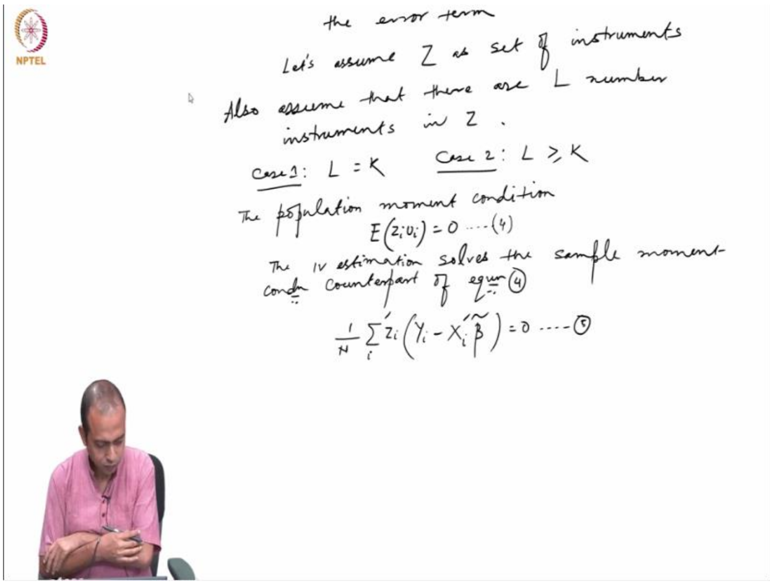
Applied Econometrics

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

Lecture - 43
Dynamic Panel data Model: Part VI

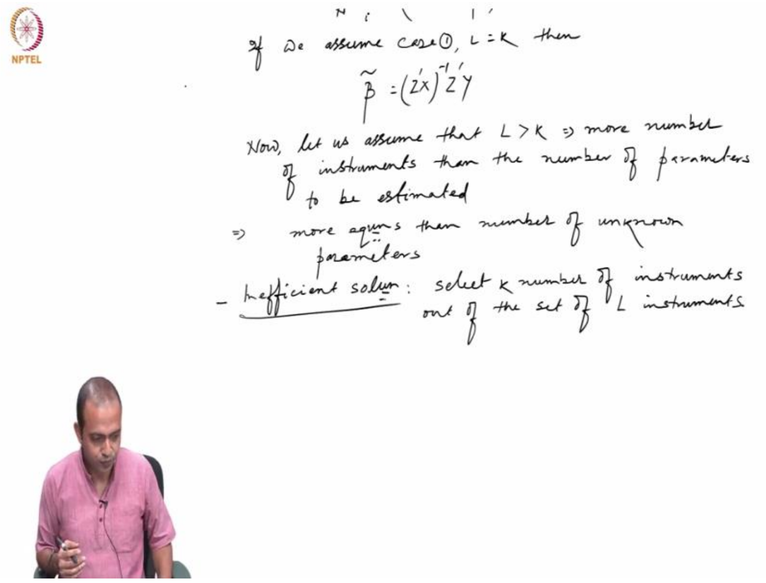
(Refer Slide Time: 00:16)



So, let us assume Z is set of instruments and assume that there are L number of instruments in the set. Now there can be two cases. So, case 1 is where L is equal to K and case 2 is where L is greater than or equal to K, where K is basically total number of parameters to be estimated in the model. So, I have K number of parameters to be estimated, but I have L number of instruments within this Z instrument matrix and then L can be either equal to K number of parameters to be estimated or L can be greater than or equals to K. Now the population moment condition is $E(Z_i\, u_i) = 0$ ---- which is equation 4. So, the IV estimation solves the sample moment condition counterpart of equation 4 which is basically the population moment condition. The sample moment condition is given by: $\frac{1}{N}\sum_{t}' Z_i (Y_i - \acute{X_\iota}\tilde{\beta}) = 0$ ----- (5)

Let us say this is equation 5.

