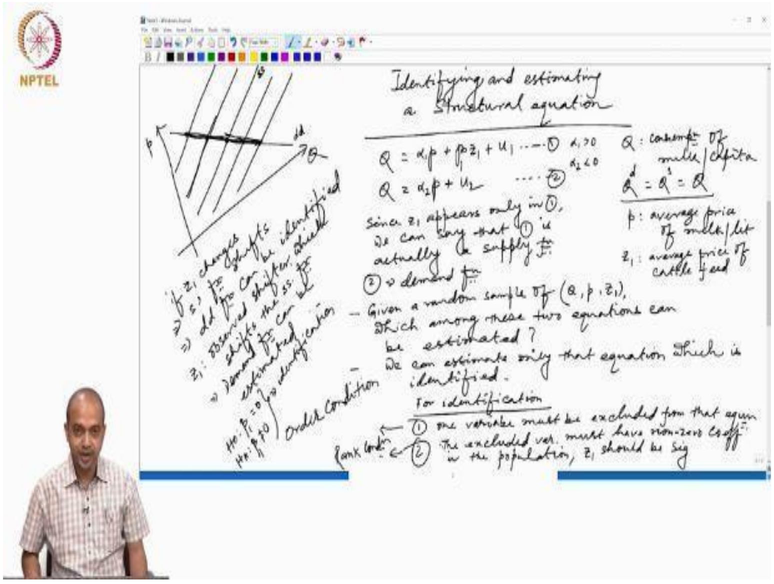
Applied Econometrics

Prof. Sabuj Kumar Mandal
Department of Humanities and Social Sciences
Indian Institute of Technology-Madras

Lecture-15
Simultaneous Equation Model-Part IV

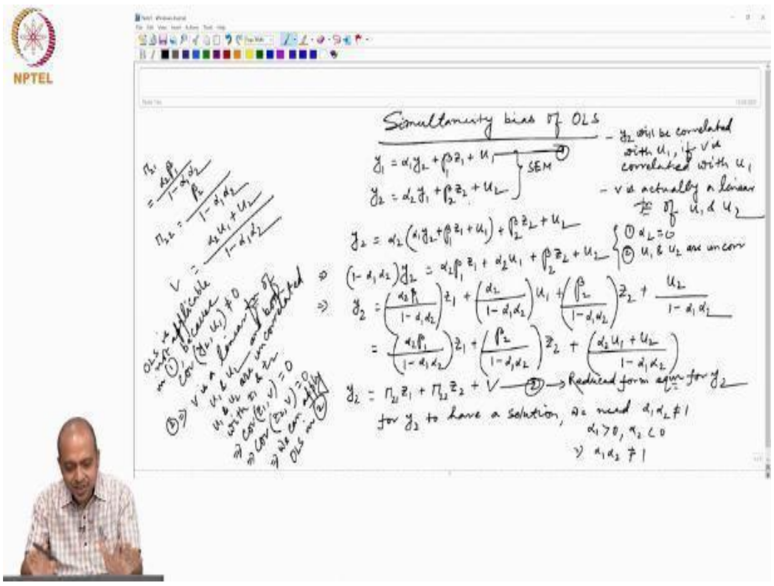
So, this is the structure, then the question is what is the solution. So far we understood there is no simultaneous bias but what could be the solution of that.

(Refer Slide Time: 00:43)



Now, to understand the solution what we need to do is identifying and estimating a structural equation.

(Refer Slide Time: 01:42)



By structural equation I actually mean these two equations $y_1 = \alpha_1 y_2 + \beta_1 z_1 + u_1$ and $y_2 = \alpha_2 y_1 + \beta_2 z_2 + u_2$.. And these two equations both are structural equation, because they have independent meaning. So, if I need to estimate the structural equation number 1, then I need to first of all identify that equation identification comes prior to estimation.

So, if you do not identify an equation you do not know which particular equation you are estimating. For example, demand and supply they constitute a simultaneous equation model. Both demand and supply are function of price. So, unless I can identify the demand function or supply function, I cannot estimate them because it does not have any meaning and how will you identify.

I need to have some variable, which will appear only in one equation, but not in the other. So, that example we will take once again and we will try to understand the identification more clearly even today. So, let us say that this is an equation $Q = \alpha_1 p + \beta_1 z_1 + u_1$and this is $\alpha_2 p + u_2$. Let us say this is equation 1 and this is equation 2 and Q here is actually consumption of milk per capita. And we assume that demand for milk, supply of milk equals to Q. So, we assume this equilibrium condition. That is why I am writing both the equation in terms of only Q. So, please do not get confused. So, instead of writing Q_d or Q_s let us say this is the milk consumption per capita in a particular state and I have assumed this equilibrium condition $Q_d = Q_s = Q$ that is why I have written both the equation in terms of Q.

So, this is consumption of milk per capita. What is p? p is average price of milk per capita and $\mathbf{z_1}$ is actually the average price of cattle feed. Now, if I ask you think about these two equations and tell me which equation is the demand for milk and which equation is the supply of milk. Because, both the equations are presented in terms of Q and at equilibrium $Q_d = Q_s = Q$ is the assumption we made.

