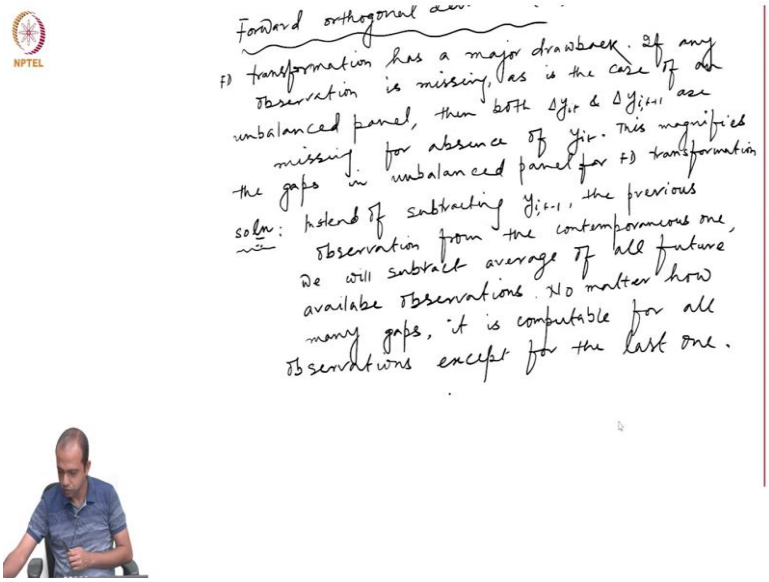
Applied Econometrics
Prof. Sabuj Kumar Mandal
Department of Humanities and Social Sciences
Indian Institute of Technology – Madras

Lecture - 49
Dynamic Panel Data Model - Part XII

(Refer Slide Time: 00:19)



Xtabond this command is more powerful than xtabond because xtabond 2 it has some additional features. So, what I will do I will first estimate the model same type of model system and difference GMM first I will start with difference GMM using the xtabond command. So, this is xtabond 2 and then what I need to specify is my dependent variable which is n and then I am taking its lag also nL1 and then let me take nL2 also look at now the difference between xtabond and xtabond 2 command.

In xtabond I was specifying only in after xtabond command but here I am putting nL1 and nL2 as well because I would like to use two periods lag right. So, that is the difference between the command of xtabond 2 and xtabond and then I will be taking w then wL1 then I will take capital and then kL1 and kL2 first and second total lag of capital and then I am taking ws also okay and then wsL1 ws sorry ysL2.

So, first and second order lag of industrial output and then what else I will take and then I will put y r that means your specific dummy y L start. Now in this xtabond 2 after specifying your basic model that means what I have specified my dependent and all my independent variables

I need to specify which are my GMM type instruments coming Arellano Bonds logic and which are the instruments coming from the standard logic.

Each of the predetermined variables which includes lag dependent and other exogenous variables they are actually they themselves or their instrument following that standard logic right. so. my GMMs type instrument I need to specify GMM and then nL1. So, this is my only instruments coming from that Arellano Bonds logic GMM type and then after that I will be specifying IV style.

That means IV within the bracket w all the variables w then wL1 then I have capital then Capital 1 kL1 and kL2 and then I have y s then I have ys of lag one ys of lag 2. And then I have year dummies. So, these are all my standard instrument. So, that means basically I am specifying which is my endogenous variable and what are the variables which are predetermine which may include lag dependent and exogenous variables.

And then to specify that I am considering only my difference equation I need to specify that I do not need no level no level equation. By specifying no level equation I am basically specifying a difference GMM I am asking status that I do not need the level equation to be considered. If you consider then it will become a system GMM. Since I do not need that I am specifying no level equation I also need robust estimates I also correct for small sample bias.

If you recall we initially said the dynamic panel that a model is best suited when t is small in is large. So, that is why I need to correct for small sample bias and then lastly I also need to put no cons command OK why no cons because this equation is in difference form. So, when the equation is in difference form obviously you do not get a constant for that I need to put no cons command right.

If you put that it is not able to recognize xtabond 2. So, I need to install. This command is also not taking ssc install xtabond2. Yes, now it is installed. So, I will put the command now. So, xtabond 2 so, please keep in mind if you have a license to version of this data and at any point of time if you get this unrecognized command then automatically you should install that using xtabond sorry ssc install and that particular command.

Now look at I got the same estimates but yeah so, this is basically the difference GMM but you see in the difference GMM the coefficient is 0.68 that means it is still lower than the lower limit specified by 0.73 a lower limit was specified by the fixed effect model. So, that means all my difference GMM you see the coefficient is not lying exactly within the interval as I have specified anyway for the time being let us see what I have installed here.

So, I have estimated this model this is this is this is what I have estimated the dynamic panel data estimation one step difference GMM right one step difference GMM and these are the coefficients and once again how many number of instrument is 41 and these are the instruments for first difference is difference of all these that is a standard. And what are the GMM type from lag 1 to lag 8 nL1.

