done, we say that the particular equation is *identified.* If this cannot be done, then we say that the equation under consideration is *unidentified,* or *underidentified.* An identified equation may be either *exactly* (or fully or just) *identified* or *overidentified*. It is said to be exactly identified if unique numerical values of the structural parameters can be obtained. It is said to be overidentified if more than one numerical value can be obtained for some of the parameters of the structural equations.

The identification problem arises because different sets of structural coefficients may be compatible with the same set of data. To put the matter differently, a given reduced-form equation may be compatible with different structural equations or different hypotheses (models), and it may be difficult to tell which particular hypothesis (model) we are investigating. In the remainder of this section we consider several examples to show the nature of the identification problem.

2.3.1 UNDER IDENTIFICATION:

Consider once again the demand-and-supply model together with the market-clearing, or equilibrium, condition that demand is equal to supply. By the equilibrium condition, we obtain

$$\alpha_0 + \alpha_1 P_t + u_{1t} = \beta_0 + \beta_1 P_t + u_{2t} \qquad \textbf{(1)}$$

Solving (19.2.1), we obtain the equilibrium price

$$P_t = \prod_0 + v_t \qquad \textbf{(2)}$$

where

$$\prod_0 = \frac{\beta_0 - \alpha_0}{\alpha_{0_1} - \beta_1} \qquad \textbf{(3)}$$

$$v_t = \frac{u_{2t} - u_{1t}}{\alpha_1 - \beta_1} \qquad \textbf{(4)}$$

Incidentally, note that the error terms *vt* and *wt* are linear combinations of the original error terms *u*1 and *u*2. Equations (2) and (5) are reduced-form equations. Now our demand-and-supply model contains four structural coefficients $\alpha 0$, $\alpha 1$, $\beta 0$, and $\beta 1$, but there is no unique way of estimating them. Why? The answer lies in the two reduced-form coefficients given in (3) and (6). These reduced-form coefficients contain all four structural parameters, but there is no way in which the four structural unknowns can be estimated from only two reduced-form coefficients. Recall from high school algebra that to estimate four unknowns we must have four (independent) equations, and, in general, to estimate *k* unknowns we must have *k* (independent) equations. Incidentally, if we run the reduced-form regression (2) and (5), we will see that there are no explanatory variables, only the *constants*, and these *constants* will simply give the mean values of *P* and *Q* (why?). What all this means is that, given time series data on *P* (price) and *Q* (quantity) and no other information, there is no way the researcher can guarantee whether he or she is estimating the demand function or the supply function. That is, a given *Pt* and *Qt* represent simply the point of intersection of the appropriate demand-and-supply curves because of the equilibrium condition that demand is equal to supply.

For an equation to be identified, that is, for its parameters to be estimated, it must be shown that the given set of data will not produce a structural equation that looks similar in appearance to the one in which we are interested. If we set out to estimate the demand function, we must show that the given data are not consistent with the supply function or some mongrel equation.

2.3.1.1 UNDER IDENTIFICATION`

$$Q_t = \alpha_0 + \alpha_1 P_t + u_{1t} \qquad \textbf{(1)}$$

$$Q_t = \beta_0 + \beta_1 P_t + u_{2t} \qquad \textbf{(2)}$$

Equation (1) as demand function and equation (2) as supply function

$$\alpha_0 + \alpha_1 P_t + u_{1t} = \beta_0 + \beta_1 P_t + u_{2t} \qquad \textbf{(3)}$$

$$\alpha_1 P_t - \beta_1 P_t = \beta_0 - \alpha_0 - u_{1t} + u_{2t} \quad \textbf{(4)}$$

$$(\alpha_1 - \beta_t) P_t = \beta_0 - \alpha_0 - u_{1t} + u_{2t} \quad \textbf{(5)}$$

$$P_t = \frac{\beta_0}{\alpha_1 - \beta_1} - \frac{\alpha_0}{\alpha_1 - \beta_1} - \frac{U_{1t} + U_{2t}}{\alpha_0 - \beta_0} \quad \textbf{(6)}$$

$$P_t = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \frac{U_{2t} + U_{1t}}{\alpha_0 - \beta_0} \quad \textbf{(7)}$$

$$P_t = \pi_0 + V_t \quad \textbf{(8)}$$

$$\pi_0 = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1}$$

$$V_t = \frac{U_{2t} + U_{1t}}{\alpha_1 - \beta_1}$$

2.3.2 JUST, OR EXACT, IDENTIFICATION:

The reason we could not identify the preceding demand function or the supply function was that the same variables *P* and *Q* are present in both functions and there is no additional information, such as that indicated in .But suppose we consider the following demand-and-supply model:

$$\textit{Demand function:} \quad Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + u_{1t} \qquad \alpha_1 < 0, \alpha_2 > 0$$

$$\textit{Supply function:} \quad Q_t = \beta_0 + \beta_1 P_t + u_{2t} \qquad \beta_1 > 0$$

(Demand function as eq. 1 and supply function as eq. 2)

where I = income of the consumer, an exogenous variable, and all other variables are as defined previously.

Notice that the only difference between the preceding model and our original demand-and-supply model is that there is an additional variable in the demand function, namely, income. From economic theory of demand we know that income is usually an important determinant of demand for most goods and services. Therefore, its inclusion in the demand function will give us some additional information about consumer behavior. For most commodities income is expected to have a positive effect on consumption ($\alpha 2 > 0$).

Using the market-clearing mechanism, quantity demanded = quantity supplied, we have

$\alpha_0 + \alpha_1 P_t + \alpha_2 I_t + u_{1t} = \beta_0 + \beta_1 P_t + u_{2t}$ **(3)**

Solving Eq. (3) provides the following equilibrium value of Pt :

$P_t = \Pi_0 + \Pi_1 I_t + v_t$ **(4)**

where the reduced-form coefficients are

$$\Pi_0 = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1}$$

$$\Pi_1 = -\frac{\alpha_2}{\alpha_1 - \beta_1}$$

and

$$v_t = \frac{u_{2t} - u_{1t}}{\alpha_1 - \beta_1} \qquad (5)$$

Substituting the equilibrium value of *point* into the preceding demand or supply function, we obtain the following equilibrium quantity:

$$Q_t = \Pi_2 + \Pi_3 I_t + w_t$$

where

$$\Pi_2 = \frac{\alpha_1 \beta_0 - \alpha_0 \beta_1}{\alpha_1 - \beta_1}$$

$$\Pi_3 = -\frac{\alpha_2 \beta_1}{\alpha_1 - \beta_1}$$

and

$$w_t = \frac{\alpha_1 u_{2t} - \beta_1 u_{1t}}{\alpha_1 - \beta_1}$$

Since (4) and (6) are both reduced-form equations, the OLS method can be applied to estimate their parameters. Now the demand-and supply model (1) and (2) contains five structural coefficients— α_0, α_1, α_2, β_1, and β_2. But there are only four equations to estimate them, namely, the four reduced-form coefficients $\prod_0$, $\prod_1$, $\prod_2$, and $\prod_3$ given in (5) and (7).

Hence, unique solution of all the structural coefficients is not possible. But it can be readily shown that the parameters of the supply function can be identified (estimated) because

$$\beta_0 = \Pi_2 - \beta_1 \Pi_0$$

$$\beta_1 = \frac{\Pi_3}{\Pi_1}$$

The demand-and-supply model given in Eqs and contain six structural coefficients—α0, α1, α2, β0, β1, and β2—and there are six reduced form coefficients to estimate them. Thus, we have six equations in six unknowns, and normally we should be able to obtain unique estimates. Therefore, the parameters of both the demand and supply equations can be identified, and the system as a whole can be identified. To check that the preceding demand-and-supply functions are identified, we can also resort to the device of multiplying the demand equation by λ ($0 \leq \lambda \leq 1$) and the supply equation by $1 - \lambda$ and add them to obtain a mongrel equation. This mongrel equation will contain both the predetermined variables It and Pt−1; hence, it will be observationally different from the demand as well as the supply equation because the former does not contain Pt−1 and the latter does not contain it.

2.3.3 OVERIDENTIFICATION:

For certain goods and services, income as well as wealth of the consumer is an important determinant of demand. Therefore, let us modify the demand function (1 previous) as follows, keeping the supply function as before:

Demand function: $Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + \alpha_3 R_t + u_{1t}$ **(1)**

Supply function: $Q_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + u_{2t}$ **(2)**

where in addition to the variables already defined, R represents wealth; for most goods and services, wealth, like income, is expected to have a positive effect on consumption.

Equating demand to supply, we obtain the following equilibrium price and quantity:

$$P_t = \Pi_0 + \Pi_1 I_t + \Pi_2 R_t + \Pi_3 P_{t-1} + v_t$$

$$Q_t = \Pi_4 + \Pi_5 I_t + \Pi_6 R_t + \Pi_7 P_{t-1} + w_t$$

where

$$\Pi_0 = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} \qquad \Pi_1 = -\frac{\alpha_2}{\alpha_1 - \beta_1}$$

$$\Pi_2 = -\frac{\alpha_3}{\alpha_1 - \beta_1} \qquad \Pi_3 = \frac{\beta_2}{\alpha_1 - \beta_1}$$

$$\Pi_4 = \frac{\alpha_1\beta_0 - \alpha_0\beta_1}{\alpha_1 - \beta_1} \qquad \Pi_5 = -\frac{\alpha_2\beta_1}{\alpha_1 - \beta_1}$$

$$\Pi_6 = -\frac{\alpha_3\beta_1}{\alpha_1 - \beta_1} \qquad \Pi_7 = \frac{\alpha_1\beta_2}{\alpha_1 - \beta_1}$$

$$w_t = \frac{\alpha_1 u_{2t} - \beta_1 u_{1t}}{\alpha_1 - \beta_1} \qquad v_t = \frac{u_{2t} - u_{1t}}{\alpha_1 - \beta_1}$$

Above are eq. 3,4, & 5

The preceding demand-and-supply model contains seven structural coefficients, but there are eight equations to estimate them—the eight reducedform coefficients given in (5);

that is, the number of equations is greater than the number of unknowns. As a result, unique estimation of all the parameters of our model is not possible, which can be shown easily. From the preceding reduced-form coefficients, we can obtain

$$\beta_1 = \frac{\Pi_6}{\Pi_2}$$

or

$$\beta_1 = \frac{\Pi_5}{\Pi_1}$$

that is, there are two estimates of the price coefficient in the supply function, and there is no guarantee that these two values or solutions will be identical. Moreover, since β_1 appears in the denominators of all the reduced-form coefficients, the ambiguity in the estimation of β_1 will be transmitted to other estimates too.

2.4 SUMMARY AND CONCLUSIONS:

The problem of identification precedes the problem of estimation.

The identification problem asks whether one can obtain unique numerical estimates of the structural coefficients from the estimated reducedform coefficients.If this can be done, an equation in a system of simultaneous equations is identified. If this cannot be done, that equation is un- or underidentified. An identified equation can be just identified or overidentified. In the former case, unique values of structural coefficients can be obtained; in the latter, there may be more than one value for one or more structural parameters. The identification problem arises because the same set of data may be compatible with different sets of structural coefficients, that is, different models. Thus, in the regression of price on quantity only, it is difficult to tell whether one is estimating the supply function or the demand function, because price and quantity enter both equations. To assess the identifiability of a structural equation, one may apply the technique of **reduced-form equations,** which expresses an endogenous variable solely as a function of predetermined variables.

2.5 LETS SUM IT UP:

The models surveyed in this lesson involve most of the issues that arise in analysis of linear equations in econometrics. Before one embarks on the process of estimation, it is necessary to establish that the sample data actually contain sufficient information to provide estimates of the parameters in question. This is the question of identification. Identification involves both the statistical properties of estimators and the role of theory in the specification of the model. Once identification is established, there are numerous methods of estimation.

2.6 EXCERCISES:

Q 1. Consider the following two-equation model:

$$y1 = \gamma 1 y2 + \beta 11 x1 + \beta 21 x2 + \beta 31 x3 + \varepsilon 1,$$

$$y2 = \gamma 2 y1 + \beta 12 x1 + \beta 22 x2 + \beta 32 x3 + \varepsilon 2.$$

a. Verify that, as stated, neither equation is identified.

b. Establish whether or not the following restrictions are sufficient to identify (or partially identify) the model:

(1) $\beta 21 = \beta 32 = 0$,

(2) $\beta 12 = \beta 22 = 0$,

(3) $\gamma 1 = 0$,

(4) $\gamma 1 = \gamma 2$ and $\beta 32 = 0$,

(5) $\sigma 12 = 0$ and $\beta 31 = 0$,

(6) $\gamma 1 = 0$ and $\sigma 12 = 0$,

(7) $\beta 21 + \beta 22 = 1$,

(8) $\sigma 12 = 0$, $\beta 21 = \beta 22 = \beta 31 = \beta 32 = 0$,

(9) $\sigma 12 = 0$, $\beta 11 = \beta 21 = \beta 22 = \beta 31 = \beta 32 = 0$.

Q.2 For the model

$y_1 = \gamma_1 y_2 + \beta_{11} x_1 + \beta_{21} x_2 + \varepsilon_1,$

$y_2 = \gamma_2 y_1 + \beta_{32} x_3 + \beta_{42} x_4 + \varepsilon_2,$

show that there are two restrictions on the reduced-form coefficients. Describe a procedure for estimating the model while incorporating the restrictions.

Q3. The model

$Y_{1t} = \beta_{10} + \beta_{12} Y_{2t} + \gamma_{11} X_{1t} + u_{1t}$

$Y_{2t} = \beta_{20} + \beta_{21} Y_{1t} + u_{2t}$

produces the following reduced-form equations:

$Y_{1t} = 4 + 8X_{1t}$

$Y_{2t} = 2 + 12X_{1t}$

a. Which structural coefficients, if any, can be estimated from the reduced-form coefficients? Demonstrate your contention.

b. How does the answer to (a) change if it is known a priori that

(1) $\beta_{12} = 0$ and (2) $\beta_{10} = 0$?

Q4. Expalin underidentified, overidentified and exactly identified problems?

Q5. Give some examples of reduced form equations?

2.7 Suggested Reading / References:

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2nd Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

LESSON 3

RULES FOR IDENTIFICATION AND A TEST OF SIMULTANEITY

STRUCTURE

3.1 INTRODUCTION:

As the examples shown previous in Section, in principle it is possible to resort to the reduced-form equations to determine the identification of an equation in a system of simultaneous equations. But these examples also show how time-consuming and laborious the process can be. Fortunately, it is not essential to use this procedure. The so-called order and rank conditions of identification lighten the task by providing a systematic routine. To understand the order and rank conditions, we introduce the following notations:

M = number of endogenous variables in the model

m = number of endogenous variables in a given equation

K = number of predetermined variables in the model including the intercept

k = number of predetermined variables in a given equation

K - k = m-1

3.2 OBJECTIVES:

1. Understand the order condition of identifiability.

2. Understand the rank condition of identifiability.

3. Understand the test of simultaneity.

3.3 THE ORDER CONDITION OF IDENTIFIABILITY :

A necessary (but not sufficient) condition of identification, known as the order condition, may be stated in two but equivalent ways as follows:

Definition 1:

In a model of M simultaneous equations, in order for an equation to be identified, it must exclude at least M-1 variables (endogenous as well as predetermined)

appearing in the model. If it excludes exactly M-1 variables, the equation is just identified. If it excludes more than M-1 variables, it is over identified.

Definition 2:

In a model of M simultaneous equations, in order for an equation to be identified, the number of predetermined variables excluded from the equation must not be less than the number of endogenous variables included in that equation less 1, that is,

$$K - k \geq m - 1$$

If $K - k = m - 1$, the equation is just identified,

But $K - k > m - 1$ it is over identified.

To illustrate the order condition, let us take examples.

EXAMPLE 1

Demand function: $Q_t = \alpha_0 + \alpha_1 P_t + u_{1t}$ 1

Supply function: $Q_t = \beta_0 + \beta_1 P_t + u_{2t}$ 2

This model has two endogenous variables P and Q and no predetermined variables. To be identified, each of these equations must exclude at least $M - 1 = 1$ variable. Since this is not the case, neither equation is identified.

EXAMPLE 2

Demand function: $Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + u_{1t}$ 3

Supply function: $Q_t = \beta_0 + \beta_1 P_t + u_{2t}$ 4

In this model Q and P are endogenous and I is exogenous. Applying the order condition given in 3, we see that the demand function is unidentified. On the other hand, the supply function is just identified because it excludes exactly $M - 1 = 1$ variable I_t.

EXAMPLE 3

Demand function: $Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + u_{1t}$ 5

Supply function: $Q_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + u_{2t}$ 6

Given that P_t and Q_t are endogenous and I_t and P_{t-1} are predetermined, Eq. 5 excludes exactly one variable P_{t-1} and Eq. 6 also excludes exactly one variable I_t. Hence each equation is identified by the order condition. Therefore, the model as a whole is identified.

