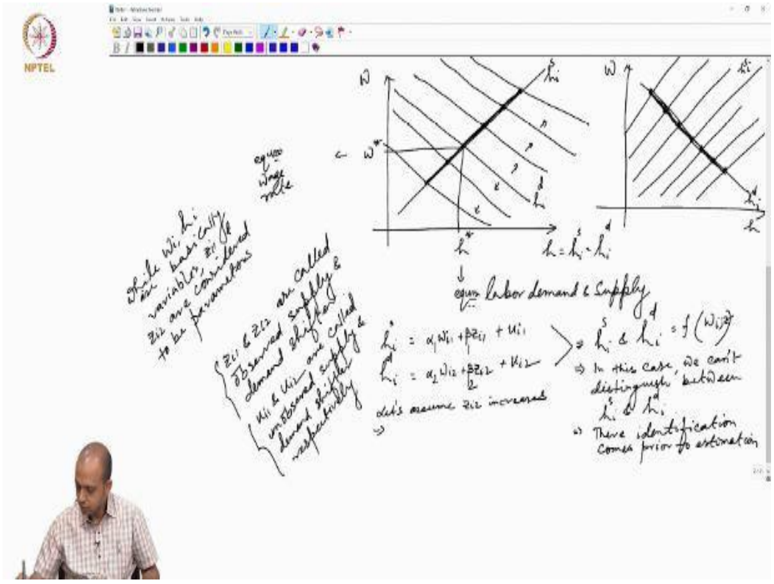
Applied Econometrics
Prof. Sabuj Kumar Mandal
Department of Humanities and Social Sciences
Indian Institute of Technology-Madras

Lecture-13
Simultaneous Equation Model-Part II

(Refer Slide Time: 00:15)



Similarly, if z_{i1} increases or z_{i1} changes, then what will happen? Since, it is appearing only in the labour supply function so, labour supply will change. So, I will draw another diagram this is h and this is w. This is demand(h_i^d) and this is supply(h_i^s). This is . So, now we are assuming z_{i1} is changing. That means, labour supply will change. If it is increasing, that means if industry wage is increasing, z_{i1} is increasing, then labour supply in agriculture will come down.

So, this will be like this. Similarly, effectively if industry wage is coming down, then labour supply in agriculture will go up. But, labour demand will remain same because z_{i1} is not appearing in the demand function. If you collect all this new equilibrium then that will help us trace out the labour demand function. This is the labour demand function which explains the identification problem and conceptually make us understand what exactly we are identifying.

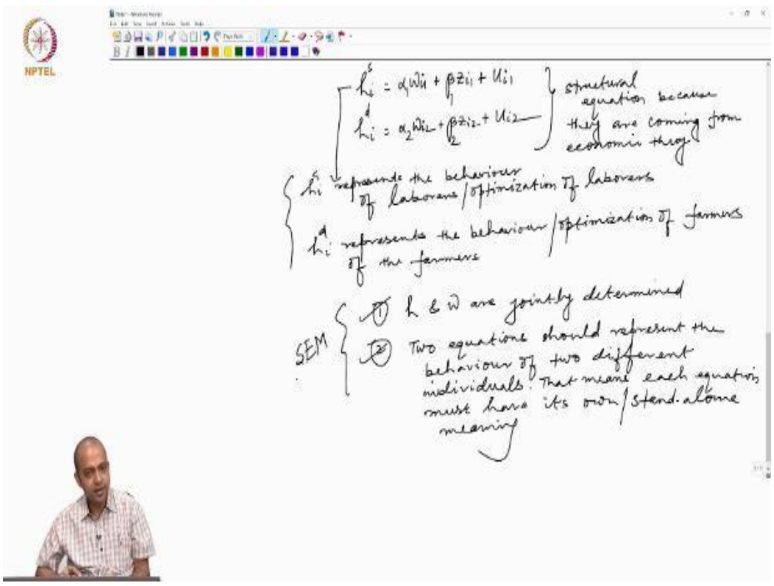
So, to identify a function we need to have a variable, which is not appearing in the supply function, but must appear in the second function. Secondly, for identifying the demand function we need to have a variable, which is appearing in the demand function, but not appearing in the supply function. That is how we can identify the equations. If we need to identify the labour supply function what we need to do?

We need to change the labour demand function such that it will give different intersection point. So, to identify h_i^s we need to have a variable like z_{i2}, which will appear only in the demand function, but not in the supply function. To identify the supply function, we need to have a variable like z_{i1}, which will appear in the supply function but not in the demand function.