This is how we have specified p and $\mathbf{z_1}$. Then my first question is which is the demand function, which is the supply function?. Out of this equation 1 and equation 2 you need to identify them as a demand function and as a supply function. Now, if you think then you can easily understand that both demand and supply of milk is function of price, that is fine that is why price is common in both the equations. But, in the first equation z_1 is appearing. What is z_1?

It is the Average price of cattle feed. Now, do you think cattle feed has anything to do with demand for milk? No, cattle feed has no role in determining demand for milk. Does it have any role in supply of milk? Yes. Higher the cattle feed price, then it is it becomes less profitable for the milk supplier to supply milk. So, that is the reason since z_1 is appearing only equation 1 we can say that equation 1 is actually a supply function and equation 2 is actually demand function because here the demand is a function of price. Now, Given a random sample on p, Q and z_1, which among these two equations can be estimated?

Now, to answer this question I have already mentioned that identification comes prior to estimation. So, we can estimate only that equation, which is identified. That is the hint**.** So, now you have to think out of this equation 1 and 2, which equation is identified? We have already discussed this identification in our previous class.

Now, I will explain once again using simple diagram. Let us say in the y-axis we are measuring price of milk, in the x-axis this is quantity, this is the demand and this is the supply**.** So, α_1 is actually positive, so we assume that here α_1 is greater than 0 and α_2 is less than 0- upward slopping supply and downward slopping demand. Now, in equation 1, that means in supply equation, there is a factor called z_1 which means cattle feed.

If there is a change in the in the z_1, so when price of cattle feed changes, what will happen? So, z_1 is actually a parameter we have already discussed. Apart from price and quantity we assume all other factors to be constant. So, a change in z_1 means there would be a parametric shift in the supply function.

When z_1 increases, supply of milk becomes less profitable, supply function will shift upward. When z_1 decreases, that means price of cattle feed decreases, supply of milk become more profitable it will shift like this**.** Now, if the supply function is shifting either upward or downward, that means the demand function is same and we are getting new equilibrium from the intersection point.

And if we take all this new equilibrium and connect them with a line, then basically I am identifying the demand function. So, supply function shifts and as a result of which demand function can be identified. So, that is why we said if you recall once again z_1 is called an observed shifter which shifts the supply function.

And because the supply function is shifting, I am getting new equilibrium by the intersection of the supply and demand and I am able to trace out the demand function. So, identification means I am actually able to trace out the demand function that is why it implies demand function can be estimated also since it is identified**.** So, we can only estimate the demand function, because it is identified.

But, to estimate the supply function we should have some variable, which will shift the demand function. But, in the demand function if you look at there is no observed demand shifter. There is only unobserved demand shifter in terms of u_2. That is the reason why we cannot estimate the supply function, because supply function is not identified. So, that means for identification there are two conditions.

See here the demand function is identified because demand function excludes a variable which is z_1. So, one variable must be excluded from that equation. Here, z_1 is excluded from equation 2 that is why I am able to identify. This is condition number 1. That is called necessary condition. Condition number 1 is called necessary condition, which is also known as order condition.

If you recall we have already discussed this in our previous section, but there is one more condition that I am going to tell that we have not discussed. The variable must be excluded and the second condition is that the excluded variable must have non-zero coefficient in the population. That means β_1 should be significant and non-zero. If it is not significant, that means it has no meaning. This second condition is known as rank condition.

These are the two conditions. So, that means z_1 is excluded from equation 2, but appearing in equation 1. What is z_1 actually? That is also we have discussed in our previous class previous module actually. So, when z_1 is excluded, excluded exogenous variable is known as instrument. So, that means I am able to estimate equation 2- the demand function, because I got an instrument for the endogenous variable p.

So, z_1 is basically an instrument, z_1 is an instrument for p that is why this equation 2 is identified. But, to estimate equation 1 I have no instrument to be used for p in equation 1. which is an Excluded exogenous variable that is significant. This is the identification condition we discussed. H_o: $\beta_1 = 0$ against H_a: β_1 not equal to 0 is the identification condition. So, when z_1 is excluded from equation 2 and have a non-zero coefficient, we say that equation 2 is identified and we are able to estimate that also. So, there are two conditions. One is called rank condition, the way one variable must be excluded since we have only one endogenous variable.

We need only one excluded exogenous variable and second condition is that the excluded variable must be significant also. So, that means the same instrumental variable estimation technique is applicable in the context of simultaneous equation model also. So, that is the reason we say that the equation must be identified, otherwise we cannot estimate the equation.

And just I would like to mention one more thing why z_1 is not used as an instrument for equation 1. Since, z_1 is already there in the model, since z_1 is an explanatory variable appearing in equation 1; z_1 cannot be used at an instrument for the identifying the first equation. A variable which is already included cannot be used as instrument.

The variable must be excluded that is why according to the exogeneity condition, that variable must be excluded from the equation to be used as an instrument. That is why while z_1 is used as an instrument for p in the second equation it cannot be used as an instrument to identify the first equation.