(Refer Slide Time: 05:35)



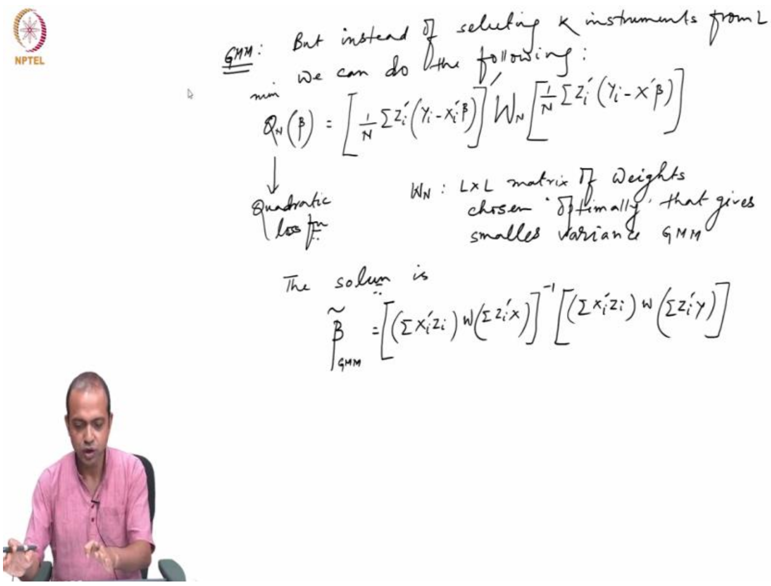
And if we assume case 1 where L equals to K then $\acute{\tilde{\beta}} = (\acute{Z}X)^{-1}\acute{Z}Y$

Now again we have derived the IV estimators using this moment condition. So, this is fine that means as long as my number of instruments are equals to number of parameters to be estimated from the model. I am fine that means my IV estimation also looks like the OLS estimate what we have derived earlier. In place of $(\acute{X}X)^{-1}\acute{X}Y$ $\acute{X}X$, this is like $(\acute{Z}X)^{-1}\acute{Z}Y$. So, Z is basically set of instruments so we have substituted X with Z that is all. So, that means in case of exactly identified system, my method of moments estimation in presence of endogeneity or that means my IV estimates can be derived by using the sample moment condition which looks like $\acute{\tilde{\beta}} = (\acute{Z}X)^{-1}\acute{Z}Y$

Now the case becomes complicated when we assume that L is actually greater than K that means we have more number of instruments than the number of parameters to be estimated that means more equations and each instrument will give rise to each one sample moment condition like this $\frac{1}{N}\sum_t' Z_i(Y_i - \acute{X_\iota}\tilde{\beta})$

So, here I have assumed that number of equations are actually equals to the number of number of instruments are basically equals to number of parameters to be estimated then the moment we get more instruments than number of parameters to be estimated then what will happen that will lead to a situation more equations than number of unknown parameters. So, what we can do in this case inefficient solution is when you have more instruments than number of parameters to be estimated then we can always select K number of instruments out of the set of L instruments. L is greater than K. so what I can do? I will take only K number of instruments from this set of L and that will lead to an inefficient solution.

(Refer Slide Time: 11:45)



Now here comes GMM what is that, but instead of doing this, instead of selecting K instruments from L what we can do we can do can do the following that is more efficient actually and that is the GMM estimator which is nothing, but what we actually doing I will write it here

$$Q_N\beta = [\frac{1}{N}\sum\nolimits_t Z_i{}'(Y_i - X_\iota'\beta)]'W_N[\frac{1}{N}\sum\nolimits_t Z_i{}'(Y_i - X'\ \beta)]$$

We will take see this is the sample moment condition which we have derived earlier we will use that take transpose of that and then we will use one W weight matrix and then we will multiply these $\frac{1}{N}\sum_t\ Z_i{}'(Y_i - X_\iota'\beta$. So, we will minimize this. So, this is called quadratic loss function. So, I will select the beta so what the idea here is the W_N is basically Lx L matrix of weights that chosen optimally that gives smallest variance GMM estimate.

