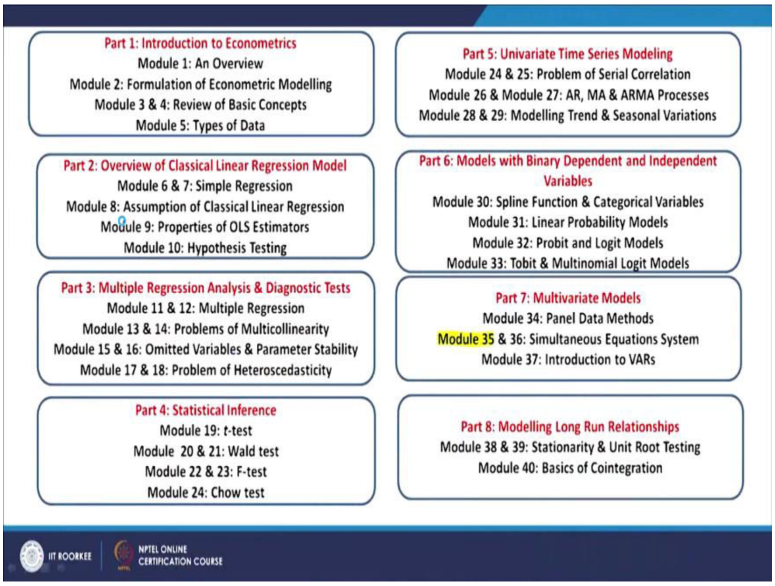
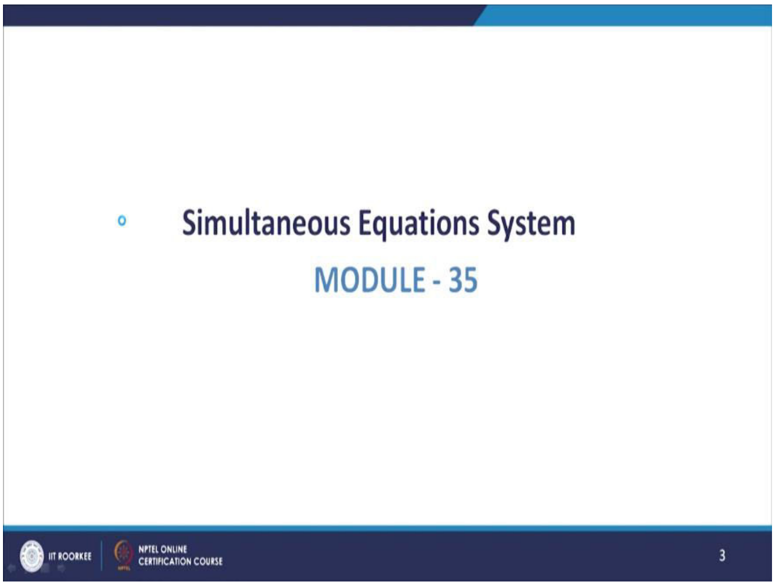
Econometric Modelling
Professor. Sujata Kar
Department of Management Studies
Indian Institute of Technology, Roorkee
Lecture No. 35
Stimulus Equation System-I

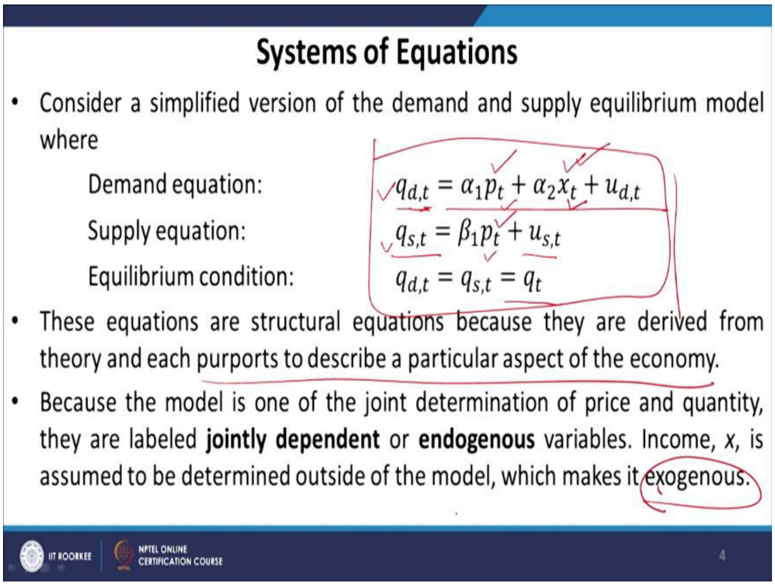
(Refer Slide Time: 0:33)





Hello, and welcome to the course on econometric modelling, this is Module 35. Module 35 is part of multivariate modeling. And in m 35 and 36 we are going to discuss the simultaneous equations system. So, Module 35 is the first part of some simultaneous equation system. So, first of all, we define what is our systems of equations.

(Refer Slide Time: 0:55)



Consider a simplified version of the demand and supply equilibrium model where we have a demand equation. So, here d refers to demand, and this actually with respect to a particular time period *(refer to slide time 0:55)*. So, demand for a particular product at time period t is expressed as a function of the price of that particular product, and another exogenous variable is Xt.

And similarly, we have a supply equation where we have qs,t denoting the supply of that output or product at time period t, which is again a function of Pt that is the price of the product and we do not assume any other exogenous variables, so we simply have the error terms. And finally, we have an equilibrium condition, which states that in equilibrium the quantity of demand will be equal to the quantity of supply.

