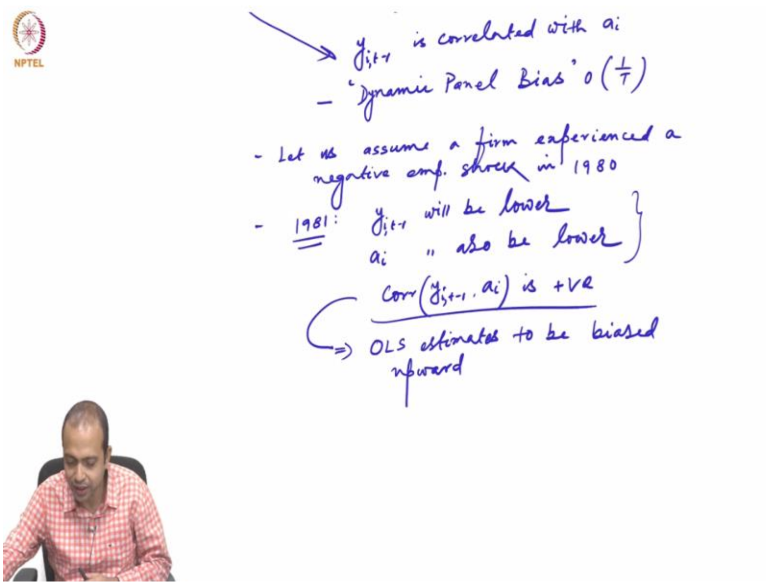
Applied Econometrics
Prof. Sabuj Kumar Mandal
Department of Humanities and Social Sciences
Indian Institute of Technology – Madras

Lecture - 45
Dynamic Panel data Model: Part VIII

Now the question is should this dynamic panel bias be upward or downward? Should OLS estimates of dynamic panel data model be a case of overestimation or underestimation.

(Refer Slide Time: 00:31)



To understand this let us assume a firm experienced a negative employment shock in 1980 a negative employment shock. So, if that is the case what will happen in 1981? In, 1981 if this is a negative employment shock in 1981 your y i t - 1 will be lower ai will also be lower then what you observe in 1980. There is a negative employment shock. So, throughout the study period these negatives impact of this negative shock will be there.

So, that means that means what will happen what is ai basically? ai is the unexplained employment effect which you cannot explain which is unabsorbed, unexplained employment effect will be lower than the sample average because as a result of this negative employment shock. So, if you consider only one period there is a shock in 1980, negative employment stock as a result of which when I go to next period the model will try to adjust.

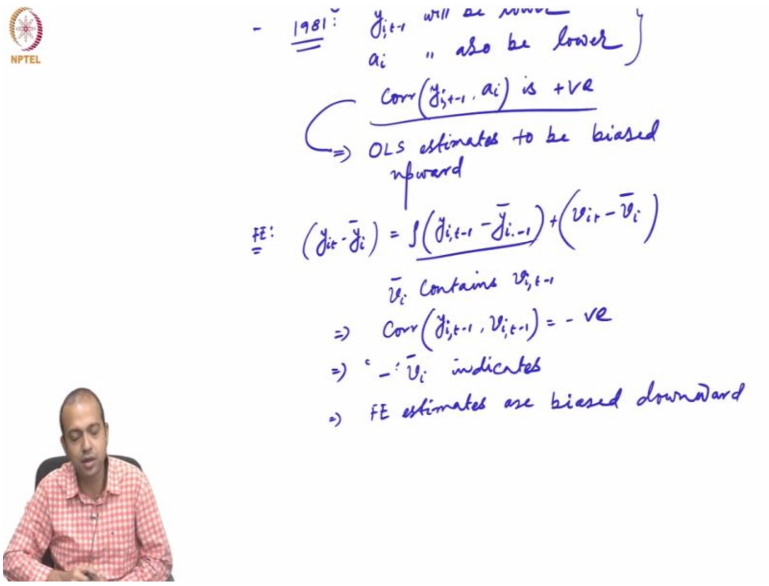
How it will adjust y i t - 1 will be lower and this a i will also be lower that is how lower than the previous period that is how it will slowly adjust. So, since y i t - 1 is lower a i is also lower.

