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

Lecture - 44
Dynamic Panel data Model: Part VII

Welcome once again to our discussion to Applied Econometrics and we are planning to discuss about the empirical estimation the preliminary empirical estimation of the models that we have discussed in the context of dynamic panel data model. And before we go back to empirics let me just recap quickly what we have discussed in our previous session. So, in our previous session basically we were discussing about the theoretical foundation of generalized method of moments.

And we discussed, first of all the method of moments conditions and method of moments estimation basically and we derived that OLS is also a specific case OLS estimator beta hat equals to x prime x inverse x prime y that we have derived using moment method of moments. So, basically the idea is in the OLS we minimize the sum of errors and in the context of method of moments what we do we use the condition that no explanatory variable should be correlated with the error term.

That means we first derive a theoretical moment condition or population moment condition using the condition expectation of x and u equals to 0 and then we derive the sample counter part of that population moment condition and then from there we try to derive the estimator and then we said that in the context when some of your explanatory variables are actually correlated with the error term.

Then we use instruments and let us assume the instrument matrix is Z and then we use the condition again expectation Z u equals to 0 and from there we derived the estimator as Z prime x inverse Z prime y instead of x prime x inverse x prime y and in that context when we discussed about method of moments, we said that when there are more number of instruments than the number of parameters to be estimated there actually we can have two cases.

In case 1, we can take only K number of parameters from that Z matrix, but that will lead to inefficient solution. So, that means an inefficient way of estimating is now taking only K

number of instruments from the set of Z. The other way is basically to use quadratic loss function which says that in the context of number of instruments to be more than number of parameters to be estimated.

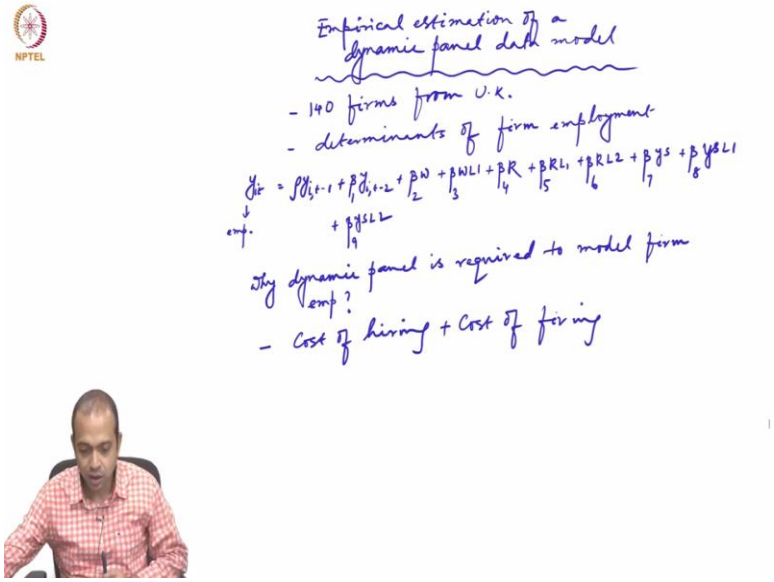
In that context we discussed about generalized method of moment which is nothing, but taking the same population moment condition what we have derived earlier and then multiplying that with some W matrix which is weight matrix we say that Wn is basically L cross L weight matrix choose an optimally that gives minimum variance GMM. So, this is basically the context.

So, in a nutshell we can say that GMM is a method of technique which is the generalized version of method of moments when you have number of instruments more than number of parameters to be estimated that is all we have discussed yesterday. Now in today's class what we will do we will take an example and then we will discuss all the preliminary models that we have discussed.

First of all, we will see when there is lag dependent variable present in your model that means in a context of dynamic panel data model what happens if we apply OLS, what happens if we apply fixed effect transformation and then how to overcome those problems. So, it is very important to understand the implications of OLS and fixed effect transformation in a dynamic panel context then only we will be able to appreciate the development of specific technique by Arellano and Bond and that is called the dynamic panel data model.