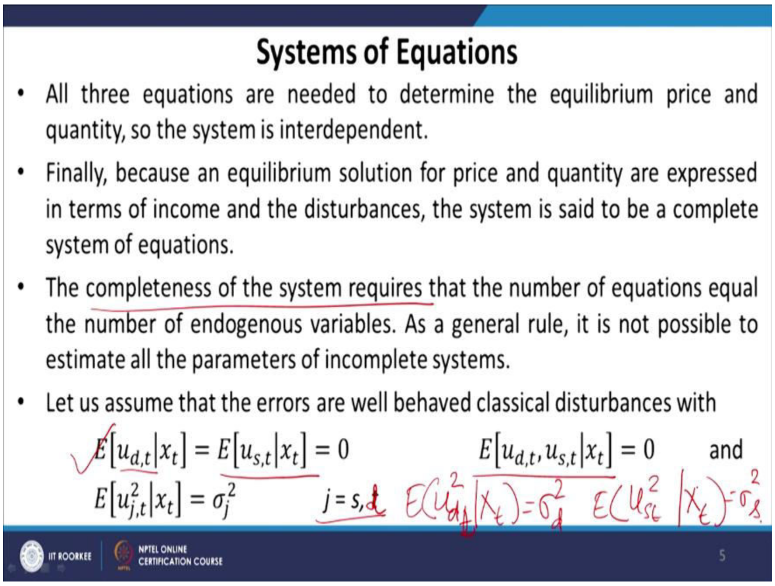
So, these equations are structural equations because they are derived from theory, and each purports to describe a particular aspect of the economy *(refer to slide time 0:55)*. So, this is a simple example. There could be many other examples where we have a structural equation, which is basically obtained from a set of theory or deduced from a theory. And then each equation, when he specifically refers to a situation where each purports to describe a particular aspect of the economy, a better example would be, a macroeconomic model.

The very simple macroeconomic model where probably we consider GDP equals consumption plus investment plus government expenditure, plus exports minus imports. And then we can have individual, we can have equations for individual components. Like one equation for consumption function, another equation for individual investment function, then the third equation for government expenditure, fourth for exports, and the fifth for imports. And then combining them together, we will be having structural equations where it is a complete model, and in each of these equations representing one sector or one aspect of the economy.

Because the model is one of the joint determinations of price and quantity in this example specifically, they are leveled jointly dependent or endogenous variables. So, here, in this case, we can see that both price and quantity both are appearing, and we actually determine both. Like this structural equation is actually designed to determine both the equilibrium price and the quantity. And that is why price and quantity are jointly dependent or endogenous variables *(refer to slide time 0:55).*

Now, x could be the income of individuals, so income is assumed to be determined outside the model, which makes it an exogenous variable. So, x is an exogenous variable q is in endogenous p is also endogenous.

(Refer Slide Time: 4:06)



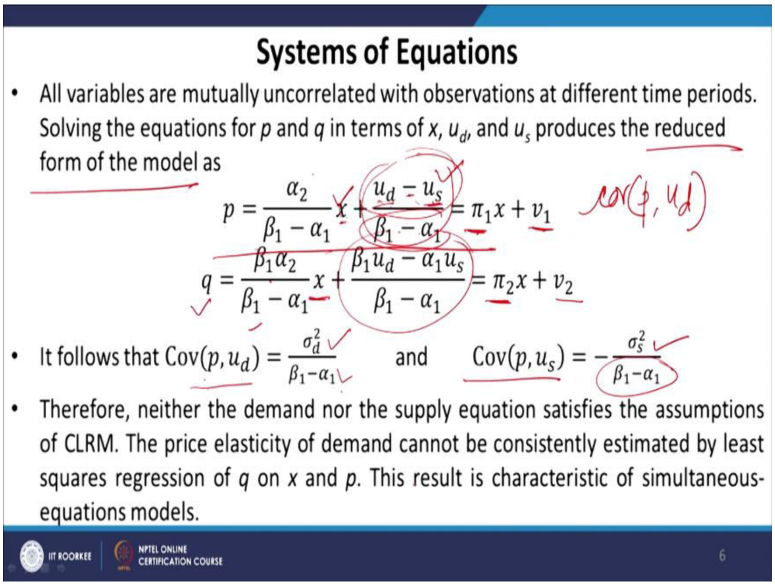
Now, all three equations are needed to determine the equilibrium price and quantity, so the system is actually interdependent. Finally, because an equilibrium solution for price and quantity is expressed in terms of income and the disturbances, the system is said to be a complete system of equations. The completeness of the system requires that the number of equations is equal to equal the number of endogenous variables. As a general rule, it is not possible to estimate all the parameters of incomplete systems.

So first of all, we will talk about the completeness of the system the first thing that we require is that, for each and every endogenous variable, there should be an equation. Basically, endogenous variables imply that this variable is explained by the system. So, unless and until we have an equation for one of the variables which we consider it to be endogenous, the variable actually becomes an exogenous variable. So, completeness implies that we specify the model in its entirety completely and that is why I include equations for each and every endogenous variable.

However, it is not possible to estimate all the parameters of incomplete systems. Now, let us assume that the errors are well behaved classical disturbances so that the expected value of the error term associated with the demand equation conditional upon the values of x is equal to 0 and so is the error term associated with the supply equation again conditional upon the values of x.

Now, we also assume the expected values of Edt, Est, which implies that the covariance between these two error terms is also 0 *(refer to slide time 04:06).* And finally, the expected value of uj square t conditional upon the values of Xt is sigma j square. So, here, this actually implies that the expected value of j is equal to s and d. So, this implies that the expected value of Ud square t conditional upon the values of Xt is equal to sigma square d, and the expected value of Ust square conditional upon the value of Xt is equal to sigma square s. So, this simply implies that the errors are heteroskedastic and errors are also not cross-correlated.

(Refer Slide Time: 6:46)



So, all variables are mutually uncorrelated with observations at different time periods. Solving the equations for p and q in terms of x, Ud and Us and us produce the reduced form of the model. So, what is the reduced form is the form where we have only the endogenous variables or one endogenous variable on the left-hand side, and we have only the exogenous variables and the error terms on the right-hand side. So, by doing this, we find that p can be expressed as a function of x and the error terms.

Suppose, we denote alpha 2 upon beta 1 minus alpha 1 by pi 1, and this entire component is denoted by v1. Similarly, q is also expressed as a function of income, and a combination of the error terms. And again, we denote beta 1, alpha 2 divided by beta 1 minus alpha 1 by pi 2 and this entire error term is denoted by v2. So, this is a reduced form, which has been obtained from the structure form, which is basically expert explaining or expressing all the endogenous variables in terms of only the exogenous variables and the disturbance terms*(refer to slide time 06:46).*

Now, it follows that, covariance between p and Ud will be sigma square d divided by beta 1 minus alpha 1. So, this happens because if I consider covariance between with p and Ud, so I am considering covariance between u and p and Ud. And what I am getting? First of all covariance between x and Ud is 0, so this term disappears and then taking out beta 1 minus alpha 1 constant I will be having covariance between p, Ud minus covariance between p, Us.