So, that means there is a pattern a similar pattern that we observe in y i t - 1 and unobserved effect a i and as a result of which correlation between y i t - 1 and a i is actually positive a positive correlation between a i and y i t – 1.

And because of this positive correlation OLS estimates biased upward. So, this positive correlation basically leads to OLS estimates to be biased upward. OLS estimates to be biased upward and that is what that means this value what I am getting 1.04 is actually a biased estimate which is biased upward. We can easily check that it is not less than 1 for dynamic stability. So, we can understand this is biased upward.

Now what will happen if we apply fixed effect transformation. Fixed effect transformation will also lead to biased and inconsistent estimate we have already proved, but the question is will this fixed effect estimates be biased upward or downward that we need to understand. So, to understand that we will once again derive the fixed effect transformation.

(Refer Slide Time: 04:43)



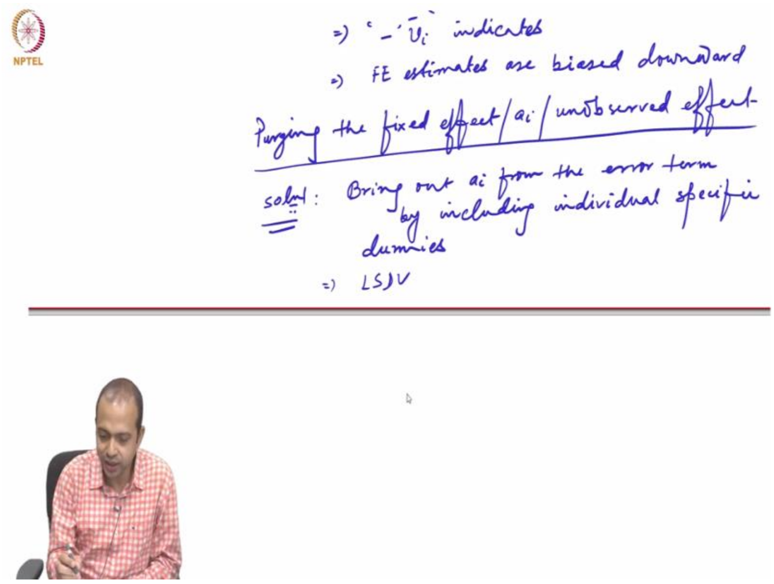
So, FE transformation is basically y i t - y i bar equals to rho y i t - 1 – y i dot - 1 bar and what is this? We basically say this is the mean of this value + a i will get canceled out and then you will have v i t plus minus v i bar and we said this v i bar contains v i t - 1 and that is again correlated with y i t – 1, but this negative sign before v i bar means the nature of the correlation is negative.

So, what I am, saying v i bar contains v i t - 1 and that is the reason there is a correlation between y i t - 1 and v i bar or sorry we can write v i t - 1 and that means we can say that actually this entire variable y i t - 1 is actually correlated with this transform error term which is v i - v i bar. So, this results in y i t - 1 is correlated. So, this is correlated with the v i t - 1 so that means we can say that this is correlated with this.

And since there is a negative sign the negative sign before v i bar indicates that this correlation is actually negative, this is negative because of this negative correlation FE estimates are biased downward. So, this is also dynamic panel bias only thing is that this is bias downward unlike the OLS estimates which are biased upward that is the reason. So, what you will do we can do a couple of things over here to understand the consequences.

See when we have this type of unobserved effect present in the model before estimating the direct fixed effect transformation what we can do.

(Refer Slide Time: 08:20)



So, this is basically we are trying to purge this is basically purging the fixed effect or a i or unobserved effect. How will you do that? Solution 1 what you can do? We can actually bring out a i from the error term by including individual specific dummy. So, what we are doing here all these problems are arising because this fellow y i t - 1 is getting correlated with a i and that is leading to some kind of dynamic panel bias if we apply OLS.