EXAMPLE .4

$$\text{Demand function:} \quad Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + \alpha_3 R_t + u_{1t} \qquad 1$$

$$\text{Supply function:} \quad Q_t = \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + u_{2t} \qquad 2$$

In this model P_t and Q_t are endogenous and I_t, R_t, and P_{t-1} are predetermined. The demand function excludes exactly one variable P_{t-1}, and hence by the order condition it is exactly identified. But the supply function excludes two variables I_t and R_t, and hence it is overidentified. As noted before, in this case there are two ways of estimating β_1, the coefficient of the price variable.

Notice a slight complication here. By the order condition the demand function is identified. But if we try to estimate the parameters of this equation from the reduced-form coefficients given in 1 , the estimates will not be unique because β_1, which enters into the computations, takes two values and we shall have to decide which of these values is appropriate.

3.4 THE RANK CONDITION OF IDENTIFIABILITY:

The order condition discussed previously is *a necessary but not sufficient* condition for identification; that is, even if it is satisfied, it may happen that an equation is not identified. Thus, in Example, the supply equation was identified by the order condition because it excluded the income variable I_t, which appeared in the demand function. But identification is accomplished only if $\alpha 2$, the coefficient of I_t in the demand function, is not zero, that is, if the income variable not only probably but actually does enter the demand function.

More generally, even if the order condition $K - k \geq m - 1$ is satisfied by an equation, it may be unidentified because the predetermined variables excluded from this equation but present in the model may not all be independent so that there may not be one-to-one correspondence between the structural coefficients (the β's) and the reduced-form coefficients.

That is, we may not be able to estimate the structural parameters from the reduced-form coefficients, as we shall show shortly. Therefore, we need both a necessary and sufficient condition for identification. This is provided by the rank condition of identification, which may be stated as follows:

> RANK CONDITION OF IDENTIFICATION
>
> In a model containing M equations in M endogenous variables, an equation is identified if and only if *at least* one nonzero determinant of order $(M - 1)(M - 1)$ can be constructed from the coefficients of the variables (both endogenous and predetermined) excluded from that particular equation but included in the other equations of the model.

As an illustration of the rank condition of identification, consider the following hypothetical system of simultaneous equations in which the *Y* variables are endogenous and the *X* variables are predetermined.

$$
\begin{array}{lr}
Y_{1t} - \beta_{10} \qquad - \beta_{12}Y_{2t} - \beta_{13}Y_{3t} - \gamma_{11}X_{1t} = u_{1t} & 1 \\
Y_{2t} - \beta_{20} \qquad - \beta_{23}Y_{3t} - \gamma_{21}X_{1t} - \gamma_{22}X_{2t} = u_{2t} & 2 \\
Y_{3t} - \beta_{30} - \beta_{31}Y_{1t} \qquad - \gamma_{31}X_{1t} - \gamma_{32}X_{2t} = u_{3t} & 3 \\
Y_{4t} - \beta_{40} - \beta_{41}Y_{1t} - \beta_{42}Y_{2t} \qquad - \gamma_{43}X_{3t} = u_{4t} & 4
\end{array}
$$

To facilitate identification, let us write the preceding system in Table , which is self-explanatory. Let us first apply the order condition of identification. By the order condition each equation is identified. Let us recheck with the rank condition. Consider the first equation, which excludes variables $Y4$, $X2$, and $X3$ (this is represented by zeros in the first row). For this equation to be identified, we must obtain at least one nonzero determinant of order 3 × 3 from the coefficients of the variables excluded from this equation but included in other equations. To obtain the determinant we first obtain the relevant matrix of coefficients of variables $Y4$, $X2$, and $X3$ included in the other equations. In the present case there is only one such matrix, call it A, defined as follows:

$$
\mathbf{A} = \begin{bmatrix} 0 & -\gamma_{22} & 0 \\ 0 & -\gamma_{32} & 0 \\ 1 & 0 & -\gamma_{43} \end{bmatrix} \qquad 6
$$

It can be seen that the determinant of this matrix is zero:

Since the determinant is zero, the rank of the matrix (6), denoted by $\rho(\mathbf{A})$, is less than 3. Therefore, Eq. (2) does not satisfy the rank condition and hence is not identified. As noted, the rank condition is both a necessary and sufficient condition for identification. Therefore, although the order condition shows that Eq. (2) is identified, the rank condition shows that it is not. Apparently, the columns or rows of the matrix **A** given in (6) are not (linearly) independent, meaning that there is some relationship between the variables $Y4$, $X2$, and $X3$. As a result, we may not have enough information to estimate the parameters of equation (2); the reduced-form equations for the preceding model will show that it is not possible to obtain the structural coefficients of that equation from the reduced-form coefficients. The reader should verify that by the rank condition Eqs. (3) and (4) are also unidentified but Eq. (5) is identified. As the preceding discussion shows, *the rank condition tells us whether the equation under consideration is identified or not, whereas the order condition tells us if it is exactly identified or overidentified.*

To apply the rank condition one may proceed as follows:

1. Write down the system in a tabular form, as shown in table of coefficients of variables.

2. Strike out the coefficients of the row in which the equation under consideration appears.

3. Also strike out the columns corresponding to those coefficients in 2 which are nonzero.

4. The entries left in the table will then give only the coefficients of the variables included in the system but not in the equation under consideration. From these entries form all possible matrices, like **A**, of order $M-1$ and obtain the corresponding determinants. If at least one non-vanishing or nonzero determinant can be found, the equation in question is (just or over)

identified. The rank of the matrix, say, **A**, in this case is exactly equal to $M-1$. If all the possible $(M-1)(M-1)$ determinants are zero, the rank of the matrix **A** is less than $M-1$ and the equation under investigation is not identified.

Our discussion of the order and rank conditions of identification leads to the following general principles of identifiability of a structural equation in a system of M simultaneous equations:

1. If $K - k > m - 1$ and the rank of the **A** matrix is $M - 1$, the equation is overidentified.
2. If $K - k = m - 1$ and the rank of the matrix **A** is $M - 1$, the equation is exactly identified.
3. If $K - k \geq m - 1$ and the rank of the matrix **A** is less than $M - 1$, the equation is underidentified.
4. If $K - k < m - 1$, the structural equation is unidentified. The rank of the **A** matrix in this

case is bound to be less than $M - 1$.

Henceforth, when we talk about identification we mean exact identification, or overidentification. There is no point in considering unidentified, or underidentified, equations because no matter how extensive the data, the structural parameters cannot be estimated. However, as shown parameters of overidentified as well as just identified equations can be estimated. Which condition should one use in practice: Order or rank? For large simultaneous-equation models, applying the rank condition is a formidable task.

Therefore, as Harvey notes,

Fortunately, the order condition is usually sufficient to ensure identifiability, and although it is important to be aware of the rank condition, a failure to verify it will rarely result in disaster.

3.5 A TEST OF SIMULTANEITY:

If there is no simultaneous equation, or simultaneity problem, the OLS estimators produce consistent and efficient estimators. On the other hand, if there is simultaneity, OLS estimators are not even consistent. In the presence of simultaneity, as we will show in Lesson 20, the methods of two stage least squares (2SLS) and instrumental variables will give estimators that are consistent and efficient. Oddly, if we apply these alternative methods when there is in fact no simultaneity, these methods yield estimators that are consistent but not efficient (i.e., with smaller variance). All this discussion suggests that we should check for the simultaneity problem before we discard OLS in favor of the alternatives.

As we showed earlier, the simultaneity problem arises because some of the regressors are endogenous and are therefore likely to be correlated with the disturbance, or error, term.Therefore, a test of simultaneity is essentially a test of whether (an endogenous) regressor is correlated with the error term. If it is, the simultaneity problem exists, in which case alternatives to OLS must be found; if it is not, we can use OLS. To find out which is the case in a concrete situation, we can use Hausman's specification error test.

3.5.1 Hausman Specification Test:

A version of the Hausman specification error test that can be used for testing the simultaneity problem can be explained as follows:

To fix ideas, consider the following two-equation model:

Demand function:

$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + \alpha_3 R_t + u_{1t}$ (1)

Supply function:

$Q_t = \beta_0 + \beta_1 P_t + u_{2t}$ (2)

Where: P = price

Q = quantity

I = income

R = wealth

u's = error terms

Assume that I and R are exogenous. Of course, P and Q are endogenous.

Now consider the supply function (2). If there is no simultaneity problem (i.e., P and Q are mutually independent), P_t and u_{2t} should be un- correlated (why?). On the other hand, if there is simultaneity, P_t and u_{2t} will be correlated. To find out which is the case, the Hausman test proceeds as follows:

First, from (1) and (2) we obtain the following reduced-form equations:

$P_t = {}_0 + {}_1 I_t + {}_2 R_t + v_t$ (3)

$$Q_t = {}_3 + {}_4 I_t + {}_3 R_t + w_t \qquad (4)$$

where v and w are the reduced-form error terms. Estimating (3) by OLS we obtain

$$\hat{P}_t = \hat{}_0 + \hat{}_1 I_t + \hat{}_2 R_t \qquad (5)$$

$$P_t = \hat{P}_t + \hat{v}_t \qquad (6)$$

Therefore,

where $\hat{P}_t$ are estimated P_t and $\hat{v}_t$ are the estimated residuals. Substituting (6) into (2), we get

$$Q_t = \beta_0 + \beta_1 \hat{P}_t + \beta_1 \hat{v}_t + u_{2t} \qquad (7)$$

Note: The coefficients of P_t and v_t are the same.

Now, under the null hypothesis that there is no simultaneity, the correlation between $\hat{v}_t$ and u_{2t} should be zero, asymptotically. Thus, if we run the regression (7) and find that the coefficient of v_t in (7) is statistcally zero, we can conclude that there is no simultaneity problem. Of course, this conclusion will be reversed if we find this coefficient to be statistically significant.

Essentially, then, the Hausman test involves the following steps:

Step 1. Regress P_t on I_t and R_t to obtain $\hat{v}_t$.

Step 2. Regress Q_t on $\hat{P}_t$ and $\hat{v}_t$ and perform a t test on the coefficient of $\hat{v}_t$. If it is significant, do not reject the hypothesis of simultaneity; otherwise, reject it. For efficient estimation, however, Pindyck and Rubinfeld suggest regressing Q_t on P_t and $\hat{v}_t$.

3.6 SUMMARY AND CONCLUSIONS:

However, this time-consuming procedure can be avoided by resorting to either the **order condition** or the **rank condition** of identification.

Although the order condition is easy to apply, it provides only a necessary condition for identification. On the other hand, the rank condition is both a necessary and sufficient condition for identification. If the rank condition is satisfied, the order condition is satisfied, too, although the converse is not true. In practice, though, the order condition is generally adequate to ensure identifiability. In the presence of simultaneity, OLS is generally not applicable. But if one wants to use it nonetheless, it is imperative to test for simultaneity explicitly. The **Hausman specification test** can be used for this purpose.Although in practice deciding whether a variable is endogenous or exogenous is a matter of judgment, one can use the Hausman specification test to determine whether a variable or group of variables is endogenous or exogenous.

3.7 LETS SUM IT UP:

In last we can say that, Identification involves both the statistical properties of estimators and the role of theory in the specification of the model. Once identification is established, there are numerous methods of estimation. We considered a number of single equation techniques including least squares, instrumental variables, GMM, and maximum likelihood.

3.8 EXCERCISES:

Q1. Explain the Hausman specification test?

Q2. Describe the order condition of identifiability?

Q3. Describe the Rank condition of identifiability?

Q4. Consider the following extended Keynesian model of income determination:

Consumption function: $C_t = \beta_1 + \beta_2 Y_t - \beta_3 T_t + u_{1t}$

Investment function: $I_t = \alpha_0 + \alpha_1 Y_{t-1} + u_{2t}$

Taxation function: $T_t = \gamma_0 + \gamma_1 Y_t + u_{3t}$

Income identity: $Y_t = C_t + I_t + G_t$

where C = consumption expenditure

Y = income

I = investment

T = taxes

G = government expenditure

u's = the disturbance terms

In the model the endogenous variables are *C*, *I*, *T*, and *Y* and the predetermined variables are *G* and Y_{t-1} . By applying the order condition, check the identifiability of each of the equations in the system and of the system as a whole. What would happen if r_t , the interest rate, assumed to be exogenous, were to appear

on the right-hand side of the investment function?

3.9 Suggested Reading / References:

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2nd Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

LESSON-4

ESTIMATION OF A JUST IDENTIFIED EQUATION

STRUCTURE

4.1 INTRODUCTION:

Having discussed the nature of the simultaneous-equation models in the previous two lessons, in this lesson we turn to the problem of estimation of the parameters of such models. At the outset it may be noted that the estimation problem is rather complex because there are a variety of estimation techniques with varying statistical properties. In view of the introductory nature of this text, we shall consider only a few of these techniques. Our discussion will be simple and often heuristic, the finer points being left to the references.

4.2 OBJECTIVES:

1. To understand the Method Of Indirect Least Squares (Ils)
2. To understand the Method Of Two-Stage Least Squares (2sls)

4.3 THE METHOD OF INDIRECT LEAST SQUARES (ILS):

For a just or exactly identified structural equation, the method of obtaining the estimates of the structural coefficients from the OLS estimates of the reduced-form coefficients is known as the method of indirect least squares (ILS), and the estimates thus obtained are known as the indirect least squares estimates. ILS involves the following three steps:

Step 1. We first obtain the reduced-form equations. As noted in previously, these reduced-form equations are obtained from the structural equations in such a manner that the dependent variable in each equation is the only endogenous variable and is a function solely of the predetermined (exogenous or lagged endogenous) variables and the stochastic error term(s).

Step 2. We apply OLS to the reduced-form equations individually. This operation is permissible since the explanatory variables in these equations are predetermined and hence uncorrelated with the stochastic disturbances. The estimates thus obtained are consistent.9

Step 3. We obtain estimates of the original structural coefficients from the estimated reduced-form coefficients obtained in Step 2. As noted in previously, if an equation is exactly identified, there is a one-to-one correspondence between the structural and reduced-form coefficients; that is, one can derive unique estimates of the former from the latter.

As this three-step procedure indicates, the name ILS derives from the fact that structural coefficients (the object of primary enquiry in most cases) are obtained indirectly from the OLS estimates of the reduced-form coefficients.

4.4 ESTIMATION OF AN OVERIDENTIFIED EQUATION: THE METHOD OF TWO-STAGE LEAST SQUARES (2SLS)

Consider the following model:

$Y_{1t}= \beta_{10} + \quad + \beta_{11}Y_{2t}+\gamma_{2t}X_{1t}+ \gamma_{12}X_{2t}+u_{1t}$…… …1 income function

$Y_{2t}= \beta_{20} + \beta_{21}Y_{1t} \qquad +u_{1t}$………… 2 money supply function

where Y1 = income

Y2 = stock of money

X1 = investment expenditure

X2 = government expenditure on goods and services

The variables X1 and X2 are exogenous. The income equation, a hybrid of quantity-theory–Keynesian approaches to income determination, states that income is determined by money supply, investment expenditure, and government expenditure. The money supply function postulates that the stock of money is determined (by the Federal Reserve System) on the basis of the level of income.

Applying the order condition of identification, we can see that the income equation is underidentified whereas the money supply equation is overidentified. There is not much that can be done about the income equation short of changing the model specification. The overidentified money supply function may not be estimated by ILS because there are two estimates of β_{21}(the reader should verify this via the reduced-form coefficients).

As a matter of practice, one may apply OLS to the money supply equation, but the estimates thus obtained will be inconsistent in view of the likely correlation between the stochastic explanatory variable Y_1 and the stochastic disturbance term u_2. Suppose, however, we find a "proxy" for the stochastic explanatory variable Y_1 such that, although "resembling" Y_1 (in the sense that it is highly correlated with Y_1), it is uncorrelated with u_2. Such a proxy is also known as an instrumental variable. If one can find such a proxy, OLS can be used straightforwardly to estimate the money supply function. But how does one obtain such an instrumental variable? One answer is provided by the two-stage least squares (2SLS), developed independently by Henri Theil and Robert Basmann. As the name indicates, the method involves two successive applications of OLS. The process is as follows:

Stage 1. To get rid of the likely correlation between Y_1 and u_2, regress first Y1 on all the predetermined variables in the whole system, not just that equation. In the present case, this means regressing Y_1 on X_1 and X_2 as follows:

$$Y_{1t} = \hat{\Pi}_0 + \hat{\Pi}_1 X_{1t} + \hat{\Pi}_2 X_{2t} + \hat{u}_t \qquad 3$$

$$\hat{Y}_{1t} = \hat{\Pi}_0 + \hat{\Pi}_1 X_{1t} + \hat{\Pi}_2 X_{2t} \qquad 4$$

where $\hat{Y}_{1t}$ is an estimate of the mean value of Y conditional upon the fixed X's. Note that (20.4.3) is nothing but a reduced-form regression because only the exogenous or predetermined variables appear on the right-hand side. Equation (20.4.3) can now be expressed as

$$Y_{1t} = \hat{Y}_{1t} + \hat{u}_t \qquad 5$$

which shows that the stochastic Y_1 consists of two parts: $\hat{Y}_{1t}$, which is a linear combination of the nonstochastic X's, and a random component $\hat{U}_{1t}$. Following the OLS theory, $\hat{Y}_{1t}$ and $\hat{U}_{1t}$ are uncorrelated. (Why?)