Since I have specified my nL1 is the only endogenous variable then by putting that xtabond 2 command stator has actually considered from lag 1 to lag 8 as the instruments or for this nL1 these are the instruments right. And again you see Arellano Bond test AR1 which is there that means the z value probability get the P value is 0.000 and this is 0.606. What does it indicate basically it indicates that presence of AR1. What about the over identifying test that we are talking about we will look at the hands and test of over identifying restriction which is 0.177.

So, both the test Sergeant test as well as Hansen test they are test of over identifying restrictions. Please keep in mind you can go back to my earlier lectures on instruments instrumental variable estimates we need to be very careful about the null hypothesis of the Sergeant and Hansen test in both the tests actually the null hypothesis is that over identifying restrictions are valid right.

When over identifying restrictions are valid that means my null hypothesis should not be rejected while both the test ideally should produce results in such a way that my null hypothesis should not be rejected surgeon test indicate my null hypothesis is rejected but since this is not a robust test we will generally go for hands and test of over identifying restriction which is a robust test and that value P value is 0.77 that means my null hypothesis over identifying the restrictions are valid that is not actually rejected.

So, this is how we have estimated the difference GMM using the xtabond command what are the advantages of this command this command produces both the post estimation checkup results simultaneously I do not need to put again stat a bond then air test 4 separately. So, it is

giving desktop autocorrelation test of over identifying restrictions together right. Now once we estimate this model see so, far when we are taking estimating the model how we are transforming the variables?

We are transforming the variables in this way y i t - y i t - 1 this is called first difference transformation. Now this first difference first difference transformation it has a serious drawback what is that if any observation is missing if any of the y i t is missing then what happens actually uh it magnifies the gap in unbalanced panel. Unbalance panel means some of your observation might be missing.

We cannot assume for example in this particular context when you are discussing 140 forms data about the employment and all other variables it may. So, happen that a particular forms observation is missing in a particular year okay if it is missing then what will happen neither Delta y i t - 1 not Delta y i t + 1 will be available if the observation itself is not there how will you take its previous value and how will you take its next value.

If my observation for second period for a particular form is missing then when I am taking this first difference transformation then what will happen for the third period also I cannot get that right. So, first difference transformation if we apply in unbalanced panel then it basically magnifies the Gap because Delta y i t + 1 will be missing because Delta y i t itself is missing. So, to combat with this problem to combat with this problem there is a suggestion given by Arellano and Bower 1995 which is called forward orthogonal transformation.

This is very important please keep in mind forward orthogonal transformation that what we are discussing here is very important when you have unbalanced panel and trying to estimate it by fast difference transformation right. So, what we are doing here this is called forward orthogonal transformation or in sort FOD okay sorry this is not forward orthogonal transformation this is forward orthogonal deviation F O D forward orthogonal deviation this is suggested by Arellano and Bover 1995.

So, what they say basically that first difference transformation FD transformation FD transformation has a major drawback. What is the drawback if any observation is missing as is the case in an unbalanced panel as is the case of an unbalanced panel then both y i Delta y i t

and Delta y i t + 1 are missing for absence of y i t. So, basically what it says if y i t is missing then Delta y i t and Delta y i t + 1 both are missing.

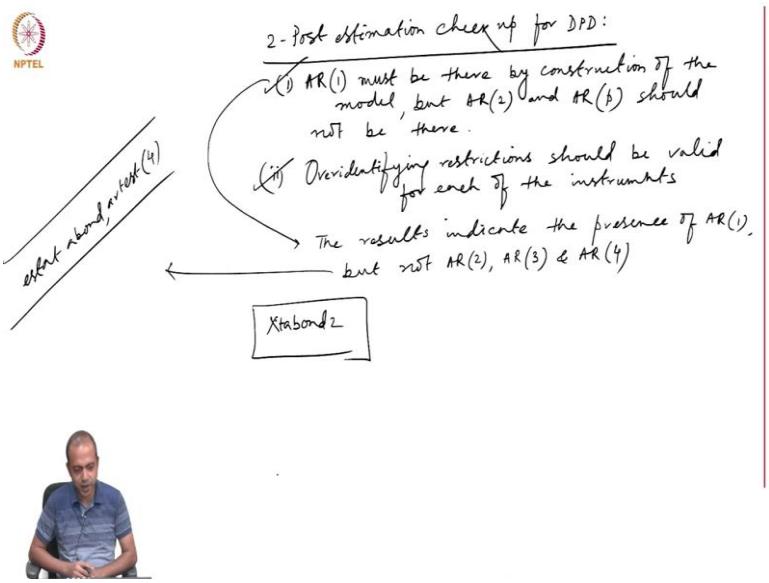
Then what will happen that means basically that means this basically magnifies the gaps in unbalanced panel foreign solution is; so in unbalanced panel for FD transformation. So, what is the solution is now that F O D what we are talking about what they say that instead of instead of subtracting y i t - 1 that means which is basically the previous observations.