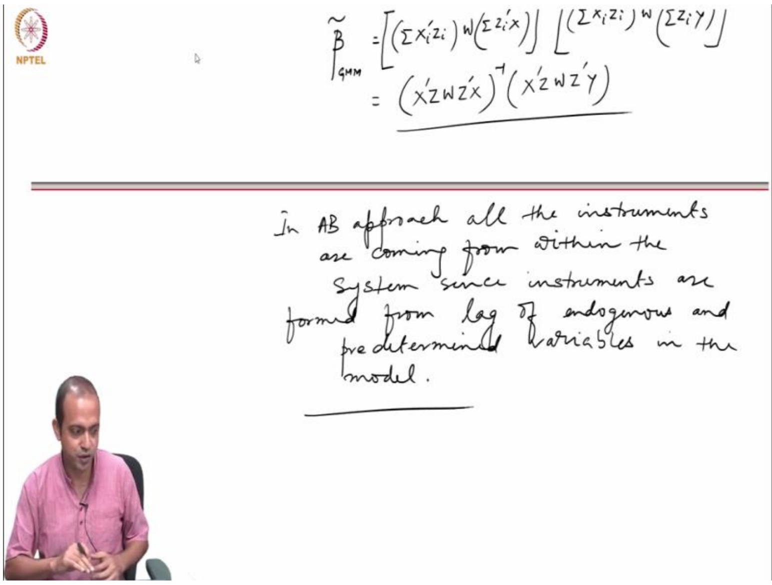
So, that means what I am actually doing when L is greater than K instead of selecting K number of instrument out of L, I can construct a quadratic loss function. That means minimizing this particular equation which is derived from the sample moment condition $= [\frac{1}{N}\sum_t\ Z_i{}'(Y_i - X_\iota'\beta)]'W_N[\frac{1}{N}\sum_t\ Z_i{}'(Y_i - X'\ \beta)]$

So, that is the idea when I have more number of instruments than the parameters to be estimated and that is basically if you solve this, if you minimize this with respect to beta then the solution will give you, the solution is basically $\hat{\beta}$ GMM or $\tilde{\beta}$ GMM you can write equal to $\widetilde{\beta_{GMM}} =$

$$[[\left(\Sigma X_i^{'} Z_i\right) W (\Sigma Z_i^{'} X]^{-1}][\left(\Sigma X_i^{'} Z_i\right) W \left(\Sigma Z_i^{'} Y\right)]$$

This will become your GMM estimate which is derived from this particular quadratic loss function which is again constructed by the sample moment condition.

(Refer Slide Time: 17:54)



$$\widetilde{\beta_{GMM}} = (X'ZWZ'X)^{-1}(X'ZWZ'Y)$$

So this is basically your GMM estimates that we are talking about and this is the GMM estimates that Arellano–Bond basically derived from this quadratic loss function. This is basically the idea of method of moment. GMM is a generalized version of method of moment when your number of instruments are actually more than number of parameters to be estimated and Method of moment condition is actually very simple if you go back which is derived from the idea that no explanatory variable should be correlated with my error term.

And if that is the case I can use that $E(X_iU_i)=0$ and if you replace U with the value then you will get a population moment condition given by equation 2 which is given by $E[X_i'(Y_i' - X_i'\beta)$

From there you just try to derive the sample moment condition then your $\hat{\beta} = (X'X)^{-1}(X'Y)$

With this I can show that OLS is basically a moment estimation.

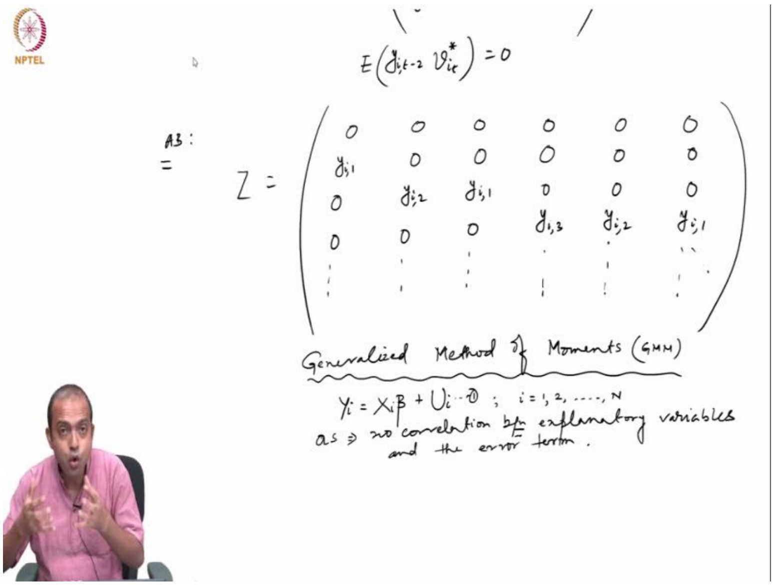
So, OLS as a method of moment estimation we can prove by using this estimation. Now, if we relax that assumption if we say that some of the explanatory variables are actually correlated with your error term then what happens you lose the consistency property that means probability limit N tends to infinity beta hat minus beta less than delta which is equals to 1 minus epsilon that is no longer are true. This is the consistency property. So, when that is the case that means when you have endogeneity issues in the model and we need to identify instruments that are highly correlated with your endogenous variable, but not at all correlated with the error term. Let us assume that Z as a set of instruments and we can assume two cases let us say that Z consists of L number of instruments in first case we assume L equals to K where K is number of parameters to be estimated from the model. Now, if K is less than L that means if you have more instruments than number of parameters to be estimated, then we can get a case of GMM. So, first we have derived L equals to K again using the expectation condition where $E(Z_iU_i)=0$. Instrument is not at all correlated with the error term. then again I replace U_i by$(Y_i - X_\iota' \tilde{\beta})$, and then by solving this we will again get $(Z'X)^{-1}(Z'Y)$.