Why dynamic panel data model is required? Because as we discussed theoretically there will be bias if we apply fixed effect transformation or if we apply OLS or even if we apply first difference or transformation. So, this particular example is taken from UK.

(Refer Slide Time: 05:42)



So, this is empirical estimation we are talking about empirical estimation of a dynamic panel data model. So, the context is we are taking 140 firms from UK and what we are trying to model here we are trying to estimate determinants of firm employment and how do you model it? We model it in this way y i t equals to lambda y i t - 1 +beta 1 y i t – 2 + beta 2 let us assume this y i t equals to employment in short.

So, employment by the ith firm in dth period depends on his previous employment which is rho y i t – 1 then its second order lag that means twice lag of employment then beta 2 wage which is denoted by W and then beta 3 W L 1; L 1 means lag 1. So, wage as well as its lag then beta 3 sorry beta 4 capital a firm employment as you know that employment depends on how much capital you have already employed in your firm + beta 5 lag of capital.

And then beta 6 second lag of capital then beta 7 beta you have output firm sector output then beta 8 first lag of firm output which is L 1 + beta 9 second lag of the output. Now I will explain why the justification of each of these variables to be included. First of all, why dynamic panel is required to model firm employment? So, why first of all there should be y i t – 1.

The answer is if the firm wants to hire new worker, then there is a cost involved, what is the cost? First of all, the firm has to advertise that there is a requirement then the firm has to conduct interview to select appropriate people and then they have to offer it. So, apart from the cost for in terms of wages and salaries that needs to be paid to the new worker there is a cost of hiring itself.

There is a cost of hiring which is called as search cost because I need to advertise, I need to conduct interview and then you are not sure even if you conduct interview whether you will be getting proper people so on and so forth. So, there is a cost of hiring this is called cost of hiring that is why you see that the firm generally stick to their present level of employment they really do not go for new recruitment quite often.

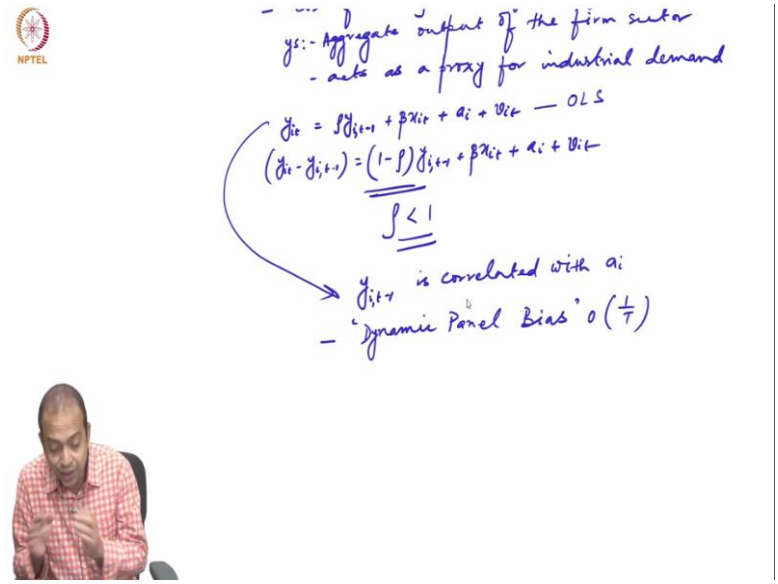
So, this is cost of hiring similarly if the firm wants to fire some of the existing laborers, then also there is a cost involved. What is the cost because if you fire some of your existing employees then what will happen? There would be protest from the labour union. Also, the existing laborer they have to work more you do not know whether they would be cope up with this new task that means extra work burden that are going to be imposed on the remaining laborers. So, there is a cost of firing also.

So, this cost of hiring and firing prevents the firms to go for either new recruitment or to fire some of the existing employment that means in every time how much employment the firm will have that very much depends on what the firm was already having in the previous period. So, this particular case actually explains why employment firm employment should be modeled as a dynamic context because of mainly cost of hiring and firing.