Stage 2. The overidentified money supply equation can now be written as

$$\begin{aligned} Y_{2t} &= \beta_{20} + \beta_{21}(\hat{Y}_{1t} + \hat{u}_t) + u_{2t} \\ &= \beta_{20} + \beta_{21}\hat{Y}_{1t} + (u_{2t} + \beta_{21}\hat{u}_t) \\ &= \beta_{20} + \beta_{21}\hat{Y}_{1t} + u_t^* \end{aligned}$$

where $u_t^* = u_{2t} + \beta_{21}\hat{u}_t$. 6

Comparing(.6) with (2), we see that they are very similar in appearance, the only difference being that Y_1 is replaced by $\hat{Y}_{1t}$. What is the advantage of (.6)? It can be shown that although Y_1 in the original money supply equation is correlated or likely to be correlated with the disturbance term u_2 (hence rendering OLS inappropriate), $\hat{Y}_{1t}$ in (6) is uncorrelated with $\hat{U}_{1t}$ asymptotically, that is, in the large sample (or more accurately, as the sample size increases indefinitely). As a result, OLS can be applied to (.6), which will give consistent estimates of the parameters of the money supply function.

As this two-stage procedure indicates, the basic idea behind 2SLS is to "purify" the stochastic explanatory variable Y_1 of the influence of the stochastic disturbance u_2. This goal is accomplished by performing the reduced-form regression of Y1 on all the predetermined variables in the system (Stage 1), obtaining the estimates $\hat{Y}_{1t}$ and replacing $\hat{Y}_{1t}$ in the original equation by the estimated $\hat{Y}_{1t}$, and then applying OLS to the equation thus transformed (Stage 2). The estimators thus obtained are consistent; that is, they converge to their true values as the sample size increases indefinitely.

To illustrate 2SLS further, let us modify the income–money supply model as follows:

$$Y_{1t} = \beta_{10} + \beta_{12}Y_{2t} + \gamma_{11}X_{1t} + \gamma_{12}X_{2t} \qquad + u_{1t} \quad 7$$

$$Y_{2t} = \beta_{20} + \beta_{21}Y_{1t} \qquad + \gamma_{23}X_{3t} + \gamma_{24}X_{4t} + u_{2t} \quad 8$$

where, in addition to the variables already defined, X3 = income in the previous time period and X4 = money supply in the previous period. Both X3 and X4 are predetermined.

It can be readily verified that both Eqs. (7) and (.8) are overidentified. To apply 2SLS, we proceed as follows: In Stage 1 we regress the endogenous variables on all the predetermined variables in the system. Thus,

$$Y_{1t} = \hat{\Pi}_{10} + \hat{\Pi}_{11}X_{1t} + \hat{\Pi}_{12}X_{2t} + \hat{\Pi}_{13}X_{3t} + \hat{\Pi}_{14}X_{4t} + \hat{u}_{1t} \qquad 9$$

$$Y_{2t} = \hat{\Pi}_{20} + \hat{\Pi}_{21}X_{1t} + \hat{\Pi}_{22}X_{2t} + \hat{\Pi}_{23}X_{3t} + \hat{\Pi}_{24}X_{4t} + \hat{u}_{2t} \qquad 10$$

In Stage 2 we replace Y1 and Y2 in the original (structural) equations by their estimated values from the preceding two regressions and then run the OLS regressions as follows:

$$Y_{1t} = \beta_{10} + \beta_{12}\hat{Y}_{2t} + \gamma_{11}X_{1t} + \gamma_{12}X_{2t} + u^*_{1t} \qquad 11$$

$$Y_{2t} = \beta_{20} + \beta_{21}\hat{Y}_{1t} + \gamma_{23}X_{3t} + \gamma_{24}X_{4t} + u^*_{2t} \qquad 12$$

where $u^*_{1t} = u_{1t} + \beta_{12}\hat{u}_{2t}$ and $\hat{u}^*_{2t} = u_{2t} + \beta_{21}\hat{u}^*_{2t}$. The estimates thus obtained will be consistent.

Note the following features of 2SLS.

1. It can be applied to an individual equation in the system without directly taking into account any other equation(s) in the system. Hence, for solving econometric models involving a large number of equations, 2SLS offers an economical method. For this reason the method has been used extensively in practice.

2. Unlike ILS, which provides multiple estimates of parameters in the overidentified equations, 2SLS provides only one estimate per parameter.

3. It is easy to apply because all one needs to know is the total number of exogenous or predetermined variables in the system without knowing any other variables in the system.

4. Although specially designed to handle overidentified equations, the method can also be applied to exactly identified equations. But then ILS and 2SLS will give identical estimates.

5. If the R^2 values in the reduced-form regressions (that is, Stage 1 regressions) are very high, say, in excess of 0.8, the classical OLS estimates and 2SLS estimates will be very close.

4.5 SUMMARY AND CONCLUSIONS:

1. Assuming that an equation in a simultaneous-equation model is identified (either exactly or over-), we have several methods to estimate it.

2. These methods fall into two broad categories: Single-equation methods and systems methods.

3. For reasons of economy, specification errors, etc. the single-equation methods are by far the most popular. A unique feature of these methods is that one can estimate a single-equation in a multiequation model without worrying too much about other equations in the system. (Note: For identification purposes, however, the other equations in the system count.)

4. Three commonly used single-equation methods are OLS, ILS, and 2SLS.

5. Although OLS is, in general, inappropriate in the context of

simultaneous-equation models, it can be applied to the so-called recursive models where there is a definite but unidirectional cause-and-effect relationship among the endogenous variables.

6. The method of ILS is suited for just or exactly identified equations. In this method OLS is applied to the reduced-form equation, and it is from the reduced-form coefficients that one estimates the original structural coefficients.

7. The method of 2SLS is especially designed for overidentified equations, although it can also be applied to exactly identified equations. But then the results of 2SLS and ILS are identical. The basic idea behind 2SLS is to replace the (stochastic) endogenous explanatory variable by a linear combination of the predetermined variables in the model and use this combination as the explanatory variable in lieu of the original endogenous

variable. The 2SLS method thus resembles the instrumental variable method of estimation in that the linear combination of the predetermined variables serves as an instrument, or proxy, for the endogenous regressor.

8. A noteworthy feature of both ILS and 2SLS is that the estimates obtained are consistent, that is, as the sample size increases indefinitely, the estimates converge to their true population values. The estimates may not satisfy small-sample properties, such as unbiasedness and minimum variance.

Therefore, the results obtained by applying these methods to small samples and the inferences drawn from them should be interpreted with due caution.

4.6 EXCERCISES:

Q1. State whether each of the following statements is true or false:

a. The method of OLS is not applicable to estimate a structural equation in a simultaneous-equation model.

b. In case an equation is not identified, 2SLS is not applicable.

c. The problem of simultaneity does not arise in a recursive simultaneous-equation model.

d. The problems of simultaneity and exogeneity mean the same thing.

e. The 2SLS and other methods of estimating structural equations have desirable statistical properties only in large samples.

f. There is no such thing as an R2 for the simultaneous-equation model as a whole.

*g. The 2SLS and other methods of estimating structural equations are not applicable if the equation errors are autocorrelated and/or are correlated across equations.

h. If an equation is exactly identified, ILS and 2SLS give identical results.

Q2. Why is it unnecessary to apply the two-stage least-squares method to exactly identified equations?

Q3. Consider the following modified Keynesian model of income determination:

$$C_t = \beta_{10} + \beta_{11}Y_t + u_{1t}$$

$$I_t = \beta_{20} + \beta_{21}Y_t + \beta_{22}Y_{t-1} + u_{2t}$$

$$Y_t = C_t + I_t + G_t$$

where C = consumption expenditure

I = investment expenditure

Y = income

G = government expenditure

G_t and Y_{t-1} are assumed predetermined

a. Obtain the reduced-form equations and determine which of the preceding equations are identified (either just or over).

b. Which method will you use to estimate the parameters of the overidentified equation and of the exactly identified equation? Justify your answer.

Q4. Explain ILS in detail

Q.5 Explain 2SLS in detail.

. 4.7 Suggested Reading / References:

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2nd Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

UNIT-3

TIME SERIES ANALYSIS

Lesson-1

STATIONARY UNIT ROOT AND CO INTEGRATION

STRUCTURE

1.1 INTRODUCTION :

A time series is a sequence of data points, measured typically at successive points in time spaced at uniform time intervals. Time series data have a natural temporal ordering. This makes time series analysis distinct from cross-sectional studies, in which there is no natural ordering of the observations (e.g. explaining people's wages by reference to their respective education levels, where the individuals' data could be entered in any order). Time series analysis is also distinct from spatial data analysis where the observations typically relate to geographical locations (e.g. accounting for house prices by the location as well as the intrinsic characteristics of the houses). A stochastic model for a time series will generally reflect the fact that observations close together in time will be more closely related than observations further apart. In addition, time series models will often make use of the natural one-way ordering of time so that values for a given period will be expressed as deriving in some way from past values, rather than from future values (see time reversibility.)

Time series analysis can be applied to real-valued, continuous data, discrete numeric data, or discrete symbolic data (i.e. sequences of characters, such as letters and words in the English language.)

Stochastic Process: A random or stochastic process is a collection of random variable ordered in time.

Stationary Stochastic Process: (SSP) A type of stochastic process that has received a great deal of attention and scrutiny by time series analyses is SSP.

A stochastic process is said to be state if its means and variance are constant over time and the value of the Cov between the two time period depends only on the distance between the two time period.

$$E(Y_t)=u$$
$$Var(Y_t)=(-Y_t-u)^2=\sigma^2$$
$$Cov.Y_k=E[(Y_t-\mu)(Y_t+k-u)]$$
$$Cov.at\ \text{log}\ K.$$

1.2 OBJECTIVES:

1. The key objective is to explain the unit root test of stationarity.
2. To understand the Dickey Fuller Test.

3. To understand the Augmented Dickey Fuller Test.

1.3 THE UNIT ROOT TEST:

A test of stationary that has become widely popular over the past several years as unit root test.

$$Y_t=\rho Y_{t-1}+u_t \quad -1\le\rho\le1 \ldots\ldots\ldots\ldots(1)$$
$$\rho=1$$

1) Become a random walk model unit drift.

Sub Y_{t-1} from b/s

$$Y_t-Y_{t-1}=\rho Y_{t-1}-Y_{t-1}+u_t$$
$$=(p-1)Y_{t-1}+u_t \ldots\ldots\ldots\ldots(2)$$
$$\Delta Y_t=\delta Y_{t-1}+u_2 \ldots\ldots\ldots\ldots(3)$$

(Using 1st differential operate)

$\delta=0, if\ \delta=0\ then\ p=1,$ that is we have unit root, time series under consideration is non-stationary.

$$\Delta Y_t=(Y_t-Y_{t-1})=u_t \ldots\ldots\ldots(4)$$

The term non-stationary, random walk and unit root can be treated as synonymous.

$|\rho| \leq 1$ i.e. the time serves It is stationary.

1.4 DICKEY FULLER (DF) TEST:

A test use to find out if the estimated coefficient of Y_{t-1} in eq(1) is zero or not.

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

Δ = ist difference operator

$\delta = (\rho - 1)$

Dicky Fuller have shown that under the null hypotheses that $\delta = 0$, the estimated t value of the coefficient of Y_{t-1} in eq. (1) follows the t (two) statictic.

In the literature t statistic test is known as the DF test.

In if $\delta = 0$ is reflected we can use ……. t test.

DF test is estimated on 3 different forms, i.e. under 3 different null hypotheses.

Y_t is a random walk

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

Y_t is a walk within drift

$\Delta Y_t = \beta_1 + \delta Y_{t-1} + u_t$ (3)

Y_t is a random walk with drift around a stochastic trend.

$\Delta Y_t = \beta_1 + \beta_2 + \delta Y_{t-1} + u_t$ (4)

1.5 THE AUGMENTED DICKEY–FULLER (ADF) TEST:

In case the u_t are correlated, Dickey and Fuller have developed a test, known as the augmented Dickey–Fuller (ADF) test. This test is conducted by "augmenting" the preceding three equations by adding the lagged values of the dependent variable ΔY_t . The ADF test here consists of estimating the following regression:

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^{m} \alpha_i \Delta Y_{t-i} + \varepsilon_t \qquad \ldots\ldots 1$$

where ε_t is a pure white noise error term and where $\Delta Y_{t-1} = (Y_{t-1} - Y_{t-2})$, $\Delta Y_{t-2} = (Y_{t-2} - Y_{t-3})$, etc. The number of lagged difference terms to include is often determined empirically, the idea being to include enough terms so that the error term in (1) is serially uncorrelated. In ADF 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.

1.6 SUMMARY AND CONCLUSIONS:

Regression analysis based on time series data implicitly assumes that the underlying time series are stationary. The classical t tests, F tests, etc. are based on this assumption. In practice most economic time series are nonstationary. A stochastic process is said to be weakly stationary if its mean, variance, and autocovariances are constant over time (i.e., they are timeinvariant). At the informal level, weak stationarity can be tested by the correlogram of a time series, which is a graph of autocorrelation at various lags. For stationary time series, the correlogram tapers off quickly, whereas for nonstationary time series it dies off gradually. For a purely random series, the autocorrelations at all lags 1 and greater are zero. At the formal level, stationarity can be checked by finding out if the time series contains a unit root. The Dickey–Fuller (DF) and augmented Dickey–Fuller (ADF) tests can be used for this purpose. An economic time series can be trend stationary (TS) or difference stationary (DS). A TS time series has a deterministic trend, whereas a DS time series has a variable, or stochastic, trend. The common practice of including the time or trend variable in a regression model to detrend the data is justifiable only for TS time series. The DF and ADF tests can be applied to determine whether a time series is TS orDS.

1.7 LETS SUM IT UP:

This lesson has completed our survey of techniques for the analysis of time-series data. Contemporary econometric analysis of macroeconomic data has added considerable structure and formality to trending variables, which are more common than not in that setting. The variants of the Dickey–Fuller tests for unit roots are an indispensable tool for the analys to time series data.

1.8 EXCERCISES:

Q1. What is meant by weak stationarity?

Q2. What is meant by an integrated time series ?

Q3. If a time series is I(3), how many times would you have to difference it to make it stationary?

Q4. What are Dickey–Fuller (DF) and augmented DF tests?

Q5. What is the meaning of a unit root?

Q6. If a time series is I(3), how many times would you have to difference it to make it stationary?

Q7. What are Dickey–Fuller (DF) and augmented DF tests?

Q8. What is the meaning of a unit root?

. 1.9 Suggested Reading / References:

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2nd Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

Lesson-2

COINTEGRATION: REGRESSION OF A UNIT ROOT

STRUCTURE

2.1 INTRODUCTION

2.2 OBJECTIVES

2.3 TIME SERIES ON ANOTHER UNIT ROOT TIME SERIES

2.4 TESTING FOR COINTEGRATION

2.5 COINTEGRATING REGRESSION DURBIN–WATSON (CRDW) TEST

2.6 COINTEGRATION AND ERROR CORRECTION MECHANISM (ECM)

2.7 SUMMARY AND CONCLUSIONS

2.8 LETS SUM IT UP

2.9 EXCERCISES

2.10 SUGGESTED READING / REFERENCES

2.1 INTRODUCTION:

Studies in empirical macroeconomics almost always involve nonstationary and trending variables, such as income, consumption, money demand, the price level, trade flows, and exchange rates. Accumulated wisdom and the results of the previous sections suggest

that the appropriate way to manipulate such series is to use differencing and other transformations (such as seasonal adjustment) to reduce them to stationarity and then to analyze the resulting series as VARs or with the methods of Box and Jenkins. But recent research and a growing literature has shown that there are more interesting, appropriate ways to analyze trending variables. In the *fully specified* regression model

$$y_t = \beta x_t + \varepsilon_t ,$$

there is a presumption that the disturbances ε_t are a stationary, white noise series. But this presumption is unlikely to be true if y_t and x_t are integrated series. Generally, if two series are integrated to different orders, then linear combinations of them will be integrated to the higher of the two orders. Thus, if y_t and x_t are $I(1)$—that is, if both are trending variables—then we would normally expect $y_t - \beta x_t$ to be $I(1)$ regardless of the value of β, not $I(0)$ (i.e., not stationary). If y_t and x_t are each drifting upward with their own trend, then unless there is some relationship between those trends, the difference between them should also be growing, with yet another trend. There must be some kind of inconsistency in the model. On the other hand, if the two series are

both $I(1)$, then there *may* be a β such that

$$\varepsilon_t = y_t - \beta x_t$$

is I (0). Intuitively, if the two series are both I (1), then this partial difference between them might be stable around a fixed mean. The implication would be that the series are drifting together at roughly the same rate. Two series that satisfy this requirement are said to be **cointegrated,** and the vector $[1,-\beta]$ (or any multiple of it) is a **cointegrating vector.** In such a case, we can distinguish between a long-run relationship between *yt* and *xt* , that is, the manner in which the two variables drift upward together, and the short-run dynamics, that is, the relationship between deviations of *yt* from its long-run trend and deviations of *xt* from its long-run trend. If this is the case, then differencing of the data would be counterproductive, since it would obscure the long-run relationship between *yt* and *xt* . Studies of cointegration and a related technique, **error correction,** are concerned with methods of estimation that preserve the information about both forms of covariation.