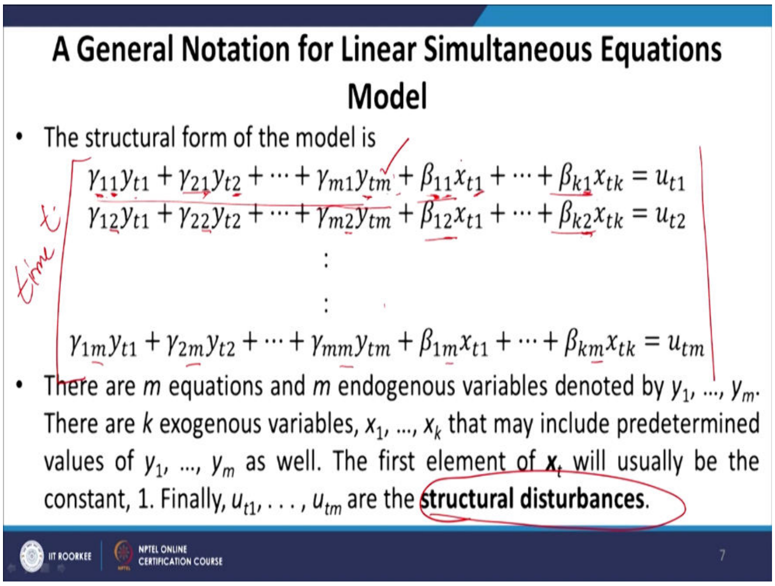
Now, by considering covariance between p and this component I will be having covariance between Ud, Ud, and covariance between Ud, Us. Now, the covariance between Ud, Us is 0 and covariance between Ud, Ud is actually variance of Ud. So, that is why we have sigma d square on the numerator and beta 1 minus alpha 1, as it is in the denominator.

In a similar fashion, we will be also deriving the covariance between p and Us as minus sigma s square, which is the variance of Us divided by beta 1 minus alpha 1. Therefore, neither the demand nor the supply equation satisfies the assumptions of CLRM, that is, a very basic assumption that the variables or the variables on the right-hand side should be uncorrelated with the error terms.

Now, if I go to the structural form, what we observe is that, p appears here on the right-hand side and Ud,t appears also here *(refer to slide time 06:46)*. Similarly, p appears here and Us,t appears here. We have just observed that Pt is correlated with Ud,t and Pt is also correlated with Us,d. As a result of which we actually cannot apply OLS on this set of equations because they do not fulfill the standard classical linear regression model assumptions, that is what it is mentioned here that neither the demand or the supply equation satisfies the assumptions of CLRM.

The price elasticity of demand, which is basically denoted by alpha 1 here and price elasticity of supply is beta 1 here, cannot be consistently estimated by least squares regressions of q on x and p four for the demand equation and q on p for the supply equation. This result is characteristic of the simultaneous equations model in that it does not satisfy the basic CLRM assumption.

(Refer Slide Time: 10:59)



Now, when we go for a generalized version of the simultaneous equations model or systems of equation *(refer to slide time 10:59)*. So, the structure form of the model is like this, where we can see that several endogenous variables are denoted by Y and we are having several endogenous variables. How many endogenous variables? So, there are m endogenous variables. And you can see that this actually refers to the time period t. So, the entire system of the equation has a reference to a particular time period which is time t.

Now, this is the first equation for the first variable, the first endogenous variable. Alternatively, we have said that if there are m endogenous variables and there should be at least m equations. So, that is why we have m equations having m endogenous variables, and then we have k exogenous variables and their associated parameters are beta 11 to beta k1 beta 12 to beta k2, and so on. On the right-hand side, we have the error terms associated with the individual equations observed for a particular time period.

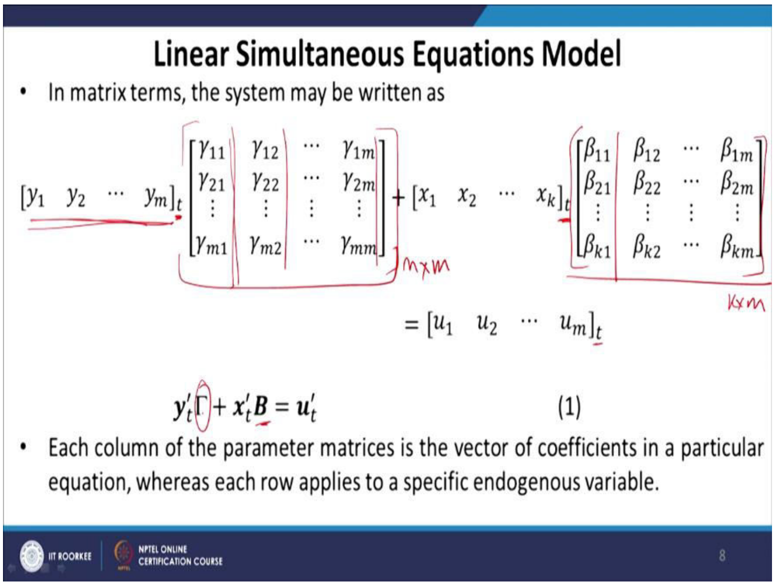
So, as you can see here gamma 11 basically refers to the second subscript refers to the equation, and the first subscript refers to the variable. So, gamma 11 refers to the first equation and the first variable gamma 21 refers to the second variable, but first equation, and similarly.

Similarly, we can also observe for the beta or the coefficients of the independent variables or the exogenous variables rather than the first reference to the variable that is, this is the first equation, the first variable, and the second reference to the equation. So, this actually is the second equation, the first equation *(refer to slide time 10:59)*.