Then whatever other variables are very straightforward. Employment definitely is a function of wage. Market wage should determine how much the firm will hire the ith firm will higher in Tth period. Then of course the lag of wage what is the lag that was prevailing in the previous period and then what is the capital because labor alone cannot produce anything you need to have a capital to work with the labour.

And then again, the lag of all those capitals and then y s is basically what is y s? y s is the aggregate output from the firm sector.

(Refer Slide Time: 13:34)



So, y s is aggregate output of the firm sector and why I am adding this aggregate output because aggregate output of the firm sector basically acts as a proxy for demand for industrial product. It acts as a proxy for industrial demand because employment how much employment will happen in a particular year that very well depends on what is the demand for that industrial product.

And that is why we are taking aggregate industrial output which acts as a proxy for industrial demand and then we add the first and second lag of all these variables and as you know the lag of all these exogenous variables are mostly predetermined in nature. They are called predetermined variable and then we can also add beta 9 or beta 10 y r dummy. So, this is the model that we are specifying, this is basically an employment function.

Employment is a function of its previous own lagged value, wage, capital and then industrial output act as a proxy for demand for industrial product. So, now what we will do? We will first estimate I will show you. **(Video Starts: 15:45)** So this is the data if you are using this data. This is the original data Arellano and Bond they have used which is their data set. So, when you are using strata if it is a licensed version.

And if you use we have web use ab data then you will get to know about the data will be automatically installed and then you can always see the data also you see all the variables are included here. So, individual, year, employment wage, capital so industrial output everything is included here and I have 1976 to 1984. So, this 9 years data we have. So, as we have

discussed earlier that dynamic panel is basically a case where t is small and n is large that is why see here t is only for 9 years.

Now what we will do? We will estimate this model by simply OLS by applying OLS what will happen. So, we are regressing we put the regress command and then my dependent variable is basically employment n and then I am using n L 1 that means lag of employment to make the model dynamic then I have also taken in n L 2 second lag of employment and then and then lag of w.

So, instead of typing we can simply kick over here so w L 1 and then I will take capital; is denoted by small k and where is my k? k is here I have taken capital then first lag of capital and then I will also take second lag of capital then I will take y s that means industrial output. This is my industrial output and then first lag of output, second lag up output and then what I will do? I will put the yr dummy.

This is purely a simple OLS model applying here ignoring the fact that there is dynamism in the system. Why we are doing this? We are trying to understand the consequence of applying OLS when your model is dynamic in nature look at this; what is happening. So, first of all before going to any other coefficients my prime variable of interest is lag of employment which is n L 1 look at the coefficient; coefficient is 1. 04. **(Video Ends: 19:21)**

Now is this 1.04 measure acceptable? The question is no. Why this is no? I will explain once again if you do not recall it if you remember that we said y i t equals to rho y i t - 1 + beta x i t + a i + v i t if this is the model subtracting y i t - 1 from both the side then you will have y i t - y i t - 1 equals to 1 - rho y i t - 1 + beta x i t + a i + v i t. Now look at the coefficient of y i t - 1.

So, rho must be less than 1 to have partial adjustment and stability of the dynamic panel data model. So, rho must be less than 1 that is the condition we require. If rho is greater than 1 then this becomes negative and then it will not lead to a stable dynamic panel data model. So, stability requires less than 1 and that is where we said that this is partial adjustment. So, the moment I see my estimate is 1.04.

That means we can say that it is not this coefficient is not supporting my theory rho cannot be more than 1, it should be it cannot go beyond 1. So, it is basically kind of overestimated value

now why is this overestimation happening? That is also something we need to understand. Again we will go back and recall what is happening here when we are applying OLS in this model.

What is the problem here of applying OLS my y i t - 1 is actually correlated with a i if y i t is a function of a i then y i t - 1 is also a function of a i that means y i t - 1 is correlated with a i. OLS is basically ignoring that fact and when you ignore this correlation that is actually leading to a dynamic panel bias. You understand what I am saying dynamic panel bias which is of order 1 by T we have discussed.

So, my explanatory variable is correlated with the unobserved effect, but OLS is ignoring that fact.