2.2 OBJECTIVES:

1. To understand the meaning of cointegration.

2. To understand the Engel Granger (EG) or Augmented Engle Granger (AEG)

3. To understand the cointegrating regression Durbin–Watson (CRDW) test.

4. Cointegration and Error Correction Mechanism (ECM).

2.3 TIME SERIES ON ANOTHER UNIT ROOT TIME SERIES:

We have warned that the regression of a non-stationary time series on another non-stationary time series may produce a spurious regression. Suppose, then, that we regress PCE on PDI as follows:

$$PCE_t = \beta_1 + \beta_2 PDI_t + u_t$$

Let us write this as:

$$u_t = PCE_t - \beta_1 - \beta_2 PDI_t$$

………1, 2

Suppose we now subject u_t to unit root analysis and find that it is stationary; that is, it is I(0). This is an interesting situation, for although PCEt and PDIt are individually I(1), that is, they have stochastic trends, their linear combination (.2) is I(0). So to speak, the linear combination cancels out the stochastic trends in the two series. If you take consumption and income as two I(1) variables, savings defined as (income & consumption) could be I(0). As a result, a regression of consumption on income as in (1) would be meaningful (i.e., not spurious). In this case we say that the two variables are cointegrated. Economically speaking, two variables will be cointegrated if they have a long-term, or equilibrium, relationship between them. Economic theory is often expressed in equilibrium terms, such as Fisher's quantity theory of money or the theory of purchasing parity (PPP), just to name a few.

2.4 TESTING FOR COINTEGRATION:

A number of methods for testing cointegration have been proposed in the literature. We consider here two comparatively simple methods: (1) the DF or ADF unit root test on the residuals estimated from the co-integrating regression and (2) the cointegrating regression Durbin–Watson (CRDW) test.

Engle Granger (EG) or Augmented Engle Granger (AEG) Test.: We already know how to apply the DF or ADF unit root tests. All we have to do is estimate a regression like (1), obtain the residuals, and use the DF or ADF tests. There is one precaution to exercise, however. Since the estimated u_t are based on the estimated cointegrating parameter β_2, the DF and ADF critical significance values are not quite appropriate. Engle and Granger have calculated these values, which can be found in the references. Therefore, the DF and ADF tests in the present context are known as Engle–Granger (EG) and augmented Engle–Granger (AEG) tests. However, several software packages now present these critical values along with other outputs.

Let us illustrate these tests. We first regressed PCE on PDI and obtained the following regression:

$$\text{PCE}_t = -171.4412 + 0.9672\,\text{PDI}_t$$

$$t = (-7.4808)\ (119.8712) \qquad \ldots.\quad \textbf{(3)}$$

$$R^2 = 0.9940 \quad d = 0.5316$$

Since PCE and PDI are individually nonstationary, there is the possibility that this regression is spurious. But when we performed a unit root test on the residuals obtained from (3), we obtained the following results:

$$\Delta\hat{u}_t = -0.2753\,\hat{u}_{t-1}$$

$$t = (-3.7791) \ldots\ldots\ldots\ldots\ldots\ldots \qquad \textbf{(4)}$$

$$R^2 = 0.1422 \quad d = 2.2775$$

The Engle–Granger 1 percent critical τ value is −2.5899. Since the computed τ (= t) value is much more negative than this, our conclusion is that the residuals from the regression of PCE on PDI are $I(0)$; that is, they are stationary. Hence, (3) is a cointegrating regression and this regression is not spurious, even though individually the two variables are nonstationary. One can call (3) the **static** or **long run** consumption function and interpret its parameters as long run parameters. Thus, 0.9672 represents the long-run, or equilibrium, marginal propensity to consumer (MPC).

2.5 COINTEGRATING REGRESSION DURBIN–WATSON (CRDW) TEST:

An alternative, and quicker, method of finding out whether PCE and PDI are cointegrated is the CRDW test, whose critical values were first provided by Sargan and Bhargava. In CRDW we use the Durbin–Watson *d* obtained from the cointegrating regression, such as $d = 0.5316$ given in (3). But now the null hypothesis is that $d = 0$ rather than the standard $d = 2$. This is because we observed that $d \approx 2(1 - \hat{\rho})$, so if there is to be a unit root, the estimated ρ will be about 1, which implies that *d* will be about zero. On the basis of 10,000 simulations formed from 100 observations each, the 1, 5, and 10 percent critical values to test the hypothesis that the true $d = 0$ are 0.511, 0.386, and 0.322, respectively. Thus, if the computed *d* value is smaller than, say, 0.511, we reject the null hypothesis of cointegration at the 1 percent level. In our example, the value of 0.5316 is above this critical value, suggesting that PCE and PDI are cointegrated, thus reinforcing the finding on the basis of the EG test. *To sum up,* our conclusion, based on both the EG and CRDW tests, is that PCE and PDI are cointegrated. Although they individually exhibit random walks, there seems to be a stable long-run relationship between them; they will not wander away from each other.

2.6 COINTEGRATION AND ERROR CORRECTION MECHANISM (ECM):

We just showed that PCE and PDI are cointegrated; that is, there is a longterm, or equilibrium, relationship between the two. Of course, in the short run there may be disequilibrium. Therefore, one can treat the error term in

(2) as the "equilibrium error." And we can use this error term to tie the short-run behavior of PCE to its long-run value. The error correction mechanism (ECM) first used by Sargan and later popularized by Engle and Granger corrects for disequilibrium. An important theorem, known as the Granger representation theorem, states that if two variables Y and X are cointegrated, then the relationship between the two can be expressed as ECM. To see what this means, let us revert to our PCE–PDI example. Now consider the following model:

$$PCE_t = \alpha_0 + \alpha_1 \_PDI_t + \alpha_2 u_{t-1} + \varepsilon_t \;\ldots\ldots\; (5)$$

where as usual denotes the first difference operator, εt is a random error term, and ut−1 = (PCEt−1 − β1 − β2PDIt−1), that is, the one-period lagged value of the error from the cointegrating regression (1).

ECM equation (5) states that _PCE depends on PDI and also on the equilibrium error term.47 If the latter is nonzero, then the model is out of equilibrium. Suppose PDI is zero and ut−1 is positive. This means PCEt−1 is

too high to be in equilibrium, that is, PCEt−1 is above its equilibrium value of (α0 + α1PDIt−1). Since α2 is expected to be negative, the term α2ut−1 is negative

and, therefore,_PCEt will be negative to restore the equilibrium. That is, if PCEt is above its equilibrium value, it will start falling in the next period to correct the equilibrium error; hence the name ECM. By the same token, if ut−1 is negative (i.e., PCE is below its equilibrium value), α2ut−1 will be positive, which will cause _CPEt to be positive, leading PCEt to rise in period t. Thus, the absolute value of α2 decides how quickly the equilibrium is restored. In practice, we estimate ut−1 by ˆut−1 = (PCEt − ˆ β1 − ˆ β2PDIt). Returning to our illustrative example, the empirical counterpart of

(5) is:

PCEt = 11.6918 + 0.2906_PDIt − 0.0867 ˆut−1

t = (5.3249) (4.1717) (−1.6003) ………..(6)

R2 = 0.1717 d = 1.9233

Statistically, the equilibrium error term is zero, suggesting that PCE adjusts to changes in PDI in the same time period. As (6) shows, short-run changes in PDI have a positive impact on short-run changes in personal consumption. One can interpret 0.2906 as the short-run marginal propensity to consume (MPC); the long-run MPC is given by the estimated (static)equilibrium relation (3) as 0.9672.

2.7 SUMMARY AND CONCLUSIONS:

Regression of one time series variable on one or more time series variables often can give nonsensical or spurious results. This phenomenon is known as **spurious regression.** One way to guard against it is to find out if the time series are cointegrated. **Cointegration** means that despite being individually nonstationary, a linear combination of two or more time series can be stationary. The EG, AEG, and CRDW tests can be used to find out if two or more time series are cointegrated. Cointegration of two (or more) time series suggests that there is a long-run, or equilibrium, relationship between them. The **error correction mechanism (ECM)** developed by Engle and Granger is a means of reconciling the short-run behavior of an economic variable with its long-run behavior. The field of time series econometrics is evolving. The established results and tests are in some cases tentative and a lot more work remains. An important question that needs an answer is why some economic time series are stationary and some are nonstationary.

2.8 LETS SUM IT UP:

This modelling framework is a distinct extension of the regression modeling where this discussion began. Cointegrated relationships and equilibrium relationships form the basis the timeseries counterpart to regression relationships. But, in this case, it is not the conditional mean as such that is of interest. Here, both the long-run equilibrium and short-run relationships around trends are of interest and are studied in the data.

2.9 EXCERCISES:

Q1. Find the autocorrelations and partial autocorrelations for the MA(2) process

$$\varepsilon t = vt - \theta 1vt-1 - \theta 2vt-2.$$

Q2. What are Engle–Granger (EG) and augmented EG tests?

Q3. What is the meaning of cointegration?

Q4. What is the difference, if any, between tests of unit roots and tests of cointegration?

Q5. What is spurious regression?

Q6. What is the connection between cointegration and spurious regression?

Q7. Descricbe Cointegration and Error correction mechanism (ECM)?

. **2.10 Suggested Reading / References:**

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2nd Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

LESSON-3

RANDOM WALK MODEL (RWM)

STRUCTURE

3.7 EXCERCISES

3.8 SUGGESTED READING / REFERENCES

3.1 INTRODUCTION:

In discussing the nature of the unit root process, we noted that a random walk process may have no drift, or it may have drift or it may have both deterministic and stochastic trends. To allow for the various possibilities, the DF test is estimated in three different forms, that is, under three different null hypotheses.

***Yt* is a random walk:** $Yt = \delta Yt{-}1 + ut$ **(2)**

***Yt* is a random walk with drift:** $Yt = \beta 1 + \delta Yt{-}1 + ut$ **(4)**

around a stochastic trend: $Yt = \beta 1 + \beta 2t + \delta Yt{-}1 + ut$ **(5)**

where t is the time or trend variable. In each case, the *null hypothesis* is that $\delta = 0$; that is, there is a unit root—the time series is nonstationary. The alternative hypothesis is that δ is less than zero; that is, the time series is stationary. If the null hypothesis is rejected, it means that Yt is a stationary time series with zero mean in the case of (2), that Yt is stationary with a nonzero mean [$= \beta 1/(1 - \rho)$] in the case of (4), and that Yt is stationary around a deterministic trend in (5). *It is extremely important to note that the critical values of the tau test to test the hypothesis that* $\delta = 0$, *are different for each of the preceding three specifications of the DF test.* Moreover, if, say, specification (4) is correct, but we estimate (2), we will be committing a specification error, whose consequences. The same is true if we estimate (4) rather than the true (5). Of course, there is no way of knowing which specification is correct to begin with. Some trial and error is inevitable, data mining notwithstanding. The actual estimation procedure is as follows: Estimate (2), or (4) by OLS; divide the estimated coefficient of $Yt{-}1$ in each case by its standard error to compute the (τ) tau statistic; and refer to the DF tables (or any statistical package). If the computed absolute value of the tau statistic ($|\tau|$) exceeds the DF or MacKinnon critical tau values, we reject the hypothesis that $\delta = 0$, in which case the time series is stationary. On the other hand, if the computed $|\tau|$ does not exceed the critical tau value, we do not reject the null hypothesis, in which case the time series is nonstationary. Make sure that you use the appropriate critical τ values.

3.2 OBJECTIVES:

1. Understand the concept of non-stationary time series.

2. Understand Random walk model with drift and Random walk model without drift.

3. Understand the trend stationary process.

3.3 RANDOM WALK MODEL:

3.3.1 Non-stationary:- If a time series is not stationary it is called non-stationary time series. RWM is a classical example. It is said that asset purchased such as stock purchases follows a random walk i.e. they are non- stationary..

Two types:

3.3.1.1 Random walk without drift (i.e. no constant)

3.3.1.2 Random walk with drift (i.e. a contant term)

3.3.1.1 RANDOM WALK WITHOUT DRIFT:

Y_t -Y_{t-1} + u_t (1) U_t = white nose.

Y_t = RW

We can write eq. (1) as

$Y_1 = Y_0 + u_1$

$Y_2 = Y_1 + u_2 = Y_0 + u_1 + u_2$

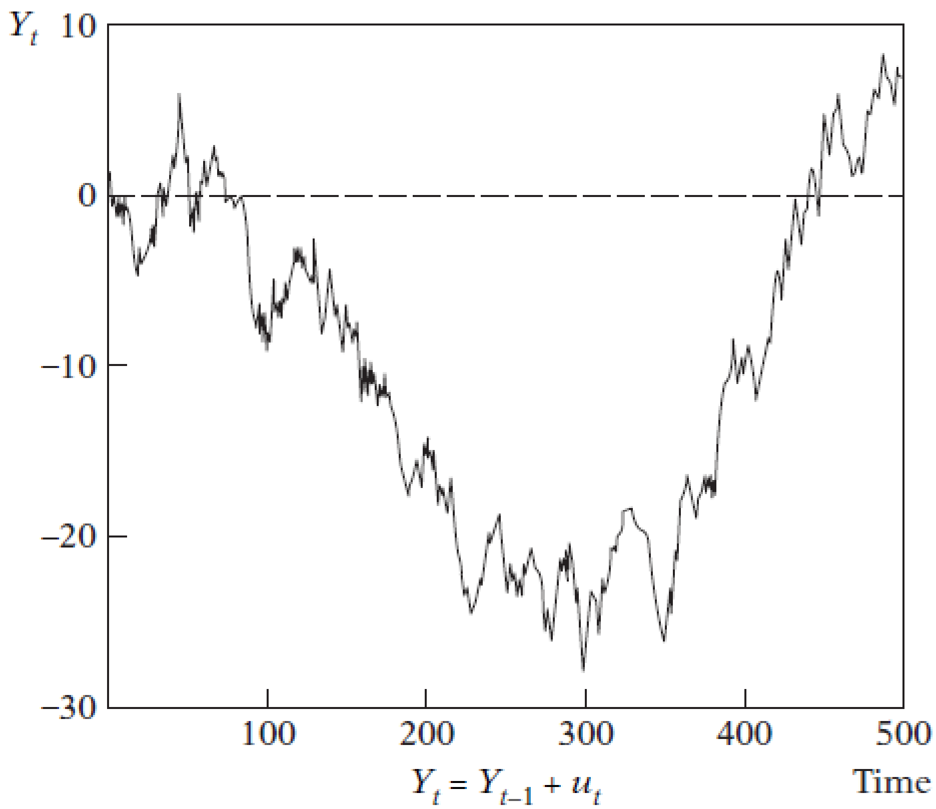
$Y_3 = Y_2 + u_3 = Y_0 + u_1 + u_2 + u_3$

$Y_t = Y_0 + \Sigma u_1$ (2)

$\therefore E(Y_t)= E(Y_0 + \Sigma u_t) = Y_0$ (6)

$Var\ Y_t = t\sigma^2$

RWM without drift is a non stationary stochastic process.



A random walk without drift.

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

3.3.1.2 RANDOM WALK WITH DRIFT:

$Y_t = \delta + Y_{t-1} + u_t$ (1) δ = drift parameter

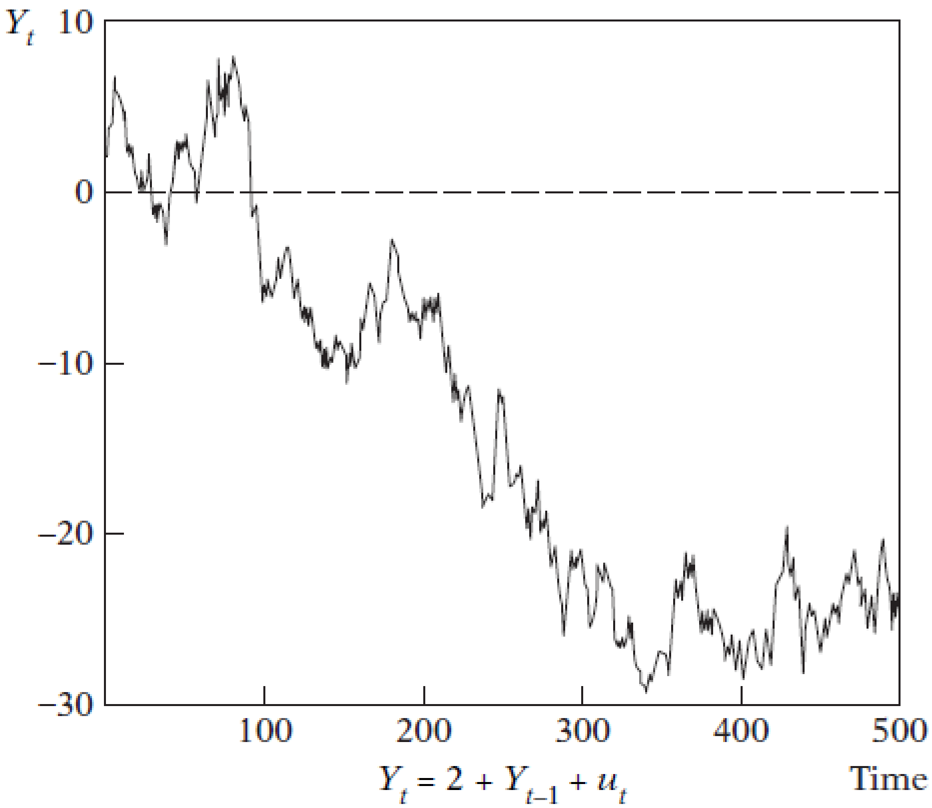
$Y_t + Y_{t-1} = \Delta Y_t = \delta + u_t$ (2)

Yt = upward or downward depending on δ being +ve or -ve

$E(Y_t) = Y_0 + t\delta$ (3)

$Var(Y_t) = = t\,\sigma^2$ (4)

RWM with drift mean the variance increasing over time, again violating the condition of (weak) stationary. In other word RWM with or root drift is a nonstationary stochastic process the RWM is an example of unit root process.