As we saw already that it is upward even when we are trying to transform this fixed effect then also there is some kind of bias, but it is downward. So, what we are trying to do over here? We are trying to understand that when you have a i in the model there are couple of ways by which you can actually get rid of this unobserved effect. The solution number one is you bring out a i from the error term how do you do?

You consider that individuals are different and how do you model the situation? You simply include individual specific dummies in your model and then actually that will lead to LSDV situation. Least square dummy variable we all know from our earlier discussion on panel data that older version of the fixed effect model is basically LSDV. What does LSDV do? LSDV basically brings out the a i component from the error term by including individual specific dummies in the model.

Why you are doing this because sum of my explanatory variable y i t - 1 was correlated with the error term basically with the a i. So, if you do that then you will get the LSDV model. **(Video Starts: 11:46)** And how do you estimate that LSDV model? Again what we can do we can regress your dependent variable is n and then you have so many n L 1 and then again n L 2 and then you will have w what is your w is wage this is w and then w L 1 then you have capital then lag of capital then second lag of capital and you have output y s.

Then first lag of output second lag of output and then you have your dummies and then you need to include individual specific dummy which is i dot id if you remember i dot id command if we put in strata that will include individual specific damage in the model and we will get least square dummy variable model. So, this is my model. So, if you do that then look at what is happening?

What is the variable 0.73. So that means the coefficient of the lagged employment is now 0.73 so which is just the opposite of the previous case. In the context of OLS it was biased upward. In the context of FE transformation my estimates are biased downward why it is downward? Because as I explained y 1 t - 1 is correlated with v i t - 1 which is there in v i bar and since there is a negative sign over there before v i bar this correlation is negative.

And this negative correlation is leading to a situation where estimates are biased downward. So, what we have learned over here? We have learned two things. First of all how to bring that a i component out from your error term to estimate an older version of the fixed effect model which is least square dummy variable model by putting i dot id command and if we do so then we will get the coefficient which is less than 1 definitely at least dynamic stability is maintained, but it is biased downward.

Now there is another way what we can do. Another way of estimating this is basically we can we can directly put x t reg command x t reg and then again your n then n L 1 then n L 2 then w then w L 1 lag of which is w L 1 and then you have capital then first lag and second lag of capital then you have y s industrial output fast and second lag of that and then you have your dummy and you have to put FE.

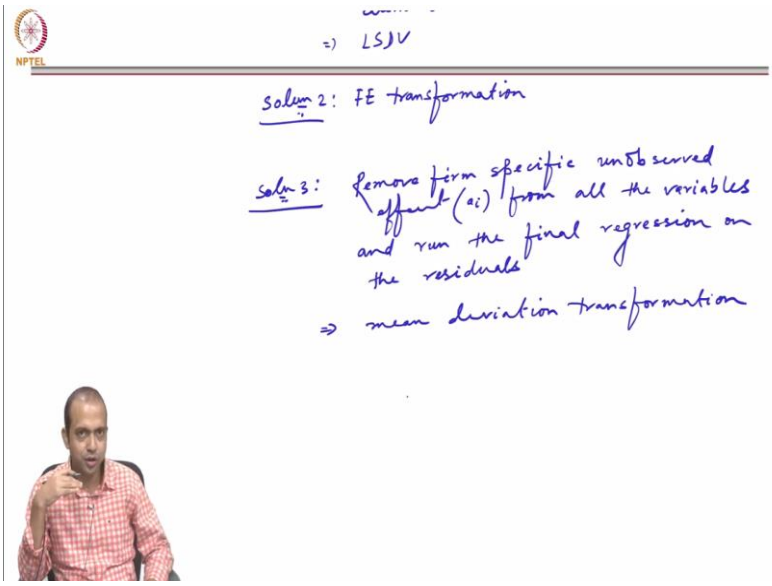
So, instead of regress this is the same standard fixed effect transformation, fixed effect model. So, basically what we are doing we are if you do so then again similar estimate look at 0.73294. So, this is the newer version of the fixed effect model LSDV is basically older version of the fixed effect model. So, what do we have to observe over here look at the F test which is, basically 1.89 and highly significant.