The second equation has 2 in the subscript, the second subscript. Similarly, the second equation has 2 as the second subscript the mth equation has m, as the second subscript. So, there are m equations and m endogenous variables denoted by Y1, Y2 to Ym, there are k exogenous variables X1 to Xk that may include predetermining variables of Y1 to Ym as well. By predetermined, we mean, the lag values of Y's which we have already been observed.

The first element of Xt will usually be the constant 1. Finally, Ut1 to Utm is called structural disturbances. So, one important thing to be noted here is that this Xts also include the constant term, so one of the Xts would generally be equivalent to 1.

(Refer Slide Time: 13:50)

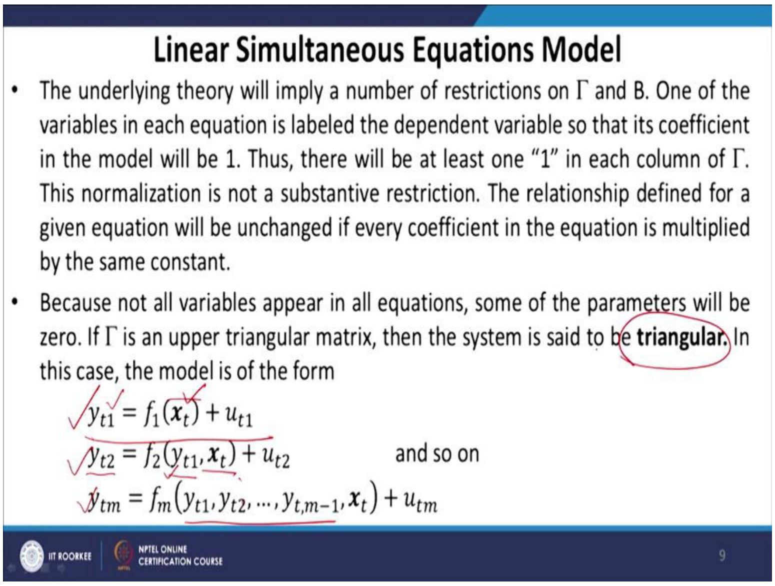


Now, in matrix form, the system may be written as Y1 to Ym it is a row vector actually, and that has reference to time period t. So, it has also referenced a time period t this also has reference to time period t. And then you can observe that this is the matrix of gamma coefficients and this is the matrix are beta coefficients *(refer to slide time 13:50)*.

Now, I if go for multiplication, then I will be multiplying this row with this column vector, which implies that the first column refers to the coefficients associated with the first equation. Similarly, here the first column refers to the coefficients associated with the first equation. The second column refers to the second equation, the third column has an association with the third equation and mth column has association with the mth equation.

Further, we denoted by gamma. So, the gamma matrix is denoted by capital gamma, and this matrix consisting of the beta parameters is denoted by capital beta. And also need to be noted that gamma is a m by m matrix, and beta is a k by m matrix. Now, each column of the parameter matrices is a vector of coefficients in a particular equation, where each row applies to a specific endogenous variable.

(Refer Slide Time: 15:18)



The underlying theory will imply a number of restrictions on gamma and beta. One of the variables in each equation is leveled the dependent variable so that its coefficient in the model will be 1. Thus, there will be at least one 1 column of gamma, this normalization is not a substantive restriction. The relationship defined for a given equation will be unchanged if every coefficient in the equation is multiplied or divided by the same constant.

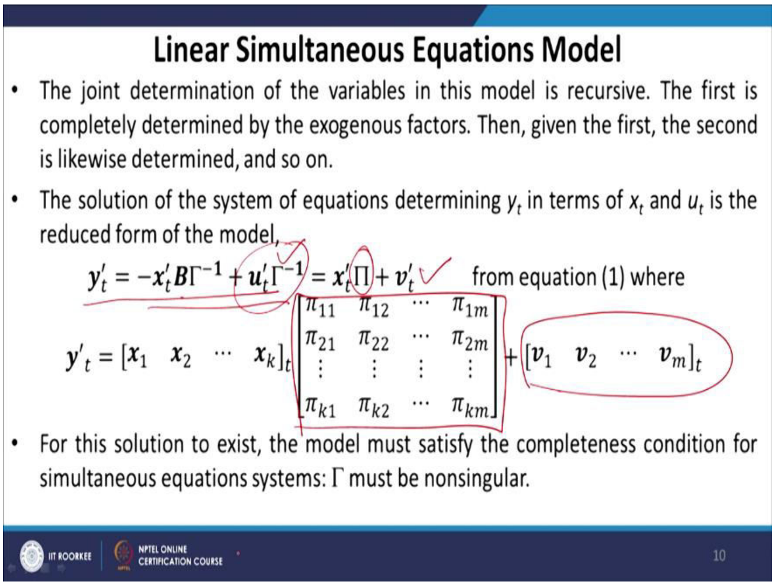
So essentially, the way we generally write an equation, what it says is that, of that m endogenous variable, each equation will have one of the endogenous variables to the left-hand side, and the rest of the endogenous variables could be on the right-hand side. And the variable on the left-hand side will have a coefficient that is equal to gamma. This means that each column of that gamma matrix will have one 1 at some place or other most often for the first equation, it will be 1 the very first coefficient for the second location it could be the second 1, which is, which will take the value 1 and so on.

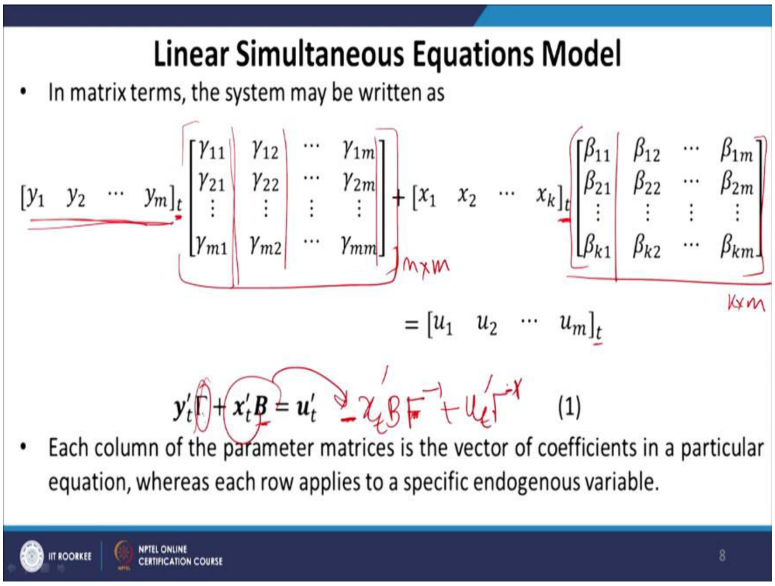
So, that is how it is, it could be expressed because not all variables appear in all equations, some of the parameters will be 0. Of course, it is quite possible that not all of them are there in each and every equation. If gamma is an upper triangular matrix then the system is said to be triangular. In this case, the model is of the form. So, now here the thing is that first of all, we are writing the first observation or the first endogenous variable Yt1 as a function of only the exogenous variables.