That is why this z_{i1} and z_{i2} they are called observed supply and demand shifters. Since, because of a change in z_{i1} and z_{i2} this demand curve and supply curve they are shifting either leftward or rightward, they are called shifter. Observe demand shifter is called observed because we can always observe z_{i1} and z_{i2}. Is there anything called unobserved demand and supply shifter?

Yes. This u_{i1} is called unobserved supply shifter and u_{i2} is called unobserved demand shifter, because we cannot observe them. Because u_{i1} and u_{i2} will also capture certain variables which are not appearing here. And if there is a change in those variables, obviously these functions will change, but we cannot observe. That is why we say that u_{i1} and u_{i2} are called unobserved supply and demand shifter respectively. Now, these two equations h_i^d and h_i^s are called structural equations. Once again, if you remember, we introduced this term structural equation, why this is called structural equation?

(Refer Slide Time: 06:24)



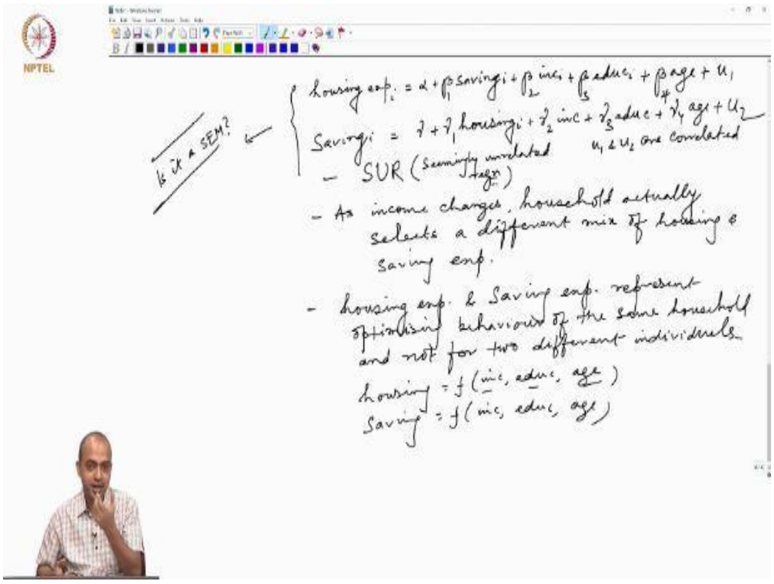
So, this $h_i^s = \alpha_1 w_{i1} + \beta_1 z_{i1} + u_{i1}$ and $h_i^d = \alpha_2 w_{i2} + \beta_2 z_{i2} + u_{i2}$ are called structural equations because, they are coming from economic theory. What is the economic theory? The economic theory here is the labour supply functions- when the labourers are supplying labour. So, obviously labourers are doing also an optimization in their mind or maximization**.** So, labourer will supply that much labour that will maximize their utility or satisfaction. And of course they will optimize by looking at the wage rate which is the prevailing wage rate in the market**.** Looking at the prevailing wage rate in the market, labourers will maximize their utility and determine the optimum labour supply.

Similarly**,** h_i^d equation represents another optimization, which the farmers are doing. Farmers will also demand that much labour to maximize their production or optimize their production given this wage rate. So, that means while h_i^s represents the behaviour of labourers, h_i^drepresents the behaviour of the farmers. So, two equations are representing optimizing behaviour of two different individuals that we have to keep in mind. Because, both equilibrium labour supply and labour demand are coming from the optimizing behaviour of the labour and the farmer. So, first equation is the supply function, which is representing the optimizing behaviour of the labourers. Second equation is the labour demand function, which is the optimizing behaviour of the farmers. Now this is called the second condition for a system of two equations to be qualified as simultaneous equation model.

What I am saying is very important thing. Previously, we mentioned one condition, we said that the two variables should be jointly determined. what are the two conditions here? h and w are jointly determined is condition number 1. Second condition is that two equations should represent the behaviour of 2 different individuals, that means each equation must have its own or stand alone meaning, this is important.

So, these two conditions if satisfied then only will say that the system of two equations is simultaneous equation model or SEM. So, first of all the two variables should be jointly determined and two equations should represent the behaviour of two different individuals. In case, only the first equation is satisfied but second equation is not then we cannot say that is actually simultaneous equation model. Sometimes it may so happen that the two variables are jointly determined, but they are not representing the behaviour of two different individual. Rather, they represent the behaviour of a single individual only.