A random walk with drift.

3.4 COMMON TREND:

Consider an example. Suppose that two $I(1)$ variables

have a linear trend,

$$y_{1t} = \alpha + \beta t + u_t ,$$

$$y_{2t} = \gamma + \delta t + v_t ,$$

where u_t and v_t are white noise. A linear combination of y_{1t} and y_{2t} with vector $(1, \theta)$ produces the new variable,

$$z_t = (\alpha + \theta\gamma) + (\beta + \theta\delta)t + u_t + \theta v_t ,$$

which, in general, is still $I(1)$. In fact, the only way the z_t series can be made stationary is if $\theta = -\beta/\delta$. If so, then the effect of combining the two variables linearly is *to remove the common linear trend,* which is the basis of Stock andWatson's (1988) analysis of the problem. But their observation goes an important step beyond this one. *The only way that y_{1t} and y_{2t} can be cointegrated to begin with is if they have a common trend of some sort.* To continue, suppose that instead of the linear trend t, the terms on the right-hand side, y_1 and y_2, are functions of a random walk, $w_t = w_{t-1} + \eta_t$, where η_t is white noise.

The analysis is identical. But now suppose that each variable y_{it} has its own random walk component w_{it} , $i = 1, 2$. Any linear combination of y_{1t} and y_{2t} must involve *both* random walks. It is clear that they cannot be cointegrated unless, in fact, $w_{1t} = w_{2t}$. That is, once again, they must have a **common trend.** Finally, suppose that y_{1t} and y_{2t} share two common trends,

$$y_{1t} = \alpha + \beta t + \lambda w_t + u_t ,$$

$$y_{2t} = \gamma + \delta t + \pi w_t + v_t .$$

We place no restriction on λ and π. Then, a bit of manipulation will show that it is not possible to find a linear combination of $y1t$ and $y2t$ that is cointegrated, even though they share common trends. The end result for this example is that if $y1t$ and $y2t$ are cointegrated, then they must share exactly one common trend.

As Stock and Watson determined, the preceding is the crux of the cointegration of economic variables. A set of M variables that are cointegrated can be written as a stationary component plus linear combinations of a smaller set of common trends. If the cointegrating rank of the system is r , then there can be up to $M-r$ linear trends and

$M-r$ common random walks. [See Hamilton (1994, p. 578).] (The two-variable case is special. In a two-variable system, there can be only one common trend in total.) The effect of the cointegration is to purge these common trends from the resultant variables.

3.4.1 TREND:

If the trend in a time series is completely predictable and not variable, we call it **as deterministic trend.**

Whereas if it is not predictable, we call it **a stochastic trend**.

$$Y_t = \beta_1 + \beta_2 t + \beta_3 Y_{t-1} + \mu_t \ldots\ldots\ldots\ldots (1)$$

T= Time measured chronologically

Rare random walk

$$Y_t = Y_{t-1} + u_t \ldots\ldots\ldots\ldots (2)$$

$$\beta_1 = 1, \beta_2 = 0, \beta_3 = 1, \text{ in eq}(1) \text{ we get } (4)$$

$$\Delta Y_t = (Y_t - Y_{t-1}) = u_t \ldots\ldots\ldots (2a) \,\&\, (2) \text{ eq}$$

Random walk with drift

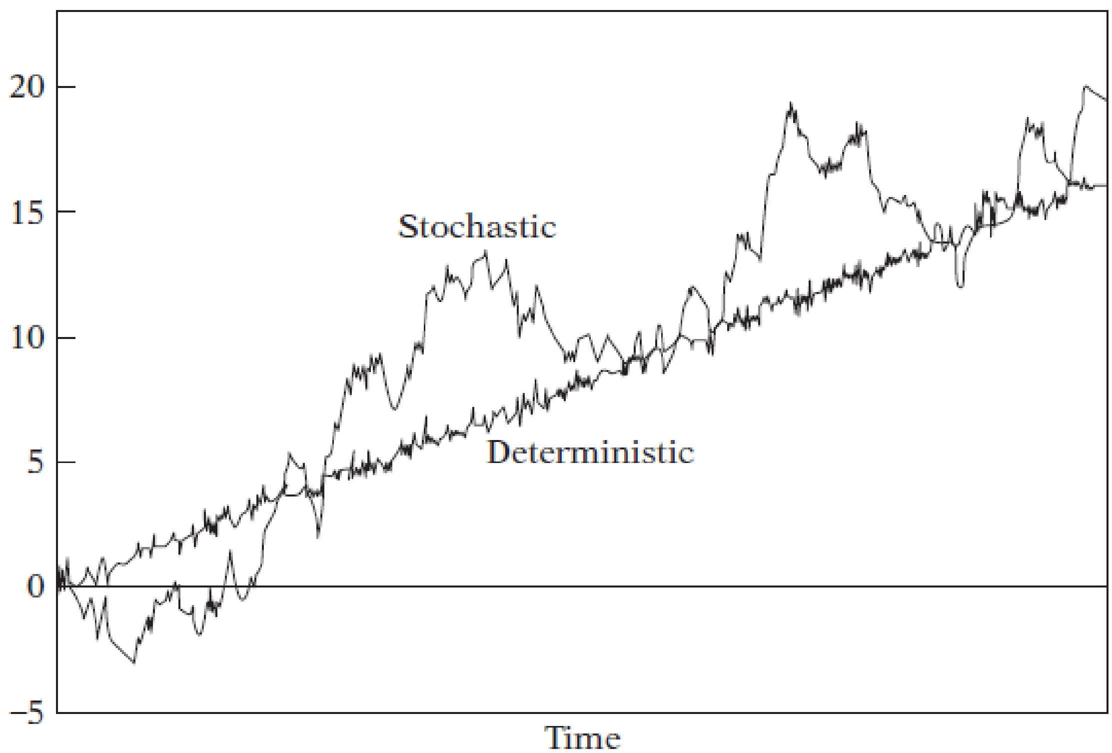
$\beta_1 \neq 0, \beta_2 = 0, \beta_3 = 1.$
$Y_t = \beta_1 + Y_{t-1} + u_t \ldots\ldots\ldots\ldots(3)$
$Y_t Y_{t-1} = \Delta Y_t = \beta_1 + ul \ldots\ldots(3a)$
$(B_1 > 0) positi\&(\beta_1 < 1) negati$

This is stochastic trend

Determistic trend

$\beta_1 \neq 0, \beta_2 \neq 0, \beta_3 = 0$
$Y_t = \beta_1 + \beta_2 t + u_t \ldots\ldots\ldots\ldots(4)$

Called as trend stationary process (TSP).



Deterministic versus stochastic trend.

3.4.2 PROPERTIES OF INTEGRATED SERIES :

Let Y_t, Y_t & Z_t be 3 time

1) If $Y_t \sim I(o) \& Y_t \sim I(1) then Z_t = (X_t + Y_t) = I(1)$ then $Z_t = (X_t + Y_t) = I(1)$.

i.e. Linear combination or sim of stationary and non stationary time series is nonstationary.

2) If $X_t \sim I(d), then Z_t = (a + bX) = I(d$

Where a & b are constant.

$\therefore \quad Z_t = (a + bX_t) \sim I(o$

3) $X_t \sim I(d_1) \& Y_t \sim 1(d_2)$

$\& Z_t = (a x_t + b y) \sim 1(d_2)$ where $d_1 < d_2$.

4) If $X_t \sim I(d) \& Y_t \sim I(d) the$

$Z = (aX_t + bY_t) \sim 1(d)$

d^* is generated cq = d

$d^* < d$

3.5 SUMMARY AND CONCLUSIONS:

This lesson has completed our survey of techniques for the analysis of time-series data. While was about extensions of regression modeling to time-series setting, most of the results in this Lesson focus on the internal structure of the individual time series, themselves.We began with the standard models for time-series processes.While the empirical distinction between, say AR*(p)* and MA*(q)* series may seem ad hoc, the word decomposition assures that with enough care, a variety of models can be used to analyze a time series. This lesson described what is arguably the fundamental tool of modern macroeconometrics, the tests for nonstationarity.

3.6 LETS SUM IT UP:

In the last we conclude that if a time series is not stationary it is called non-stationary time series. Random Walk Model is a classical example. And further if two *I (*1*)* variables are cointegrated, then some linear combination of them is *I (*0*)*. Intuition should suggest that the linear combination does not mysteriously create a well-behaved new variable; rather, something present in the original variables must be missing from the aggregated one.

3.7 EXCERCISES:

Q1. What is pure random walk?

Q.2 What is trend?

Q.3. What is random walk with drift?

Q4. Describe the various properties of Integrated Series?

Q5. What do you mean by trend stationary process?

. **3.8 Suggested Reading / References:**

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2nd Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

Lesson-4

VECTOR AUTOREGRESSION (VAR) AND MEASURING VOLATILITY IN FINANCIAL TIME SERIES:THE ARCH AND GARCH MODELS

STRUCTURE

4.8 EXCERCISES

4.9 SUGGESTED READING / REFERENCES

4.1 INTRODUCTION:

Forecasting is an important part of econometric analysis, for some people probably the most important. How do we forecast economic variables, such as GDP, inflation, exchange rates, stock prices, unemployment rates, and myriad other economic variables? In this lesson we discuss two methods of forecasting that have become quite popular: (1) **autoregressive integrated moving average (ARIMA),** popularly known as the **Box–Jenkins** methodology,1 and (2) **vector autoregression (VAR).**

In this lesson we also discuss the special problems involved in forecasting prices of financial assets, such as stock prices and exchange rates. These asset prices are characterized by the phenomenon known as **volatility clustering,** that is, periods in which they exhibit wide swings for an extended time period followed by a period of comparative tranquility. One only has to look at the Dow Jones Index in the recent past. The so-called **autoregressive conditional heteroscedasticity (ARCH)** or **generalized autoregressive conditional heteroscedasticity (GARCH)** models can capture such volatility clustering.

The topic of economic forecasting is vast, and specialized books have been written on this subject. Our objective in this lesson is to give the reader just a glimpse of this subject. The interested reader may consult the references for further study. Fortunately, most modern econometric packages have user-friendly introductions to several techniques discussed in this lesson The linkage between this lesson and the previous lesson is that the forecasting methods discussed below assume that the underlying time series are stationary or they can be made stationary with appropriate transformations. As we progress through this lesson, you will see the use of the several concepts that we introduced in the last lesson

4.2 OBJECTIVES:

1. The key objective is to understand the economic forecasting based on time series data.

2. To understand the Vector Autoregression(VAR).

3. To understand the Autoregressive conditional heteroscedasticity(ARCH) and Generalised autoregressive conditioinal heteroscedasticity(GARCH).

4.3 VECTOR AUTOREGRESSION (VAR):

According to Sims, if there is true simultaneity among a set of variables, they should all be treated on an equal footing; there should not be any a priori distinction between endogenous and exogenous variables. It is in this spirit that Sims developed his **VAR** model.

4.3.1 Estimation or VAR:

Returning to the Canadian money interest rate, we saw that when we introduced six lags of each variable as regressors, we could not reject the hypothesis that there was bilateral causality between money (M_1) and interest rate, R (90-day corporate interest rate). That is, M_1 affects R and R affects M_1. These kinds of situations are ideally suited for the application of VAR.

To explain how a VAR is estimated, we will continue with the preceding example. For now we assume that each equation contains k lag values of M(as measured by M_1) and R. In this case, one can estimate each of the following equations by OLS.

$$M_{1t} = \alpha + \sum_{j=1}^{k} \beta_j M_{t-j} + \sum_{j=1}^{k} \gamma_j R_{t-j} + u_{1t}$$

$$R_t = \alpha' + \sum_{j=1}^{k} \theta_j M_{t-j} + \sum_{j=1}^{k} \gamma_j R_{t-j} + u_{2t} \qquad \text{.......1,2}$$

where the u's are the stochastic error terms, called **impulses** or **innovations** or **shocks** in the language of VAR.

4.4 SOME PROBLEMS WITH VAR MODELING:

The advocates of VAR emphasize these virtues of the method:

(1) The method is simple; one does not have to worry about determining which variables are endogenous and which ones exogenous. All variables in VAR are endogenous.

(2) Estimation is simple; that is, the usual OLS method can be applied to each equation separately.

(3) The forecasts obtained by this method are in many cases better than those obtained from the more complex simultaneous-equation models.

But the critics of VAR modeling point out the following problems:

1. Unlike simultaneous-equation models, a VAR model is a theoretic because it uses less prior information. Recall that in simultaneous-equation models exclusion or inclusion of certain variables plays a crucial role in the identification of the model.

2. Because of its emphasis on forecasting, VAR models are less suited for policy analysis.

3. The biggest practical challenge in VAR modeling is to choose the appropriate lag length. Suppose you have a three-variable VAR model and you decide to include eight lags of each variable in each equation. You will have 24 lagged parameters in each equation plus the constant term, for a total of 25 parameters. Unless the sample size is large, estimating that many parameters will consume a lot of degrees of freedom with all the problems associated with that.

4. Strictly speaking, in an m-variable VAR model, all the m variables should be (jointly) stationary. If that is not the case, we will have to transform the data appropriately (e.g., by first-differencing). As Harvey notes, the results from the transformed data may be unsatisfactory. He further notes that “The usual approach adopted by VAR aficionados is therefore to work in levels, even if some of these series are nonstationary. In this case, it is important to recognize the effect of unit roots on the distribution of estimators.”Worse yet, if the model contains a mix of I(0) and I(1) variables, that is, a mix of stationary and nonstationary variables, transforming the data will not be easy.

5. Since the individual coefficients in the estimated VAR models are often difficult to interpret, the practitioners of this technique often estimate the so-called impulse response function (IRF). The IRF traces out the response of the dependent variable in the VAR system to shocks in the errorterms, such as u_1 and u_2 in Eqs. (.1) and (2). Suppose u1 in the M_1equation **<u>increases</u>** by a value of one standard deviation. Such a shock or change will change M_1 in the current as well as future periods. But since M_1 appears in the R regression, the change in u1 will also have an impact on R. Similarly, a change of one standard deviation in u2 of the R equation will have an impact on M1. The IRF traces out the impact of such shocks for several periods in the future. Although the utility of such IRF analysis has been questioned by researchers, it is the centerpiece of VAR analysis.

<u>4.5 MEASURING VOLATILITY IN FINANCIAL TIME SERIES: THE ARCH AND GARCH MODELS</u>

As noted in the introduction to this lesson, financial time series, such as stock prices, exchange rates, inflation rates, etc. often exhibit the phenomenon of **volatility clustering,** that is, periods in which their prices show wide swings for an extended time period followed by periods in which there is relative calm. As Philip Franses notes: Since such [financial time series] data reflect the result of trading among buyers and sellers at, for example, stock markets, various sources of news and other exogenous economic events may have an impact on the time series pattern of asset prices. Given that news can lead to various interpretations, and also given that specific economic events like an oil crisis can last for some time, we often observe that large positive and large negative observations in financial time series tend to appear in clusters.

Knowledge of volatility is of crucial importance in many areas. For example, considerable macroeconometric work has been done in studying the variability of inflation over time. For some decision makers, inflation in itself may not be bad, but its variability is bad because it makes financial planning difficult.

The same is true of importers, exporters, and traders in foreign exchange markets, for variability in the exchange rates means huge losses or profits. Investors in the stock market are obviously interested in the volatility of stock prices, for high volatility could mean huge losses or gains and hence greater uncertainty. In volatile markets it is difficult for companies to raise capital in the capital markets.