So, it is completely determinable by simply applying OLS. We are not specifying the functional form of the exogenous variables, it should be generally a linear one, but we do not specify how many exogenous variables should be there. And then Yt2 is basically the second variable at time period t is actually a function of Yt1 and the exogenous variables.

So, since we have already determined Yt1 from the first equation, now that value of Yt1 or predicted Yt1 could be used here in order to determine Yt2, and so on. Finally, we will be having Ytm, which is a function of all the Y's up to Ytm minus 1 and the exogenous variables plus, of course, the disturbance term. So, this is actually a triangular system.

(Refer Slide Time: 18:02)



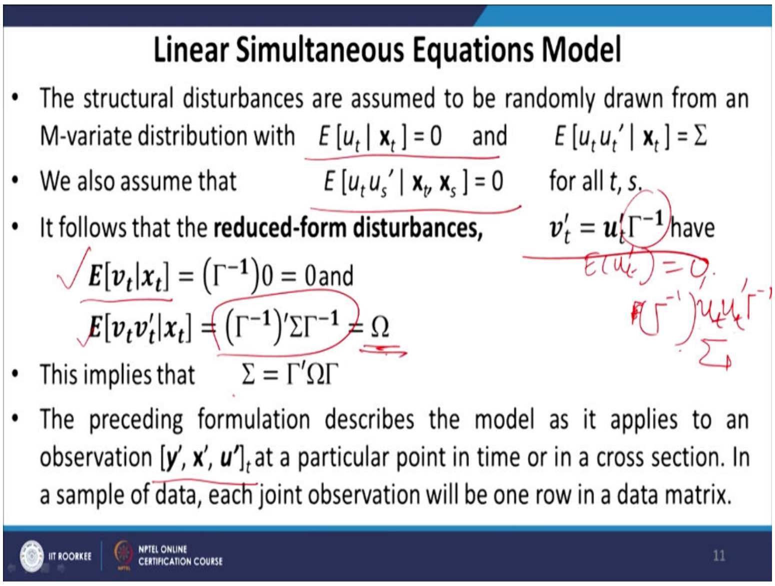


The joint determination of the variables in this model is recursive. The first is completely determined by the exogenous factors, then given the first the second is likewise determined and so on. The solution of this system of equations determining Yt terms of Xt and Ut is the reduced form of the model *(refer to slide time 18:02)*.

So, we are first of all taking the right-hand side and then post multiplying it by gamma, so gamma inverse. So, first of all, when we multiply the gamma inverse, we do not have anything on the left-hand side other than Yt prime, and on the right inside we will be having Xt prime beta post gamma plus Ut, this is minus Ut prime gamma inverse. And what we say, we call minus beta-gamma inverse pi and this entire thing is denoted by Vt prime. So, this is how, we have this reduced form, which is the pi matrix. And of course, these disturbance terms, which involves, which are basically combinations of the error terms or the structural disturbances along with the coefficients of the gamma inverse matrix *(refer to slide time 18:02)*.

For the solution to exist, the model must satisfy the completeness condition for simultaneous equation systems, which implies that gamma must be non-singular, that is, gamma must be an invertible matrix.

(Refer Slide Time: 19:55)



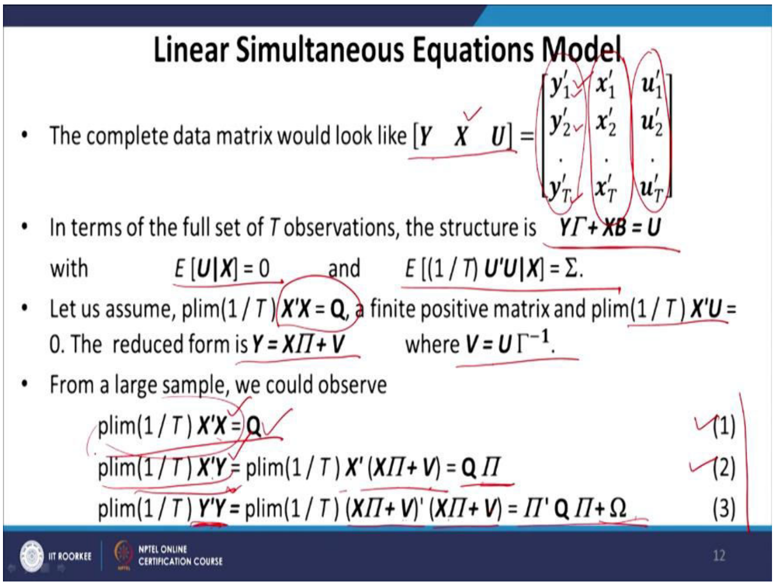
The structural disturbances are assumed to be randomly drawn from an embedded district equation with the expected value of Ut conditional upon Xt equals 0 and the expected value of Ut, Ut prime conditional upon Xt is sigma which is basically a various covariance matrix. We also assume that the expected value of Ut, Uts prime conditional upon Xt, Xs is equaled to 0 for all t, s, which implies that there is no covariance between the error terms, the cross error terms.

Now, it follows that the reduced form disturbances that is Vt prime which is equal to Ut prime gamma inverse have expected value of Vt conditional upon Xt equals to 0 because this is actually equal to the expected value of Ut prime and expected value of gamma inverse would be gamma inverse because it is a matrix of population parameters, so they are constant.