(Refer Slide Time: 13:51)



Take an example, let us say that we are talking about the housing expenditure and savings function. We know that how much a household will spend for housing and how much the household will go for saving, they are jointly determined from a set of other variables, let us say, income, education and wage. So, we can very well say that these are constituting a simultaneous equation model is that true.

Let us try to understand. This is a housing expenditure of the i^{th} family, which is a function of let us say $\alpha + \beta_1 saving_i + \beta_2 income_i + \beta_3 education_i + \beta_4 age_i + u_i$

Similarly, we have a saving function, which is let us say, $\gamma + \gamma_1 houseing_i + \gamma_2 income_i + \gamma_3 education_i + \gamma_4 age_i + u_2$

So, can we say that is it a SEM, because there are two equations, two variables housing and saving. They are jointly determined. Can you say that this is a simultaneous equation model? Actually, this is not a simultaneous equation model, because the second condition is not satisfied. How can we say that each equation has it is own ceteris paribus condition satisfied. That means, what would be the interpretation of beta2?

As income increases by one unit how the household expenditure is changing. Similarly, here as income increases how saving is changing. Keeping in the first equation what I am saying, as income increases how housing expenditure is changing, keeping other factors constant including saving. Similarly, here as income increases γ_2 measures, how savings on an average changes.

Keeping other factors constant including housing, but does it really make sense because, as income changes household actually selects a different mix of housing and saving expenditure**.** So, we cannot actually keep saving constant when income changes to identify its impact of housing. Similarly, we cannot keep housing constant to find out the change of income on saving, because the moment income changes household will select or choose a different mix of this housing and saving**.**

A different mix of housing and saving expenditure would be selected at each and every level of income meaning these two equations cannot stand alone. These two equations can be very well reduced to only one. Even though they look like two different equations, they are actually representing one individual. So, that means housing expenditure and saving expenditure represent optimizing behaviour of the same household and not for two different individuals.

So, this is not a same. And more interestingly and importantly even if we assume that these two equations form a simultaneous equation model, we cannot actually identify them because housing is a function of all this vectors-income, education, age, saving**.** So, that means at equilibrium if we equate these two, that means both housing and saving are actually the function of same set of variables.

So, same set of exogenous variables I am talking about. There is no observed shifter. So, we cannot actually identify neither the housing not the saving function based on the logic provided earlier. And as we said, identification comes prior to estimation. So, these two equations are actually a single equation.

Even though, they are jointly determined they should be estimated using different technique all together. So, that means even though we assume they are independent that means u_1 and u_2 are independent, actually u_1 and u_2 are correlated. So, that is why instead of estimating this type of equation by simultaneous equation framework rather, the appropriate methodology to estimate this type of equation is by Seemingly Unrelated Equation or SUR.

So, the appropriate methodology is SUR, which is called seemingly unrelated regression-these are not independent equations. It appears like unrelated and independent, but actually they are related. Because, they represent the optimizing behaviour of the same household that is why these two equations are not simultaneous equation example of simultaneous equation model.

Rather, that can be reduced to a single equation. And we need to estimate whenever we have this type of problem of a particular household. For example, if I introduce another equation, let us say housing expenditure, saving expenditure, expenditure for children's education, expenditure for clothe, expenditure for food in all these household given a particular level of income a particular household determines all these simultaneously.

When my income is 20,000 how much I will spend for housing? How much I will save? How much I will spend for my kid’s education? How much I will spend for five food all these things they are actually simultaneously determined. That is why those equations even though look like independent, they are actually related to each other. So, appropriate methodology should always be seemingly unrelated regression, you have to clearly keep in mind, not simultaneous equation model.

Simultaneous equation model is applicable only when variables are jointly determined. That is necessary but a sufficient condition is also required which says that the two equations must represent the behaviour or rather optimizing behaviour of two different entities, two different individuals. So, with this we are closing our discussion today, where we have just introduced what exactly is simultaneous equation model, which is the another reason for endogeneity.

We have also discussed about the structural equation and then identification, and why identification is required, how to go about identification and what is not actually a simultaneous equation model. In our next class, we will then discuss in detail about the bias that appears in simultaneous equation model. If we apply OLS and then we will try to understand the appropriate estimation methodology also. Thank you.