How do we model financial time series that may experience such volatility? For example, how do we model times series of stock prices, exchange rates, inflation, etc.? A characteristic of most of these financial time series is that in their *level form* they are random walks; that is, they are nonstationary. On the other hand, in the first difference form, they are generally stationary, as we saw in the case of GDP series in the previous lesson even though GDP is not strictly a financial time series.

Therefore, instead of modeling the levels of financial time series, why not model their first differences? But these first differences often exhibit wide swings, or **volatility,** suggesting that the variance of financial time series varies over time. How can we model such "varying variance"? This is where the so-called **autoregressive conditional heteroscedasticity (ARCH)** model originally developed by Engle comes in handy.

4.5.1 A Note on the GARCH Model:

Since its "discovery" in 1982, ARCH modeling has become a growth industry, with all kinds of variations on the original model. One that has become popular is the **generalized autoregressive conditional heteroscedasticity (GARCH)** model, originally proposed by Bollerslev. The simplest GARCH model is the GARCH(1, 1) model, which can be written as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 \sigma_{t-1}^2$$

which says that the conditional variance of *u* at time *t* depends not only on the squared error term in the previous time period [as in ARCH(1)] but also on its conditional variance in the previous time period. This model can be generalized to a GARCH(*p, q*) model in which there are *p* lagged terms of the squared error term and *q* terms of the lagged conditional variances.

We will not pursue the technical details of these models, as they are involved, except to point out that a GARCH(1, 1) model is equivalent to an ARCH(2) model and a GARCH(*p*, *q*) model is equivalent to an ARCH(*p*+ *q*) model.

For our U.S./U.K. exchange rate and NYSE stock return examples, we have already stated that an ARCH(2) model was not significant, suggesting that perhaps a GARCH(1, 1) model is not appropriate in these cases.

4.6 SUMMARY AND CONCLUSIONS:

1. Box–Jenkins and VAR approaches to economic forecasting are alternatives to traditional single- and simultaneous-equation models.

2. To forecast the values of a time series, the basic Box–Jenkins strategy is as follows:

a. First examine the series for stationarity. This step can be done by computing the autocorrelation function (ACF) and the partial autocorrelation function (PACF) or by a formal unit root analysis. The correlograms associated with ACF and PACF are often good visual diagnostic tools.

b. If the time series is not stationary, difference it one or more times to achieve stationarity.

c. The ACF and PACF of the stationary time series are then computed to find out if the series is purely autoregressive or purely of the moving average type or a mixture of the two. From broad guidelines given in Table 22.1 one can then determine the values of p and q in the ARMA process to be fitted. At this stage the chosen ARMA(p, q) model is tentative.

d. The tentative model is then estimated.

e. The residuals from this tentative model are examined to find out if they are white noise. If they are, the tentative model is probably a good approximation to the underlying stochastic process. If they are not, the process is started all over again. Therefore, the Box–Jenkins method is iterative.

f. The model finally selected can be used for forecasting.

3. The VAR approach to forecasting considers several time series at a time. The distinguishing features of VAR are as follows:

a. It is a truly simultaneous system in that all variables are regarded as endogenous.

b. In VAR modeling the value of a variable is expressed as a linear function of the past, or lagged, values of that variable and all other variables included in the model.

c. If each equation contains the same number of lagged variables in the system, it can be estimated by OLS without resorting to any systems method, such as two-stage least squares (2SLS) or seemingly unrelated regressions (SURE).

d. This simplicity of VAR modeling may be its drawback. In view of the limited number of observations that are generally available in most economic analyses, introduction of several lags of each variable can consume a lot of degrees of freedom.

e. If there are several lags in each equation, it is not always easy to interpret each coefficient, especially if the signs of the coefficients alternate. For this reason one examines the impulse response function (IRF) in VAR modeling to find out how the dependent variable responds to a shock administered to one or more equations in the system.

f. There is considerable debate and controversy about the superiority of the various forecasting methods. Single-equation, simultaneous-equation, Box–Jenkins, and VAR methods of forecasting have their admirers as well as detractors. All one can say is that there is no single method that will suit all situations. If that were the case, there would be no need for discussing the various alternatives. One thing is sure: The Box–Jenkins and VAR methodologies have now become an integral part of econometrics.

4. We also considered in this lesson a special class of models, ARCH and GARCH, which are especially useful in analyzing financial time series, such as stock prices, inflation rates, and exchange rates. A distinguishing feature of these models is that the error variance may be correlated over time because of the phenomenon of volatility clustering.

In this connection we also pointed out that in many cases a significant Durbin–Watson *d* may in fact be due to the ARCH or GARCH effect.

4.7 LETS SUM IT UP:

This lesson examined different model of stochastic volatility, the VAR,ARCH and GARCH model. These models has proved especially useful for analyzing financial data such as exchange rates, inflation, and market returns.

4.8 EXCERCISES:

Q.1. What is VAR (vector auto regression)?

Q2. What is Autoregressive conditional heteroscedasticity(ARCH)?

Q3. Explain the Generalised Autoregressive conditional heteroscedasticity(GARCH)?

Q4. Explain the Box-Jenkins Methodology?

Q5. Elaborate the estimation process of VAR model?

4.9 Suggested Reading / References:

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2^{nd} Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

Unit -4

Panel Data

Lesson-1

PANEL DATA TECHNIQUES

STRUCTURE

4.1 INTRODUCTION

1.2 OBJECTIVES

1.3 PANEL DATA

1.4 POOLING DATA

1.5 CHOW TEST

1.6 DIFFERENCE WITH MORE THAN TWO TIME PERIODS

1.7 SUMMARY AND CONCLUSIONS

1.8 LETS SUM IT UP

1.9 EXCERCISES

1.10 SUGGESTED READING / REFERENCES

4.1 INTRODUCTION:

There are three types of data that are generally available for empirical analysis, namely, time series, cross section, and panel. In time series data we observe the values of one or more variables over a period of time (e.g., GDP for several quarters or years). In cross-section data, values of one or more variables are collected for several sample units, or entities, at the same point in time (e.g., crime rates for 50 states in the United States for a given year). In panel data the same cross-sectional unit (say a family or a firm or a state) is surveyed over time. In short, panel data have space as well as time dimensions. There are other names for panel data, such as **pooled data** (pooling of time series and cross-sectional observations), **combination of time series and cross-section data, micropanel data, longitudinal data** (a study over time of a variable or group of subjects), **event history analysis** (e.g., studying the movement over time of subjects through successive states or conditions), **cohort analysis** (e.g., following the career path of 1965 graduates of a business school). Although there are subtle variations, all these names *essentially connote movement over time of cross-sectional units.*We will therefore use the term panel data in a generic sense to include one or more of these terms. *And we will call regression models based on such data panel data regression models.* Panel data are now being increasingly used in economic research. Some of the well-known panel data sets are:

1. The **Panel Study of Income Dynamics (PSID)** conducted by the Institute of Social Research at the University of Michigan. Started in 1968, each year the Institute collects data on some 5000 families about various socioeconomic and demographic variables.

2. The Bureau of the Census of the Department of Commerce conducts a survey similar to PSID, called the **Survey of Income and Program Participation (SIPP).** Four times a year, the respondents are interviewed about their economic condition.

There are also many other surveys that are conducted by various governmental agencies. At the outset a warning is in order. The topic of panel data regressions is vast, and some of the mathematics and statistics involved is quite complicated.

We only hope to touch on some of the essentials of the panel data regression models, leaving the details for the references. But be forewarned that some of these references are highly technical. Fortunately, user-friendly software packages such as Limdep, PcGive, SAS, STATA, Shazam, and Eviews, among others, have made the task of actually implementing panel data regressions quite easy.

1.2 OBJECTIVES:

Our objective is to get familiar with all these types of data:

1. Panel data
2. Pooling data
3. The chow test
4. Difference with more than two time periods

1.3 PANEL DATA :

In statistics and econometrics, the term panel data refers to multi-dimensional data frequently involving measurements over time. Panel data contain observations of multiple phenomena obtained over multiple time periods for the same firms or individuals. In biostatistics, the term longitudinal data is often used instead, wherein a subject or cluster constitutes a panel member or individual in a longitudinal study.

Time series and cross-sectional data are special cases of panel data that are in one dimension only (one panel member or individual for the former, one time point for the latter).

Example

balanced panel:

person	year	income	age	sex
1	2001	1300	27	1
1	2002	1600	28	1
1	2003	2000	29	1
2	2001	2000	38	2
2	2002	2300	39	2
2	2003	2400	40	2

unbalanced panel:

person	year	income	age	sex
1	2001	1600	23	1
1	2002	1500	24	1
2	2001	1900	41	2
2	2002	2000	42	2
2	2003	2100	43	2
3	2002	3300	34	1

In the example above, two data sets with a two-dimensional panel structure are shown, although the second data set might be a three-dimensional structure since it has three people. Individual characteristics (income, age, sex. education) are collected for different persons and different years. In the left data set two persons (1, 2) are observed over three years (2001, 2002, 2003). Because each person is observed every year, the left-hand data set is called a balanced panel, whereas the data set on the right hand is called an unbalanced panel, since Person 1 is not observed in year 2003 and Person 3 is not observed in 2003 or 2001.

There are other names for panel data, such as pooled data(pooling of time series and cross-sectional observations), combination of time series and cross-section data, micro panel data, longitudinal data(a study over time of a variable or group of subjects), event history analysis(e.g., studying the movement over time of subjects through successive states or conditions), cohort analysis(e.g., following the career path of 1965 graduates of a business school). Although there are subtle variations, all these names essentially connote movement over time of cross-sectional units. We will therefore use the term panel data in a generic sense to include one or more of these terms. And we will call regression models based on such data panel data regression models.

Time series data we obtain from the value of one or more variable over a period of time (GDP for several quarter of year.

Cross Section: Value of one or more valuable are collected for several sample units of at same pt. in time (Crime record of 50 states in last one year).

Panel dates the same cross sectional unit is surveyed over time. In short, panel data hence space as well as time dimensions.

Other name for panel data such as pooled data (pooting of time series or cross section observation)., micropanel data, longitudinal detail, evens history available.

Well know panel data sets are:

1. Panel study of income dynamics.

2. Survey of Income and Program Participation (SIPP).

1.4 POOLING DATA :

Relationship between the dependent variable and at least some of the independent variable remains constant over time.

1.5 CHOW TEST :

Chow test, which is simply on F test can be used to determine whether a multiplier regression function differs across two groups.

It can also be completes for more than two periods. Just as in the two period case, it is usually more interesting to allow the intercepts to change over time and then test whether the slope coefficients have changed overtime.

Chow test for two periods by intersecting each variable with a year dummy for one of the two years and testing for joint significance of the year dummy and all of the interaction terms.

We can test the constancy of slope coefficient generally by intersecting all of the perior and dommirs (except that definincy the box group) with one several or all of the explanatory variables and test the joins significance of the intersection terms.

1.6 DIFFERENCE WITH MORE THAN TWO TIME PERIODS :

We can also use differencing with more than two time periods.

Example.

Suppose having N individuals and T=3 the time period for each individual.

A general fixed effects model is

$$Y_{it} = \delta_1 + \delta_2 d2_t + \delta_3 d3_t + \beta_1 x_{it1} + \ldots\ldots + \beta_k y_{itk} + u + u_{it} \quad \ldots\ldots\ldots\ldots\ldots\ldots 1$$

for t = 1, 2 & 3. $(\because$ total no. of observation $= 3N)$

We have

- Included two period dummies in addition to the intercept. B'cor its good idea to allow a separate intercept for each time period, especially when we have a small no. of them.
- Basic period is t=1
- Intercept for 2nd time period is $\delta_1 + \delta_2$ & so on.
- Primarily interested in $\beta_1, \beta_2 \& \beta_k$
- If unobserved effect ai is correlated with any of the explanatory variable then use pooled OLS on the 3 years of dated results in biased and inconsistent estimates.
- Idiosyncratic error are uncorrelated & explanation variables.
- Cov (X_{itj}, $\mu_{is}) = 0$ for all t, $s \& j$ ……………2
- Explanatory variable are strictly exogenous after we take out the unobserved effect, a_i.

Eq. (2) Rules out cases where future explanatory variables react to current changes i.d the idiosyncratic error. As must be the case of X_{itj} is a lagged dependent variable.

- If ai is correlated with X_{itj} then X_{itj} will be correlated with the composite error, $V_{it} = a_i + u_{it}$ (under 2).
- We can eliminate a_i by differencing adjacent periorts.
- In T=3 case, we subtract time period one from time period two and time period two from time period three.

$$\Delta y_{it} = \delta_2 \Delta d2_t + \delta_3 \Delta d3_t + \beta_1 \Delta x_{it} + \ldots + \beta k \Delta x_{itk} + \Delta v_{it} \ldots\ldots (3)$$

- We have no. differenced eq. for t=1
- 3 eq. represents two time period for each individual in the sample.
- If this eq. satisfies the classical linear model assumptions the pooled OLS gives unbiased estimators.
- The important requirement for OLS to be consistent is that $\Delta \mu_{it}$ is uncorrelated with ΔX_{itj} for all J & t = 2 & 3.
- This is the natural extension for 2TP cone
- d_{2t} & d_{3t} are year dummies

t = 2, $\Delta d2_t = 1 \& \Delta d3_t = C$ for t = 3, $\Delta d2_t = -1 \& \Delta d3_t = 1$

3 eq. does not contain any intercept.

It is better to estimate the 1st differenced eq. with an intercept and a single time period dummy, usually for the time period.

$$\Delta y_{it} = \alpha_o - \alpha_3 d3_t + \beta_1 \Delta x_{it1} + \beta_k \Delta X_{iti} + \beta_k \Delta X_{itk} + \Delta \mu_{it} \ldots\ldots\ldots (4)$$

For t = 2 & 3

- β_j are identical in either formulation.
- With more than 3 period, things are similar.
- If we have same T time period for each of N cross – sectional units, we say at the data set is balanced panel.

- T is small relative to N, we should include a dummy variable for each TP to a/c per secular changes that are not being modeled.
- After 1st differencing, the eq. both takes.

$$\Delta y_{it} = \alpha_o + \alpha_3 d3_t + \alpha_4 d4_t + \ldots\ldots + \alpha_t dT_t + \beta_1 \Delta x_{it1} + \ldots + \beta_k \Delta X_{itk} + \Delta u_{it} \quad t = 2, 3 \ldots T \ldots (5)$$

- We have T-1 J.P on each unit i for the 1st differential eq.
- The total no. of observation is N(T-1).
- Its is simple to estimate eq (5) by pooled OLS, provided the observations have been properly organized and the differencing carefully done.
- When doing 2TP, we assure that Δu_{it} is uncorrelated over time for usual standard errors and test statisticatic to be valid.

1.7 SUMMARY AND CONCLUSIONS:

Since panel data relate to individuals, firms, states, countries, etc., over time, there is bound to be heterogeneity in these units. The techniques of panel data estimation can take such heterogeneity explicitly into account by allowing for individual-specific variables, as we shall show shortly. We use the term *individual* in a generic sense to include microunits such as individuals, firms, states, and countries.

By combining time series of cross-section observations, panel data give "more informative data, more variability, less collinearity among variables, more degrees of freedom and more efficiency." By studying the repeated cross section of observations, panel data are better suited to study the *dynamics of change.* Spells of unemployment, job turnover, and labor mobility are better studied with panel data. Panel data can better detect and measure effects that simply cannot be observed in pure cross-section or pure time series data. For example, the effects of minimum wage laws on employment and earnings can be better studied if we include successive waves of minimum wage increases in the federal and/or state minimum wages. Panel data enables us to study more complicated behavioral models.

For example, phenomena such as economies of scale and technological change can be better handled by panel data than by pure cross-section or pure time series data. By making data available for several thousand units, panel data can minimize the bias that might result if we aggregate individuals or firms into broad aggregates.

In short, panel data can enrich empirical analysis in ways that may not be possible if we use only cross-section or time series data. This is not to suggest that there are no problems with panel data modeling.

1.8 LETS SUM IT UP:

In the conclusion we can say that Panel regression models are based on panel data. Panel data consist of observations on the same cross-sectional, or individual, units over several time periods.There are several advantages to using panel data. *First,* they increase the sample size considerably. *Second,* by studying repeated cross-section observations, panel data are better suited to study the dynamics of change. *Third,* panel data enable us to study more complicated behavioral models.Despite their substantial advantages, panel data pose several estimation and inference problems. Since such data involve both cross-section and time dimensions, problems that plague cross-sectional data (e.g., heteroscedasticity) and time series data (e.g., autocorrelation) need to addressed. There are some additional problems, such as cross-correlation in individual units at the same point in time.

1.9 EXCERCISES:

Q1. Explain chow test.

Q2. What do you mean by panel data?

Q3. Distinguish between panel data and pooling data?

Q4. What are the various types of data?

Q5. Give any three examples of panel data?

Q6. Distinguish between balanced and unbalanced panel?

Q7. What are the special features of (**a**) cross-section data, (**b**) time series data, and (**c**) panel data?