And since this equals 0, the expected value of Vt will also be equals to 0. Then we consider the variance of the expected value of Vt, Vt prime. What we observe is that Vt is actually equal to gamma, gamma inverse prime Ut and then Vt prime is Ut prime gamma inverse. Gamma inverse prime and gamma inverse remain constant and the expected value of Ut, Ut prime is actually sigma. So, we have this expression *(refer to slide time 19:55).*

Suppose we denote it by omega, so then this implies that, sigma is equal to gamma inverse omega gamma. The preceding formulation describes the model, as it applies to an observation y prime x prime u prime observed at points time t that is at a particular point in time or in a cross-section. In a sample of data, each joint observation will be one row in a data matrix.

(Refer Slide Time: 22:03)



The complete data matrix would look like, suppose we denote the capital Y, so capital Y is basically a vector of all the Ys where each one of these Y1 consists of m, y, that is, m endogenous variables for that particular period. So here, 1 refers to the first period, 2 refers to this second and capital T refers to the last period the th period. So, for each and every period, we will be having m observations.

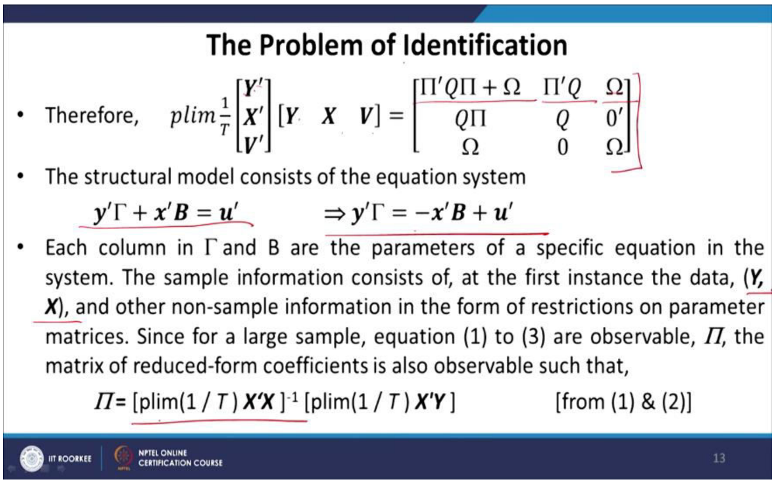
Similarly, capital X refers to the vector of all the exogenous variables, and capital U refers to the vector of all the structural disturbances observed across all time periods. So, in terms of the first set of T observations, that structure is now denoted by Y gamma plus X beta equals U. So, this is the full structure with expected value a U conditional upon X equals 0 and an expected value of 1 upon T U prime U conditional upon X is equal to the covariance variance matrix sigma.

Let us assume probity limit 1 upon T X prime X equals Q of finite positive matrix, and probability limit 1 upon T X prime U is equals to 0. This actually implies, this is the variance covariance matrix associated with the independent variables, and this says that there is no correlation or covariance between X and U. X and U the error terms and the exogenous variables are independent.

The reduced form is Y equals X pi plus V, where V equals to U gamma inverse. This has already been proved, now, we are just going for further generalization, which would not be any different. So, from a large samples, we could observe probability limit 1 upon T X prime X equals to Q. once we have the sample we can certainly calculate Q, we can also calculate this X prime Y which is basically covariance between X and Y, as X prime and then Y is replaced with X pi plus V. So, we have X prime X pi plus X prime V. X prime V is expected to be 0 because X prime U is 0, so, we are only left with Q and pi *(refer to slide time 22:03)*.

And finally, this gives us the variance of the Y matrix or Y vector. So, Y prime Y is essentially X pi plus V prime multiplied by X pi plus V. So, what we have is, pi prime X, Xpi. So, by taking property limit or expected value we have pi prime Q pi, the cross multiplications with X pi and V and V prime and X pi are all 0, what we are left with V prime V, which is essentially equal to omega that we had deduced earlier. So, pi prime Q pi plus omega is the variance of the Y vector.

(Refer Slide Time: 25:17)



Therefore, we can observe that probably limit 1 upon T, Y prime X prime V prime multiplied by Y X V essentially this implies that we are trying to obtain the variance-covariance matrix of the entire structure, and that turns out to be something like this, which is simply Y prime Y is the variance of Y, so we have pi prime Q pi plus omega just derived in the previous slide.

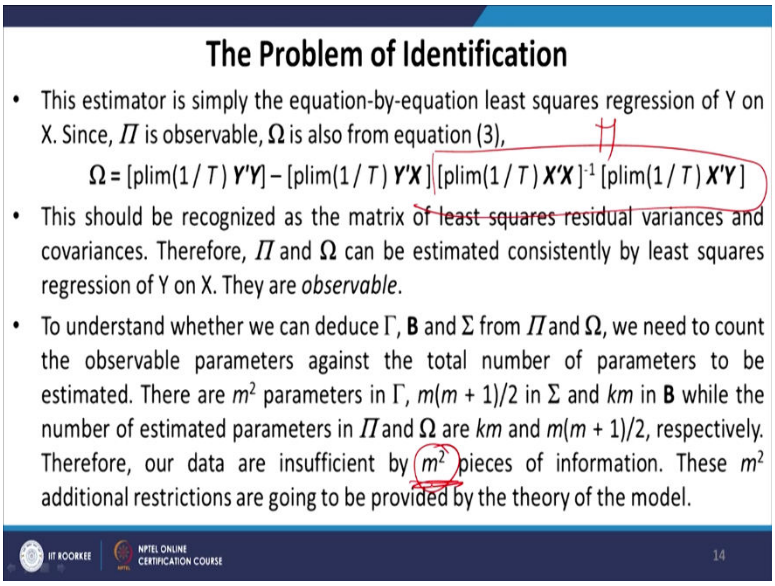
Then Y prime X will be pi prime Q, Y prime V is essentially omega, and so on, because y prime would be X pi prime plus V prime. And since X prime and V prime are independent so that would be 0, so we will be left with omega. So, these things can be, this entire various covariance matrix can be deduced.