When L is exactly equal to K, same OLS estimates that we have derived earlier in place of X we have written Z that is all, but the situation gets complicated when L is actually greater than K and we cannot select only K number of instruments out of that L because that will lead to inefficient solution.

To get an efficient solution what can we can do we can actually construct a quadratic function in this way where I am giving a weight, I am multiplying the sample moment condition its transpose within weight W_N where W is basically LxL matrix of weight choose an optimally that gives smallest variance GMM and then ultimately, we will arrive at a solution where $\widetilde{\beta_{GMM}} = (X'ZWZ'X)^{-1}(X'ZWZ'Y)$.

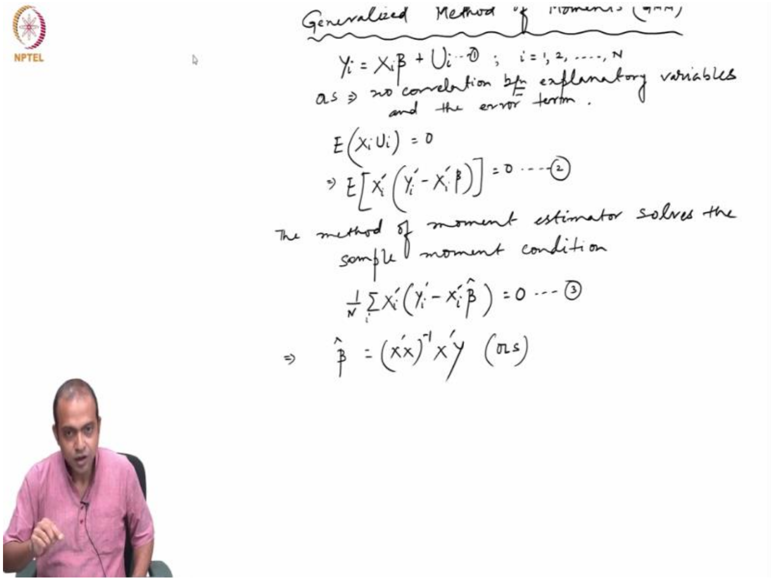
So, that means you see in between I have ZWZ' that means that instrument and that this weight matrix is playing a role to finally get you the GMM estimation in a case where number of parameters are less than number of instruments. Now why this is the case? Again, if you go back to our instrument matrix that we have derived in case of Arellano and Bond's model.

(Refer Slide Time: 23:17)



Look at here what Arellano and Bond suggested that we should use all the information available in the system we should use all types of all the moment conditions available in the system that is why he suggested you include all the lags available for each period. So, $y_{i,t-2}$ that is coming based on the idea of Holtz-Eakin they also suggested the same thing that second period lag. But additionally Arellano and Bond said apart from $y_{i,t-2}$, you see whatever lag is available for the untransformed variable. So, for endogenous variable only $y_{i,t-2}$ and above for predetermined variable it is basically your lag one is also available. So, if you include that then you see the number of instrument is 1 for second period, now 2 for third period, 3 for the fourth period and so on. So, you have enormous number of instruments. Its an over identifying system and in that case we discussed about method of moments and since this is a method of moment applied in a context where number of instruments are more than number of parameters to be estimated, that is called generalized method of moment that is all generalized method of moment. This is the idea of the Arellano and Bonds method of GMM.