The previous observations previous observations from the contemporaneous one contemporaneous is y i t from the contemporaneous one what will subtract we will subtract the average of all future observations okay all future available observations we will subtract average of all future available observations okay all available future observations. So, no matter how many gaps so, that means no matter how many missing observations are actually there it is computable.

Since I am taking average of all available observations it is computable for computable for all observations except for the last one because in the last period there is no future observations available. So, I cannot take FOD for the last periods observation otherwise last to last up to last observation whatever information is available I can take average of all the future available observations and then I can subtract that from y i t that is the difference.

So, instead of subtracting y i t - 1 from y i t which is basically magnifying the gaps when y i t is missing some y i t is missing what I am subtracting from y i t the average of all future observations. So, that way even if some of your observations are missing then it is computable for all the observations except the last one. So, in sort precisely what I am doing. So, what I am doing here.

(Refer Slide Time: 23:29)



So, my transformation will look like my transformation then will look like wit. So, what we will do w i t, so that way so, that way what will happen except the last one. So, it minimizes it minimizes it minimizes the data loss OK it minimizes the data loss and what is basically this FOD transformation to be very precise to be very precise to be precise then just one second.

So, to be precise my if w is a variable if w is a variable then the F O D transformation is I am denoting w i t + 1 with this sign this is called FOD transformation which is nothing but what I am doing c i t into w i t - 1 by capital t i t into sum of w i small s where small s is basically greater than t. So, all available future observations okay where c i t is basically equals to t i t divided by t i t + 1.

And what is T it capital T it? T it equals to number of all future observations number of number of all future observations. So, if you have unbalanced panel then we should use the F O D transformation in place of first difference transformation but if foreign is balanced then both FOD and if d produces produce similar results no need of going for F O D. So, now the question is how do you estimate this model with FOD transformation.

We will once again do that right. So, we will go to our model once again and then again that xtabond 2 xtabond 2 and then my dependent variable and again nL1 nL2 and then my W which is wedge W and then wL1 wL2 sorry wL1 and then I have capital then capital or for lag of order one order two then I have y s then y s 1 y s 2 okay then I have here. So, this is how I have specified my model then I have to put what is my GMM right.

That means what is my indigenous variable which is inL1 and then what is my IV? IV is all other variables that means w starting from w and wL1 and then capital and then capital of flag one kL1 then kL2 then y s and then y s L1 y is L2 right and then what else is there in my equation I have y r star. So, I will put wire star and then what we have to do is we have to put two step two-step robust standard order orthogonal orthogonal and no level equation.

So, this orthogonal the in that command box what is two step that I will discuss later on. For the timing please keep in mind our focus is on this orthogonal command because this orthogonal command you if you put then what is happening here we are taking forward orthogonal deviation. Orthogonal means FOD Transformations are used in place of first difference transformation right.

Now look at instruments for orthogonal deviation equation that is what we got right once again the nL1 7 2 0.72. So, that means this orthogonal deviations since some of the observations are missing here this is an unbalanced panel basically what is happening the moment we take orthogonal deviation in place of first difference transformation my coefficient of nL1 has increased and it is just closer to 0.73 it is 0.72 earlier if you look at my coefficient was 6 8. okay six eight. So, from 68 to I got 0.72.

This is the Improvement what I got just by using F O D transformation this is very important when you are working with unbalanced panel and in reality it is quite likely that many a times you may have to work with unbalanced panel. You should not expect the all the firms when you are working with a micro panel with several households several individuals several forms that each of these forms will available for each of the period there might be some missing observation.

And in missing observation it is dangerous to use first difference transformation because it enlarges the gap okay if in it enlarges the Gap if y height is missing both Delta y i t and Delta y i t + 1 both of them will be missing as a result of which what we should do is the F O D transformation that not only minimizes the data loss okay because I can now subtract what I am doing I am subtracting only average of all future observations.

So, even if some of your observations are missing I can always calculate average for all the periods except the last one and that improves the coefficient efficiency a lot. As you can see

my coefficient is moving just towards the theoretically specified lower and upper limit and then once again I got a Arellano Bonds AR1 which is again rejected indicating the presence of first order autocorrelation but second order autocorrelation is not there.

And then as far as over identifying test is concerned then Hansen test of over identifying restrictions is 0.17. So, that means what identifying restrictions are also valid I will talk about all other outputs what we are getting over here in our next class. So, with this basically we would like to close our discussion over here. So, what we have mainly discussed in today's class just the importance of that xt amount to command.

How xtabond 2 command can automatically estimate the model with the two positive estimation checkups presence of higher order correlation and overidentified restrictions all of them are readily available we do not have to put anything any specific command. For those secondly we have also talked about and another important concept which is fast in place of FD we have taken Ortho forward orthogonal deviation in case some of the observations are missing and it is applicable in presence of dynamic unbalanced panel. So, with this we are closing our discussion, thank you.