The structure model consists of the equation system, which is Y prime gamma plus x prime beta equals u prime, alternatively, this can be written like this, where we have minus x prime beta or plus u prime on the right-hand side. Now, each column of gamma and beta are parameters of a specific equation in this system. The sample information consists of at the first instance the data that is Y and X and other non-sample information in the form of restrictions on parameter matrices.

Since for a large sample, the equation 1 to 3 are observable, these are equations 1 to 3. And since we have observations on X and Y, we can also calculate these things. So, pi, the matrix of reduced form coefficients is also observable such that pi will be provided limit 1 upon T, X prime X whole inverse parameter limit 1 upon T, X prime Y. This comes from equations 1 and 2.

You can see that pi is actually equal to this expression multiplied by basically this is actually pi would be, this expression pre-multiplied by Q inverse and then we replace the value of Q inverse here *(refer to slide time 25:17)*. So, Q inverse is actually this expression whole inverse, and this is pre-multiplied. So, that is why we have this expression whole inverse multiplied by probability limit 1 upon T X by Y. So, this is the derivation of the reduced form parameters.

(Refer Slide Time: 27:43)



The estimator is simply the equation by equation least square regression of Y on X. Since pi is observable omega is also observable from equation 3 *(refer to slide time 27:43)*. From equation 3 if pi is observable Q we know is observable omega will also be observable. And this is actually expressed as this which acts, straight away comes from the equation 3 that omega is equality probability limit 1 upon T Y prime Y minus plim 1 upon T Y prime X multiplied by plim on upon T X prime X inverse, so this is actually the formula for pi and the rest follows.

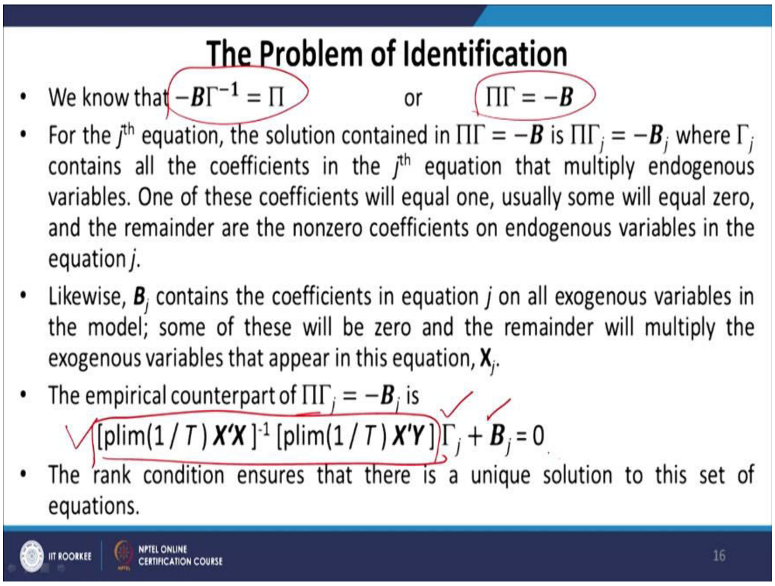
This should be recognized, as the matrix of least-squares residual variances and covariances therefore, pi and omega can be estimated consistently by least squares regression of Y on X they are observable. To understand whether we can decide gamma, beta, and sigma from pi and omega, we need to count the observable parameters against the total number of parameters to be estimated.

So, far the argument was actually revolving around the fact that the reduced form parameters which are basically the pie matrix and the variance-covariance matrix and we argued that these are actually observable. Now, the point is that, whether our structure from coefficients that is gamma, beta, and the sigma matrix whether they are observable or not.

So, first of all, there are m square parameters in gamma m into m plus 1 upon two parameters in sigma and k multiplied by m parameters in beta. While the number of estimated parameters in pi and gamma are k multiplied by m and m into m plus 1 divided by 2 respectively. Therefore, our data are insufficient by m square pieces of information.

These m square additional restrictions are going to be provided by the theory of the model. As at the outset we say that these are structural equations based on certain theoretical arguments. So, we expect that, since we do not have observations, we are running short of m square observations, that is why that must be provided with the theory or the structure.

(Refer Slide Time: 30:08)



So, we actually now face the problem of identification. The restrictions come in the form of normalizations most commonly exclusion restrictions and other relationships among the parameters such as linear relationships or specific values attached to coefficients. The conclusion is that some equations systems are identified and others are not. The formal mathematical conditions under which any equation system is identified are known as rank and order conditions.

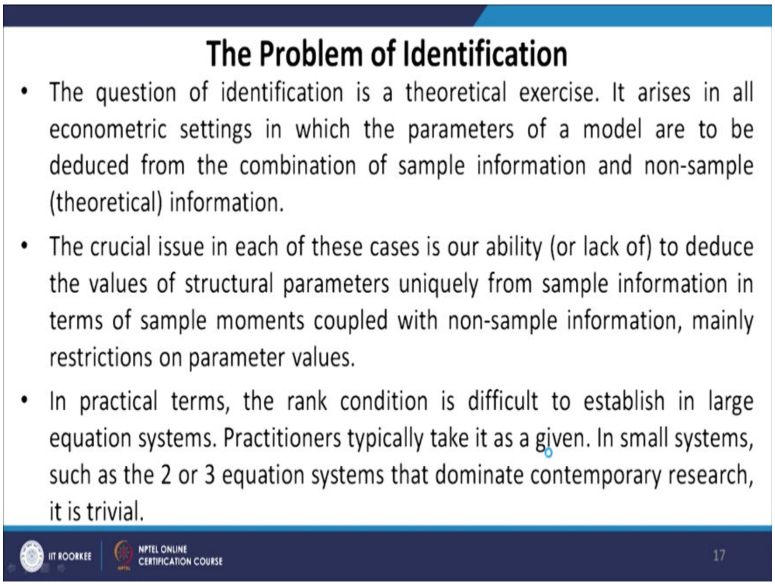
The order condition states that the number of exogenous variables that appear elsewhere in the equation system must be at least as large as the number of endogenous variables in the equation. Alternatively, for an equation system, each equation must contain its own exogenous variable that does not appear elsewhere in the system. The order condition is a necessary condition for identification, but it is not sufficient. We need a sufficient condition, and that is the rank condition.