(Refer Slide Time: 25:00)



We discuss little bit of theory of GMM also. So, that it can give you a better picture, better understanding of the Arellano and Bond model estimation which I will show you in our next class about the proper estimation of this AB model using a sample or data set. So, far then what we have discussed is we introduce a dynamic panel data model using a single left hand side variable which is dynamic in nature. And then we have also discussed how that particular model leads to persistence over a period of time because of two factors which is a_i unobserved effect and also the lag dependent variable which basically introduces the autocorrelation in the model and we prove that neither first difference nor fixed effect transformation is applicable. The moment we apply these two methods to eliminate the unobserved effect on the model, we eliminate a_i.

But at the same time we get a problem which is called biased estimation and inconsistent estimation because in both the cases in fixed effect transformation as well as first difference transformation we saw that my explanatory variable is still correlated with the error term even though error term they themselves are serially uncorrelated and then we discuss that if that is the case instead of using OLS in the transform model. What you can do when in your first difference equation you just use instrument for that difference variable which is getting correlated with the error term and then we discussed about three basically approaches started with Anderson–Hsiao who said that we can use either second period lag or third period lag, but the condition is that if we use more lag then your sample length will go down because for each panel we will lose observations either one or two depending on how many lags you are introducing.

So, to get rid of that problem there is a tradeoff between sample length and lag length. So, to get rid of that problem Holtz-Eakin et al, he said that instead of using too many lags we will use only second period lag for each time period and we will replace all the missing observation with 0 and in that way, we will maintain sample length, but if we follow that approach what is the problem. The problem is we will end up with having only second period lag that means we are not using all the moment conditions available in the system then Arellano–Bond said we can increase the efficiency of this estimate tremendously if we use all the moment conditions available, how to do it all the orthogonality condition or all the moment conditions available in the system. He suggested that you include all the lags available for the untransformed variable that means $y_{i,t-2}$ that is based on the IV style and then you have other lags available that is basically the GMM style. So two types of instruments are included here. So, for second period I have number of observation 1, from third period 2, fourth period 3 so as the number of period increases we have so many number of instruments.

It is an over identified system and in that context we introduce GMM generalized method of moment condition. When number of instruments are actually more than number of parameters to be estimated then what we will do in our next class we will take a data set and we will try to understand. So far theoretically we have estimated we have proved that neither first difference we can neither use OLS not fixed effect model to estimate a situation which is dynamic in nature because those things would be that will give you biased estimates.

We will show you what is the nature of the bias when we apply OLS in a dynamic panel data model, what is the nature of bias when we are applying fixed effect in that dynamic panel data model. We will take a dynamic panel data model and we will apply OLS and fixed effect transformation first of all to understand the bias. The magnitude of bias and then we will go to our next models this Arellano–Bond model to overcome that biasness.

So, that is basically about what we have discussed in our today's class. So, with this we are closing our discussion today and in our next class we will show the hands on experience empirical estimation of this dynamic panel data model. So far we are using please keep in mind that only the second lag of the untransformed variable. Second lag of the untransformed value that is what we had discussing over here.

But additionally, we can also take the second lag of the difference variable $\Delta y_{i,t-2}$and one more important thing is in Arellano–Bond methods since we are using instruments for endogenous variable, lag of endogenous variable, lag of predetermined variable. So, in AB approach all the instruments are coming from within the system since instruments are formed from lag of endogenous and predetermined variable in the model. So, previously when we discussed about IV estimation if you recall we were searching instruments from outside. If you recall in the context of IV we were discussing about estimating wage function.

Wage is a function of your experience, education and education we discussed that endogenous variable we were trying to replace education with father's education and mother's education which was not there in the system, which was not there in the model So, this father's education or mother's education was they were coming from outside the model to act as instrument, but here is a case $y_{i,t} = y_{i,t-1} + x_{i,t}$

My all instruments are coming either lag of this $y_{i,t-1}$or lag of the any of the predetermined variable or exogenous variables. So, that means this is a case of instruments coming from within the system. One advantage of this is that as you know the instrument requires two conditions to be satisfied. First of all they should be highly correlated with the endogenous variable. Obviously since instruments are coming from within the system $y_{i,t-2}$ would definitely be correlated with$y_{i,t-1}$, but when you have too many instruments then satisfying over identifying restrictions for each of these instruments becomes a challenge for you unlike the other case when the instruments were coming from outside the system satisfying the condition that the instruments are related with the endogenous variable. But not with the error term that it is a challenge here at least we can ensure that they are highly correlated, but there are too many in numbers and then we need to satisfy the over identifying restrictions for each of these instruments.