1.10 Suggested Reading / References:

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2^{nd} Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

Lesson-2

FIXED EFFECT OR FIRST DIFFERENCING AND RANDOM EFFECT APPROACH

STRUCTURE

2.1 INTRODUCTION:

The basic framework for this discussion is a regression model of the form

$$y_{it} = \mathbf{x}_{it}' \boldsymbol{\beta} + \mathbf{z}_i' \boldsymbol{\alpha} + \varepsilon_{it}. \quad \textbf{(1)}$$

There are K regressors in $\mathbf{x}_{it}$, *not including a constant term*. The **heterogeneity,** or **individual effect** is $\mathbf{z}_i' \boldsymbol{\alpha}$ where $\mathbf{z}_i$ contains a constant term and a set of individual orgroup specific variables, which may be observed, such as race, sex, location, and so on or unobserved, such as family specific characteristics, individual heterogeneity in skill or preferences, and so on, all of which are taken to be constant over time t. As it stands, this model is a classical regression model. If $\mathbf{z}_i$ is observed for all individuals, then the entire model can be treated as an ordinary linear model and fit by least squares. The various cases we will consider are:

1. Pooled Regression: If $\mathbf{z}_i$ contains only a constant term, then ordinary least squares provides consistent and efficient estimates of the common α and the slope vector $\boldsymbol{\beta}$.

2. Fixed Effects: If $\mathbf{z}_i$ is unobserved, but correlated with $\mathbf{x}_{it}$, then the least squares estimator of $\boldsymbol{\beta}$ is biased and inconsistent as a consequence of an omitted variable. However, in this instance, the model

$$y_{it} = \mathbf{x}_{it}' \boldsymbol{\beta} + \alpha_i + \varepsilon_{it},$$

where $\alpha_i = \mathbf{z}_i \boldsymbol{\alpha}$, embodies all the observable effects and specifies an estimable conditional mean. This **fixed effects** approach takes α_i to be a group-specific constant term in the regression model. It should be noted that the term "fixed" as used here indicates that the term does not vary over time, not that it is nonstochastic, which need not be the case.

3. Random Effects: If the unobserved individual **heterogeneity,** however formulated, can be assumed to be uncorrelated with the included variables, then the model may be formulated as

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + E[\mathbf{z}_i\boldsymbol{\alpha}] + \mathbf{z}_i\boldsymbol{\alpha} - E[\mathbf{z}_i\boldsymbol{\alpha}] + \varepsilon_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \alpha + u_i + \varepsilon_{it},$$

that is, as a linear regression model with a compound disturbance that may be consistently, albeit inefficiently, estimated by least squares. This **random effects** approach specifies that u_i is a group specific random element, similar to ε_{it} except that for each group, there is but a single draw that enters the regression identically in each period.

Again, the crucial distinction between these two cases is whether the unobserved individual effect embodies elements that are correlated with the regressors in the model, not whether these effects are stochastic or not.We will examine this basic formulation, then consider an extension to a dynamic model.

2.2 OBJECTIVES:

1. To understand the Fixed Effect Technique.

2. To get familiar with the Random Effect Approach.

2.3 FIXED EFFECT OR FIRST DIFFERENCING:

First Differencing (FD) is just one of the many ways to eliminate the Fixed Effect (FE), a_i. An alternative method, which works better under certain assumptions, is called FE transformation.

Setting aside pooled OLS, we have, seen the competing methods for estimating unobserved effect models.

1. Involves differencing the dates.

2. Other involves time demeaning.

Under following cases we will find which one to use:

Case 1: When T=2, the FE & FD estimates, as well as all test statistically are identical in this case no matter which one to use.

Between FE and FD we require the we estimate the same model in each case.

Natural to include an intercept in the FD eq. C actually the intercept for 2nd time period in the original model writers for 2 Time period.

FE must include a Dummy variable for the 2nd Time period in order to be identical to the FD estimates that include an intercept.

FD with T=2, has the advantage of being straight – forward in any econometrics and statistical purchase.

Case-2: $T \geq 3 \& FE \& FD$ **estimates are not same.**

Both are unbiased.

We can’t use unbiased as a criterion under Assumption.

Both are consistent. (With T fixed as N - ∞)

For large N and small T, the choose between FE & FD hinges on the relative efficiency of the estimator and determined by the serial correlation in the idiosyncratic error, u_{it}.

When u_{it} are serially uncorrelated FE is more efficient then FD.

If U_{it} follows a random walk i.e. there is very substantial, positive serial correlation their the difference ΔU_{it} is serially uncorrelation, FD is better.

Case 3: When T is large and especially N is not very large, (Ex T = 30 & N = 2).

FD & FE estimators can be very sensitive to clinical measurement error in one or more explonationary variable.

Important: FD does't depend upon T, while that for the F&E estimator tends to zero at the rate 1/T.

It is difficult to choose b/w FE and FD. When they give subsequently different results.

2.4 RANDOM EFFECT APPROACH:

The fixed effects model allows the unobserved individual effects to be correlated with the included variables.We then modeled the differences between units strictly as parametric shifts of the regression function. This model might be viewed as applying only to the cross-sectional units in the study, not to additional ones outside the sample. For example, an intercountry comparison may well include the full set of countries for which it is reasonable to assume that the model is constant. If the individual effects are strictly uncorrelated with the regressors, then it might be appropriate to model the individual specific constant terms as randomly distributed across cross-sectional units. This view would be appropriate if we believed that sampled cross-sectional units were drawn from a large population.

2.4.1 UNBALANCED PANEL :

Some panel data sets, especially on individuals or firms however missing years, for it least same errors sectional units in the sample. Unbal-pannel.

We Assume the

$Y_{it} = \beta_{1i} + \beta_2 X_{2it} + \beta_3 X_{3it} + u_{it} \ldots\ldots\ldots\ldots (1)$

B_{li} = intestinal of treating as fixed we treat it as random variable.

$$\therefore \beta_{1i} = \beta_1 + \epsilon_i \quad i = 1, 2, \ldots\ldots\ldots N \ldots\ldots\ldots (2)$$

(Independent of all variable) ($\in$ is a random error $\in$ = O, var $\in$ = σ_E^2

Subtracting (2) in (1) we get

$$Y_{it} = \beta_1 + \beta_2 x2_{it} + \beta_3 x3_{it} + \varepsilon_2 + u_{it}$$
$$= \beta_1 + \beta_2 x2_{it} + \beta_3 x3_{it} + ul_{it} \ldots\ldots\ldots (3)$$

where $W_{it} = \in_i + U_{it}$

2.4.2 Some of the Assumption mode by Random effect model or error component Model:

1. $\in_i \sim N(0, \sigma_E^2)$

2. $U_{it} \sim N(0, \sigma_U^2)$

3. $E(\in_1 \in_{it}) = 0$

(No … between E_i & U_{it} if it will be there middle will be heteroshadastic)

4. $E(\in_1 \in_j) = 0 \ (i \neq j)$

5. $E(W_{it}) = 0$

6. $\text{Var}(W_{it})\ \sigma_E^2 + \sigma_U^2$.

Test for Fixed Effect Model is Hausman Test

Some time its assumption gets violated so it is not a good model to

study.

2.5 SUMMARY AND CONCLUSIONS:

Panel data pose several estimation and inference problems. Since such data involve both cross-section and time dimensions, problems that plague cross-sectional data (e.g., heteroscedasticity) and time series data (e.g., autocorrelation) need to addressed. There are some additional problems, such as cross-correlation in individual units at the same point in time. There are several estimation techniques to address one or more of these problems. The two most prominent are (1) the fixed effects model (FEM) and (2) the random effects model (REM) or error components model (ECM). In FEM the intercept in the regression model is allowed to differ among individuals in recognition of the fact each individual, or crosssectional, unit may have some special characteristics of its own. To take into account the differing intercepts, one can use dummy variables. The FEM using dummy variables is known as the least-squares dummy variable (LSDV) model. FEM is appropriate in situations where the individualspecific intercept may be correlated with one or more regressors. A disadvantage of LSDV is that it consumes a lot of degrees of freedom when the number of cross-sectional units, *N*, is very large, in which case we will have to introduce *N* dummies (but suppress the common intercept term). An alternative to FEM is ECM. In ECM it is assumed that the intercept of an individual unit is a random drawing from a much larger population with a constant mean value. The individual intercept is then expressed as a deviation from this constant mean value. One advantage of ECM over FEM is that it is economical in degrees of freedom, as we do not have to estimate *N* cross-sectional intercepts.We need only to estimate the mean value of the intercept and its variance.ECMis appropriate in situations where the (random) intercept of each cross-sectional unit is uncorrelated with the regressors.

2.6 LETS SUM IT UP:

The preceding has shown a few of the extensions of the classical model that can be obtained when panel data are available. In principle, any of the models we have examined before this lesson and all those we will consider later, including the multiple equation models, can be extended in the same way. The main advantage, as we noted at the outset, is that with panel data, one can formally model the heterogeneity across groups that is typical in microeconomic data.

2.7 EXCERCISES:

Q1. What do you mean by First Differencing?

Q2. Explain the Error Components Model(ECM)?

Q3. What are the various problems which we face in dealing with panel data?

Q4. Describe the various techniques for the estimation of panel data?

Q5. W hat do you mean by unbalanced panel?

2.8 Suggested Reading / References:

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2nd Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

Lesson-3

ESTIMATION OF PANEL DATA REGRESSION MODELS

STRUCTURE

3.6 LETS SUM IT UP

3.7 EXCERCISES

3.8 SUGGESTED READING / REFERENCES

3.1 INTRODUCTION:

The basic framework for this discussion is a regression model of the form

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \mathbf{z}_i\boldsymbol{\alpha} + \varepsilon_{it}.$$

Fixed Effects: If $\mathbf{z}_i$ is unobserved, but correlated with $\mathbf{x}_{it}$, then the least squares estimator of $\boldsymbol{\beta}$ is biased and inconsistent as a consequence of an omitted variable. However, in this instance, the model

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \alpha_i + \varepsilon_{it},$$

where $\alpha_i = \mathbf{z}_i\boldsymbol{\alpha}$, embodies all the observable effects and specifies an estimable conditional mean. This **fixed effects** approach takes α_i to be a group-specific constant term in the regression model. It should be noted that the term "fixed" as used here indicates that the term does not vary over time, not that it is nonstochastic, which need not be the case.

3.2 OBJECTIVES:

1. To understand the Fixed Effects Approach.

2. Testing the significance of the group effects.

3. To understand the fixed time and group effects.

4. To understand the relationship between unbalanced panels and

fixed effects

3.3 THE FIXED EFFECTS APPROACH

$$Y_{it} = \beta_{1i} + \beta_2 X_{2it} + \beta_3 X_{3it} + u_{it} \qquad \textbf{(1)}$$

Estimation of (1) depends on the assumptions we make about the intercept, the slope coefficients, and the error term, u_{it}. There are several possibilities:

1. Assume that the intercept and slope coefficients are constant across time and space and the error term captures differences over time and individuals.

2. The slope coefficients are constant but the intercept varies over individuals.

3. The slope coefficients are constant but the intercept varies over individuals and time.

4. All coefficients (the intercept as well as slope coefficients) vary over individuals.

5. The intercept as well as slope coefficients vary over individuals and time.

3.3.1. All Coefficients Constant across Time and Individuals

The simplest, and possibly naive, approach is to disregard the space and time dimensions of the pooled data and just estimate the usual OLS regression.

3.3.2. Slope Coefficients Constant but the Intercept Varies across Individuals: The Fixed Effects or Least-Squares Dummy Variable (LSDV) Regression Model

One way to take into account the "individuality" of each company or each cross-sectional unit is to let the intercept vary for each company but still assume that the slope coefficients are constant across firms.

$$Y_{it} = \beta_{1i} + \beta_2 X_{2it} + \beta_3 X_{3it} + u_{it} \qquad \textbf{(2)}$$

Notice that we have put the subscript i on the intercept term to suggest that the intercepts of the four firms may be different; the differences may be due to special features of each company, such as managerial style or managerial philosophy.

In the literature, model (2) is known as the **fixed effects** (regression) model (**FEM**). The term "fixed effects" is due to the fact that, although the intercept may differ across individuals (here the four companies), each individual's intercept does not vary over time; that is, it is *time invariant*. Notice that if we were to write the intercept as β_{1it} , it will suggest that the intercept of each company or individual is *time variant.* It may be noted that the FEM given in (2) assumes that the (slope) coefficients of the regressors do not vary across individuals or over time.

How do we actually allow for the (fixed effect) intercept to vary between companies? We can easily do that by the dummy variable technique that we learned in Lesson 9, particularly, the **differential intercept dummies.** Therefore, we write (16.3.2) as:

$$Yi_t = \alpha_1 + \alpha_2 D_{2i} + \alpha_3 D_{3i} + \alpha_4 D_{4i} + \beta_2 X_{2it} + \beta_3 X_{3it} + u_i t \qquad \textbf{(3)}$$

where D_{2i} = 1 if the observation belongs to GM, 0 otherwise; D_{3i} = 1 if the observation belongs to US, 0 otherwise; and D_{4i} = 1 if the observation belongs to WEST, 0 otherwise. Since we have four companies, we have used only three dummies to avoid falling into the **dummy-variable trap** (i.e., the situation of perfect collinearity). Here there is no dummy for GE. In other words, α_1 represents the intercept of GE and α_2, α_3, and α_4, the *differential intercept* coefficients, tell by how much the intercepts of GM, US, andWESTdiffer from the intercept of GE. In short, GE becomes the comparison company. Of course, you are free to choose any company as the comparison company.

Incidentally, if you want explicit intercept values for each company, you can introduce four dummy variables provided you run your regression through the origin, that is, drop the common intercept in (3); if you donot do this, you will fall into the dummy variable trap.

Since we are using dummies to estimate the fixed effects, in the literature the model (3) is also known as the **least-squares dummy variable (LSDV) model.** So, the terms fixed effects and

LSDV can be used inter changeably. In passing, note that the LSDV model (3) is also known as the **covariance model** and X_2 and X_3 are known as *covariates.*

$$Y_{it} = \lambda_0 + \lambda_1 \text{Dum35} + \lambda_2 \text{Dum36} + \cdots + \lambda_{19} \text{Dum53} + \beta_2 X_{2it} + \beta_3 X_{3it} + u_{it} \quad .4$$

3.3.3 Slope Coefficients Constant but the Intercept Varies over Individuals As Well As Time

To consider this possibility, we can combine (3) and (4), as follows:

$$Y_{it} = \alpha_1 + \alpha_2 D_{\text{GM}i} + \alpha_3 D_{\text{US}i} + \alpha_4 D_{\text{WEST}i} + \lambda_0 + \lambda_1 \text{Dum}_{35} + \cdots$$

$$+\lambda_{19}\text{Dum}_{53} + \beta_2 X_{2i} + \beta_3 X_{3i} + u_{it} \qquad \textbf{(5)}$$

When we run this regression, we find the company dummies as well as the coefficients of the *X* are individually statistically significant, but none of the time dummies are.

The overall conclusion that emerges is that perhaps there is pronounced individual company effect but no time effect. In other words, the investment functions for the four companies are the same except for their intercepts. In all the cases we have considered, the *X* variables had a strong impact on *Y.*

3.3.4. All Coefficients Vary across Individuals

Here we assume that the intercepts and the slope coefficients are different for all individual, or cross-section, units. This is to say that the investment functions of GE, GM, US, and WEST are all different. We can easily extend our LSDV model to take care of this situation.

3.4 TESTING THE SIGNIFICANCE OF THE GROUP EFFECTS:

The *t* ratio for *ai* can be used for a test of the hypothesis that αi equals zero. This hypothesis about one specific group, however, is typically not useful for testing in this regression context. If we are interested in differences across groups, then we can test the hypothesis that the constant terms are all equal with an F test. Under the null hypothesis of equality, the efficient estimator is pooled least squares. The F ratio used for this test is

$$F(n-1, nT-n-K) = R2\ LSDV - R2\ Pooled/(n-1)/\ 1 - R2\ LSDV/(nT-n-K)$$

where *LSDV* indicates the dummy variable model and *Pooled* indicates the pooled or restricted model with only a single overall constant term. Alternatively, the model may have been estimated with an overall constant and $n-1$ dummy variables instead. All other results (i.e., the least squares slopes, $s2$, $R2$) will be unchanged, but rather than estimate αi , each dummy variable coefficient will now be an estimate of $\alpha i - \alpha 1$

where group “1” is the omitted group. The F test that the coefficients on these n – 1 dummy variables are zero is identical to the one above. It is important to keep in mind, however, that although the statistical results are the same, the interpretation of the dummy variable coefficients in the two formulations is different.

3.5 SUMMARY AND CONCLUSIONS:

Although easy to use, the LSDV model has some problems that need to be borne in mind.

First, if you introduce too many dummy variables, you will run up against the degrees of freedom problem. Suppose , we have 80 observations, but only 55 degrees of freedom—we lose 3 df for the three company dummies, 19 df for the 19 year dummies,

2 for the two slope coefficients, and 1 for the common intercept.

Second, with so many variables in the model, there is always the possibility of multicollinearity, which might make precise estimation of one or more parameters difficult.