Now, what is this rank condition? We know that minus gamma inverse is equals to pi or alternatively we can write that pi gamma equals to minus beta for the jth question, the solution contained in pi gamma equals minus beta is pi gamma j equals minus beta j, where gamma j contains all the coefficients in the jth equation that multiply endogenous variables. So, basically the jth column.

So, one of these coefficients will be equal to 1, usually, sum equal to 0 and the remainder is the non-zero coefficients on endogenous variables in the equation j. Likewise, beta j contains the coefficients in equation j on all exogenous variables in the model, some of these will be 0 and the remainder will multiply the exogenous variables that appear in the equation in this equation, which is the equation j and the corresponding matrix or vector of exogenous variable is Xj.

The empirical counterpart of pi gamma j equals minus beta j is basically, we are replacing pi with the estimated part of pi, which is plim 1 upon T X prime X whole inverse multiplied by plim 1 upon T X prime Y, this is we have just derived in the previous few slides. This is usually multiplied by gamma j plus beta j and that equals 0. The rank condition ensures that there is a unique solution to this set of equations, which is actually expressed in terms of the rank of the relevant matrix, and this should be ideal that the rank of the relevant matrix should be such that, we have independent rows so, that there is, all the equations are actually identified.

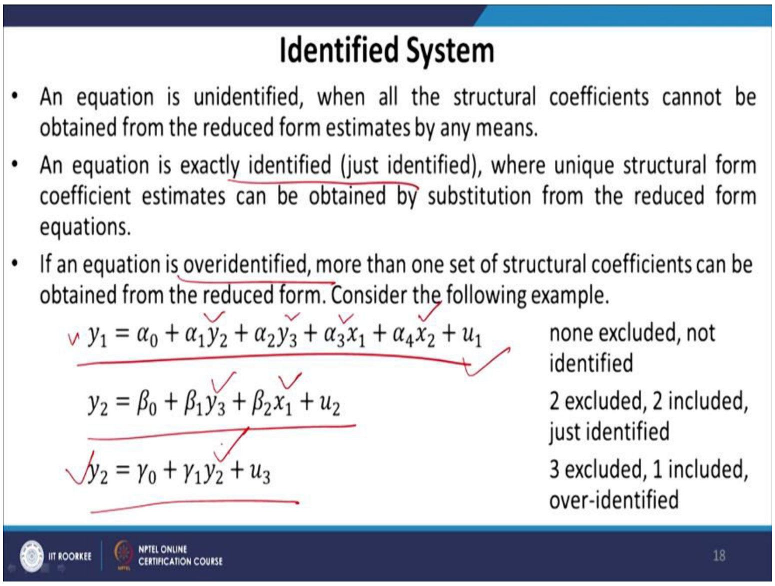
(Refer Slide Time: 33:15)



The question of identification is a theoretical exercise. It arises in all econometric settings in which the parameters of a model are to be deduced from the combination of sample information and non-samples or theoretical information. The crucial issue in each of these cases is our ability or lack thereof, to deduce the values of structural parameters uniquely from sample information in terms of sample moments coupled with non-sample information, mainly restrictions on parameter values.

In practical terms, the rank condition is difficult to establish in large equation systems. Practitioners typically take it, as a given. In small systems such as the 2 or 3 equations systems that dominate contemporary research, it is actually trivial.

(Refer Slide Time: 34:03)



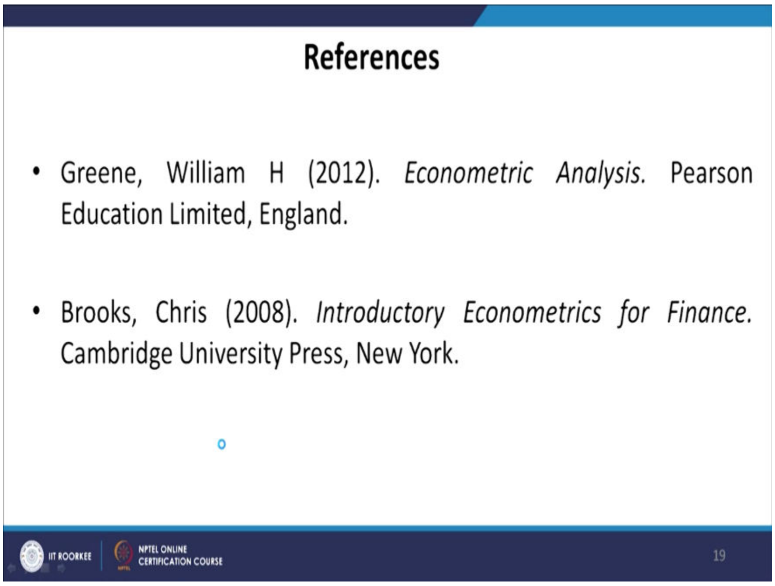
So, we talk about what is an identified system or just overidentified or under-identified system. An equation is an identity when all the structural coefficients cannot be obtained from the reduced form estimated by any means.

So, an equation is exactly identified or just identified, where unique structural form coefficient estimates can be obtained by substitution from the reduced form equations. And finally, if an equation is over-identified, more than one set of structural coefficients can be obtained from the reduced form.

So, now we consider an example *(refer to slide time 34:03)*, suppose we write Y1 as alpha naught plus alpha 1, Y2 plus alpha 2 Y3 plus alpha 3 X1 plus alpha 4 X2 plus the disturbance term. Now, you can see that we have actually 1, 2, 3, 4, 5 variables in all. And since in this equation, none is excluded, not it is not identified. This is an unidentified equation.

Now if we have Yt equals beta naught plus beta 1 Y3 plus beta 2 X1 plus Ut, so you can see that two variables are included and two are excluded. So that is why this is just identified. And the last one has only 1 variable included and 3 variables excluded. Of course, we have one here, and we are not considering that. On the right-hand side, among the four, only 1 is included 3 are excluded, and that is why this is an overidentified system.

(Refer Slide Time: 35:50)



So, that actually so far brings me to the end of the problem of identification. But in the next module, we will begin with another example of the problem of identification, and then continue with the inference and estimation of the simultaneous equation system. Thank you.