Third, suppose in the FEM (16.3.1) we also include variables such as sex, color, and ethnicity, which are time invariant too because an individual's sex color, or ethnicity does not change over time. Hence, the LSDV approach may not be able to identify the impact of such time-invariant variables.

Fourth, we have to think carefully about the error term *uit*. All the results we have presented so far are based on the assumption that the error term follows the classical assumptions, namely, $uit \sim N(0, \sigma 2)$. Since the *i* index refers to cross-sectional observations and *t* to time series observations, the classical assumption for *uit* may have to be modified. There are several possibilities.

1. We can assume that the error variance is the same for all crosssection units or we can assume that the error variance is heteroscedastic.

2. For each individual we can assume that there is no autocorrelation over time. Thus, for example, we can assume that the error term of the investment function for General Motors is nonautocorrelated. Or we could assume that it is autocorrelated, say, of the AR(1) type.

3. For a given time, it is possible that the error term for General Motors is correlated with the error term for, say, U.S. Steel or both U.S. Steel and Westinghouse.7 Or, we could assume that there is no such correlation.

4. We can think of other permutations and combinations of the error term. As you will quickly realize, allowing for one or more of these possibilities will make the analysis that much more complicated. Space and mathematical demands preclude us from considering all the possibilities. A somewhat accessible discussion of the various possibilities can be found in Dielman, Sayrs, and Kmenta.8 However, some of the problems *may* be alleviated if we resort to the so-called **random effects model,** which we discuss next.

3.6 LETS SUM IT UP:

In last we conclude that in FEM the intercept in the regression model is allowed to differ among individuals in recognition of the fact each individual, or crosssectional, unit may have some special characteristics of its own. To take into account the differing intercepts, one can use dummy variables. The FEM using dummy variables is known as the least-squares dummy variable (LSDV) model. FEM is appropriate in situations where the individualspecific intercept may be correlated with one or more regressors. A disadvantage of LSDV is that it consumes a lot of degrees of freedom when the number of cross-sectional units, *N*, is very large, in which case we will have to introduce *N* dummies (but suppress the common intercept term).

3.7 EXCERCISES:

Q.1. Discuss the test and other factors which decide whether to go for OLS or fixed effect or random effect.

Q.2. What is meant by a fixed effects model (FEM)? Since panel data have both time and space dimensions, how does FEM allow for both dimensions?

Q 3. When is fixed effect is preferred over random effect?

Q4. What is the significance of the group effects in FEM?

3.8 Suggested Reading / References:

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2nd Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.

Lesson-4

RANDOM EFFECT MODEL

STRUCTURE

4.1 INTRODUCTION:

Random Effects: If the unobserved individual **heterogeneity,** however formulated, can be assumed to be uncorrelated with the included variables, then the model may be formulated as

$$y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + E[\mathbf{z}_i\boldsymbol{\alpha}] + \mathbf{z}_i\boldsymbol{\alpha} - E[\mathbf{z}_i\boldsymbol{\alpha}] + \varepsilon_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \alpha + u_i + \varepsilon_{it},$$

that is, as a linear regression model with a compound disturbance that may be consistently, albeit inefficiently, estimated by least squares. This **random effects** approach specifies that u_i is a group specific random element, similar to ε_{it} except that for each group, there is but a single draw that enters the regression identically in each period.

Again, the crucial distinction between these two cases is whether the unobserved individual effect embodies elements that are correlated with the regressors in the model, not whether these effects are stochastic or not.We will examine this basic formulation, then consider an extension to a dynamic model.

Random Parameters: The random effects model can be viewed as a regression model with a random constant term. With a sufficiently rich data set, we may extend this idea to a model in which the other coefficients vary randomly across individuals as well. The extension of the model might appear as

$$y_{it} = \mathbf{x}_{it}(\boldsymbol{\beta} + \boldsymbol{h}_i) + (\alpha + u_i) + \varepsilon_{it},$$

where $\boldsymbol{h}_i$ is a random vector which induces the variation of the parameters across individuals. This random parameters model was proposed quite early in this literature, but has only fairly recently enjoyed widespread attention in several fields. It represents a natural extension in which researchers broaden the amount of heterogeneity across individuals while retaining some commonalities—the parameter vectors still share a common mean. Some recent applications have extended this yet another step by allowing the mean value of the parameter distribution to be person-specific, as in

$$y_{it} = \mathbf{x}_{it}(\boldsymbol{\beta} + \mathbf{z}_i + \boldsymbol{h}_i) + (\alpha + u_i) + \varepsilon_{it},$$

where $\mathbf{z}_i$ is a set of observable, person specific variables, and_is a matrix of parameters to be estimated. As we will examine later, this **hierarchical model** is extremely versatile.

4.2 OBJECTIVES:

1. To understand the Random Effect Model.

2. Distinguish between the fixed effect and random effect model.

3. To understand the Hausman Specification Test for the random effects model.

4.3 RANDOM EFFECTS MODEL:

Although straightforward to apply, fixed effects, or LSDV, modeling can be expensive in terms of degrees of freedom if we have several cross-sectional units. If the dummy variables do in fact represent a lack of knowledge about the (true) model, why not express this ignorance through the disturbance term u_{it}? This is precisely the approach suggested by the proponents of the socalled **error components model (ECM) or random effects model (REM).** The basic idea is to start with:

$$Y_{it} = \beta_{1i} + \beta_2 X_{2it} + \beta_3 X_{3it} + u_{it} \quad \textbf{(1)}$$

Instead of treating β_{1i} as fixed, we assume that it is a random variable with a mean value of β_1 (no subscript i here). And the intercept value for an individual company can be expressed as

$$\beta_{1i} = \beta_1 + \varepsilon_i \quad i = 1, 2, \ldots, N \quad \textbf{(2)}$$

where ε_i is a random error term with a mean value of zero and variance of σ^2_ε. What we are essentially saying is that the four firms included in our sample are a drawing from a much larger universe of such companies and that they have a common mean value for the intercept ($= \beta_1$) and the individual differences in the intercept values of each company are reflected in the error term ε_i.

Substituting (2) into (1), we obtain:

$$Y_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + \varepsilon_i + u_{it} = \beta_1 + \beta_2 X_{2it} + \beta_3 X_{3it} + w_{it} \quad \textbf{(3)}$$

where

$$w_{it} = \varepsilon_i + u_{it} \quad \textbf{(4)}$$

The composite error term w_{it} consists of two components, ε_i, which is the cross-section, or individual-specific, error component, and u_{it}, which is the combined time series and cross-section error component. The term *error components model* derives its name because the composite error term w_{it} consists of two (or more) error components.

The usual assumptions made by ECM are that

$$\varepsilon_i \sim N(0, \sigma^2_\varepsilon) \quad u_{it} \sim N(0, \sigma^2_u) \quad \textbf{(5)}$$

$$E(\varepsilon_i u_{it}) = 0 \quad E(\varepsilon_i \varepsilon_j) = 0 \quad (i = j)$$

$$E(u_{it} u_{is}) = E(u_{it} u_{jt}) = E(u_{it} u_{js}) = 0 \quad (i = j;\ t = s).$$

that is, the individual error components are not correlated with each other and are not autocorrelated across both cross-section and time series units.

Notice carefully the difference between FEM and ECM. In FEM each cross-sectional unit has its own (fixed) intercept value, in all N such values for N cross-sectional units. In ECM, on the other hand, the intercept β_1 represents the mean value of all the (cross-sectional) intercepts and the error component ε_i represents the (random) deviation of individual intercept from this mean value. However, keep in mind that ε_i is not directly observable; it is what is known as an **unobservable,** or **latent, variable.** As a result of the assumptions stated in (5), it follows that

$E(w_{it}) = 0$ **(6)**

$\text{var}(w_{it}) = \sigma^2_\varepsilon + \sigma^2_u$ **(7)**

Now if $\sigma^2_\varepsilon = 0$, there is no difference between models (1) and (3), in which case we can simply pool all the (cross-sectional and time series) observations and just run the pooled regression, as we did in (1). As (7) shows, the error term w_{it} is homoscedastic. However, it can be shown that w_{it} and w_{is} ($t = s$) are correlated; that is, the error terms of a given cross-sectional unit at two different points in time are correlated. The correlation coefficient, corr (w_{it}, w_{is}), is as follows:

$\text{corr}(w_{it}, w_{is}) = \sigma^2_\varepsilon \ \sigma^2_\varepsilon + \sigma^2_u$ **(8)**

Notice two special features of the preceding correlation coefficient. *First,* for any given cross-sectional unit, the value of the correlation between error terms at two different times remains the same no matter how far apart the two time periods are, as is clear from (8). This is in strong contrast to the first-order [AR(1)] scheme, where we

found that the correlation between time periods declines over time. *Second,* the correlation structure given in (8) remains the same for all crosssectional units; that is, it is identical for all individuals.

4.4 FIXED EFFECT OR RANDOM EFFECT MODEL:

Which model is better FEM or REM the answer things around the assumption one makes about the likely correlation between the individual error component E_i & the X regressors.

- If assume that E_i & the X's are uncorrelated REM.
- If assume that E_i & the X's are correlated REM.
- If T is large & the N is small. There is little diff in the value of parameters estimated by FEM and REM. Here FEM will be preferable coice here depend upon computationa; convenience.
- If T is small & N is large estimate from two will ne different

Under the ECM = $\beta_{1i} = \beta_1 + E_i$ where FEM are that β_{1i} as fixed & non random. In this case FEM.

- If the individual error component Ei 7 one or more regression are correlated then REM are baised estimates and FEM are unbiased.
- If N is large & T is small & if the assumption undertying REM held, REM estimator are more efficient the FEM estimator.
- Hausman Test will decide which one is good.

4.5 HAUSMAN SPECIFICATION TEST FOR THE RANDOM EFFECT MODEL:

$$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 I_t + \alpha_3 R_t + u_{1t} \ (DF) \ldots\ldots\ldots (1)$$

$$Q_t = \beta_0 + \beta_1 P_t + u_{2t} \ (SF) \ldots\ldots\ldots (2)$$

P=Price, Q = Queenly

I = Income, R = Wealth

U's = error term

Assume I & R are exogenous

& P & Q are endogenous

Under eq. (2), if there is no simultaneity prob., P_t & u_{2t} should be uncorrelated or will be correlated. To find the which case, the Hausman test proceeds as follows.

$$\alpha_1 P_t - \beta_1 P_t = \alpha_0 + \alpha_C I_t + \alpha_3 P_t + \mu_{1t} - \beta_0 - u_{2t}$$

$$(\alpha_1 - \beta_1) P_t = \alpha_0 - \beta_0 + \alpha_2 I_t + \alpha_3 R_t + u_{1t} - u_{2t}$$

$$P_t = \frac{\alpha_0 - \beta_0}{(\alpha_1 - \beta_1)} + \frac{\alpha_2 I_t}{(\alpha_1 - \beta_1)} + \frac{\alpha_3 R_y}{(\alpha_1 - \beta_1)} + \frac{u_t - u_{2t}}{(\alpha_1 - \beta_1)}$$

$$P_1 = \pi_0 + \pi_1 I_t + \pi_3 R_t + V_t \ldots\ldots\ldots\ldots(3)$$

$$Q_t = \beta_1 + \beta_1 \left(\frac{\alpha_0 - \beta_0}{\alpha_1 - \beta_1} \right) + \frac{\beta_1 \alpha_2 I_t}{\alpha_1 - \beta_1} - \frac{\beta_1 \alpha_3 R_t}{\alpha_1 - \beta}$$

$$+ \frac{\beta_1 (u_{1t} + u_{2t})}{\alpha_1 - \beta_1} + u_{2t}$$

$$Q_t = \pi_3 + \pi_4 I_t + \pi_S R_t + w_t \ldots\ldots\ldots\ldots\ldots(4)$$

((3) & (4) were ever identified eq.)

$$\hat{P}_t = \hat{\pi}_0 + \hat{\pi}_1 I_t + \hat{\pi}_2 R_t \qquad [By\,OLS] \ldots\ldots\ldots(5)$$

$$P_t = \hat{P}_t + \hat{V}_t \ldots\ldots\ldots\ldots(6)$$

$$Q_t = \hat{P}_t + V_t + W_t \ldots\ldots\ldots\ldots(7)$$

Putting 6 and 2

$$Q_t = \beta_0 + \beta_1(\hat{P}_t + \hat{V}_t) + \mu_{2t}$$

$$Q_t = \beta_0 + \beta_1\hat{P}_t + \beta_1\hat{V}_t + \mu_{2t} \ldots\ldots\ldots\ldots(8)$$

$(P_t \& V_t$ **are same**

Under the null hypothesis that is on simultaneity, the correlation between .. and u_{2t} should be zero, asymptotically. Thus, if we run the regression (7) is statistically zero, we can conclude that there is no simultaneity problem. Of course, this conclusion will be reversed if we find this coefficient to be statistically significant.

4.5.1 Hausman test involves the follows steps:

Step 1: Regress P_t on I_t and R_t to obtain.

Step 2: Regress Q_t on $\hat{P}_t \& \hat{V}_t$ and perform a t test on the coefficient of . If it is significant, do not reject the hypothesis of simultaneity; otherwise, reject it. For effective estimation, however, Pindyck and Rubinfeld suggest regressing Q_t on P_t and $\hat{V}_t$.

4.6 SUMMARY AND CONCLUSIONS:

1. Panel regression models are based on panel data. Panel data consist of observations on the same cross-sectional, or individual, units over several time periods.

2. There are several advantages to using panel data. *First,* they increase the sample size considerably. *Second,* by studying repeated cross-section observations, panel data are better suited to study the dynamics of change. *Third,* panel data enable us to study more complicated behavioral models.

3. Despite their substantial advantages, panel data pose several estimation and inference problems. Since such data involve both cross-section and time dimensions, problems that plague cross-sectional data (e.g., heteroscedasticity) and time series data (e.g., autocorrelation) need to addressed. There are some additional problems, such as cross-correlation in individual units at the same point in time.

4. There are several estimation techniques to address one or more of these problems. The two most prominent are (1) the fixed effects model (FEM) and (2) the random effects model (REM) or error components model (ECM).

5. In FEM the intercept in the regression model is allowed to differ among individuals in recognition of the fact each individual, or crosssectional, unit may have some special characteristics of its own. To take into account the differing intercepts, one can use dummy variables. The FEM using dummy variables is known as the least-squares dummy variable (LSDV) model. FEM is appropriate in situations where the individualspecific intercept may be correlated with one or more regressors. A disadvantage of LSDV is that it consumes a lot of degrees of freedom when the number of cross-sectional units, *N*, is very large, in which case we will have to introduce *N* dummies (but suppress the common intercept term).

6. An alternative to FEM is ECM. In ECM it is assumed that the intercept of an individual unit is a random drawing from a much larger population with a constant mean value. The individual intercept is then expressed as a deviation from this constant mean value. One advantage of ECM over FEM is that it is economical in degrees of freedom, as we do not have to estimate *N* cross-sectional intercepts.We need only to estimate the mean value of the intercept and its variance.ECMis appropriate in situations where the (random) intercept of each cross-sectional unit is uncorrelated with the regressors.

7. The Hausman test can be used to decide between FEM and ECM.

8. Despite its increasing popularity in applied research, and despite increasing availability of such data, panel data regressions may not be appropriate in every situation. One has to use some practical judgment in each case

4.7 LETS SUM IT UP:

The preceding section has shown a few of the extensions of the classical model that can be obtained when panel data are available. In principle, any of the models we have examined before this lesson including the multiple equation models, can be extended in the same way. The main advantage, as we noted at the outset, is that with panel data, one can formally model the heterogeneity across groups that is typical in microeconomic data.To some extent the model of heterogeneity can be misleading. What might have appeared at one level to be differences in the variances of the disturbances across groups may well be due to heterogeneity of a different sort, associated with the coefficient vectors. We also examined some additional models for disturbance processes that arise naturally in a multiple equations context but are actually more general cases of some of the models we looked at above, such as the model of groupwise heteroscedasticity.

4.8 EXCERCISES:

Q1. What is latent variable?

Q2. Discuss Lagrange multiplier test.

Q3. What is meant by an error components model (ECM)? How does it differ from FEM? When is ECM appropriate? And when is FEM appropriate?

Q4. Is there a difference in FEM, least-squares dummy variable (LSDV) model, and covariance model?

Q5. When are panel data regression models inappropriate? Give examples.

Q.6 Which is a better model, FEM or ECM? Justify your answer

3.8 Suggested Reading / References:

1. Baltagi, B.H.(1998). Econometrics, Springer, New York.

2. Chow,G.C.(1983). Econometrics, McGraw Hill, New York.

3. Goldberger, A.S.(1998). Introductory Econometrics, Harvard University Press, Cambridge, Mass.

4. Green, W.(2000). Econometrics, Prentice Hall of India, New Delhi.

5. Gujarati, D.N.(1995). Basic Econometrics. McGraw Hill, New Delhi.

6. Koutsoyiannis,A.(1977). Theory of Econometrics(2nd Esdn.). The Macmillan Press Ltd. London.

7. Maddala, G.S.(1997). Econometrics, McGraw Hill; New York.