And what is a test if you recall in the context of fixed effect model, we discuss this F test basically says whether the unobserved effects are significant or not. So, that means combinedly status u i is basically the a i what I am denoting over here. So, a 1 equals to a 2 equals to a 3 dot, dot, dot this a n they are combinedly significant or not. So, join significance test for this unobserved effect given by this F test.

And that is also highly significant indicating the presence of significant unobserved heterogeneity in the model. So, this is second way of bringing out the a i component from the error term because that is actually creating the problem over here because my y i t - 1 was correlated with this. Now both the models shows my estimates are actually biased downward. **(Video Ends: 18:00)**.

There is another way another solution is third solution is that removing solution 2 is; so this is solution 1 LSDV.

(Refer Slide Time: 18:13)



Solution 2 is basically fixed effect transformation and solution 3 what we can do. We can remove firm specific unobserved effect or a i from all the variables and then and then run the final regression on the residuals. So, if we do so that is basically a mean deviation transformation I am talking about that means y i t - y i bar kind of model why you are doing this because if we do so a i will get eliminated because a i does not change over time.

And mean of a i is actually a i so a i will get eliminated. This is another way and how do you estimate this type of model? What we have to do? **(Video Starts: 20:36)** We have to use data's x t data command so this is x t data then again n then n L 1 then n L 2 then w and then w L 1 and then after that we have capital first lag of capital, second lag of capital and we have output then first lag of output, second lag of output and then y r star.

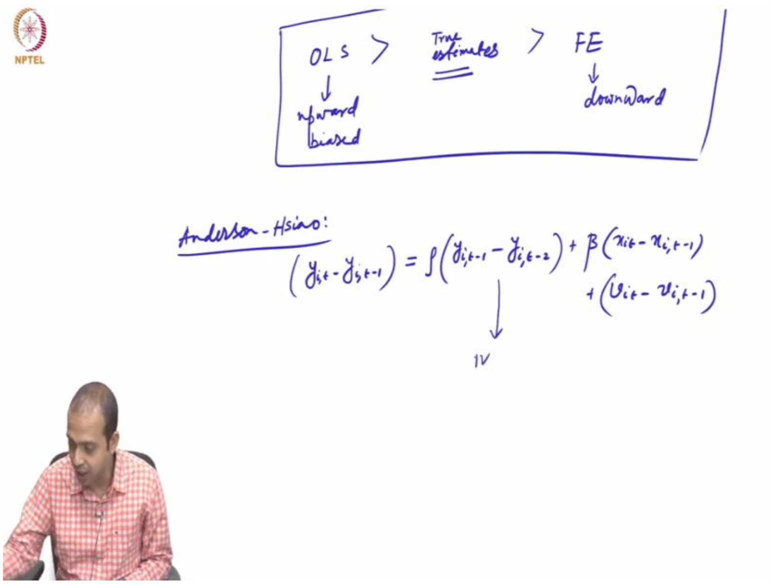
And then we will put FE. So, by putting this command what I am doing actually I am removing the a i component from all the explanatory variable. I am removing firm specific unobserved effect from all the variables and then I will be using the residuals of all those variable in my final regression. So, it is nothing, but x i t - x i bar. So, x i t minus that unobserved effect.

So, I am removing all these from my explanatory variable. This is we are doing as x t data command and then I will be using regress n and again n L 1, n L 2 and then w and then w L 1 then I have capital then capital 1, capital 2 then y s, y s 1, y s 2 and then I have y r dummy. So, look at this same coefficient 7.7329 here also it was 0.7329 and then here also it is 0.7329.

So, that means what we have proved today that when we use OLS and fixed effect transformation in a dynamic panel data model that gives rise to a dynamic panel bias why it is happening? Because in the OLS model we are ignoring the fact that y i t - 1 is actually correlated with a i and that positive correlation is leading to upward bias. In fixed effect transformation what is happening?

Y i t - 1 is correlated with - that v i bar because v i bar contains v i t - 1 which is actually correlated with y i t – 1 and a negative sign before v i bar is actually leading to underestimation of the fixed effect transformation. **(Video Ends: 24:42)** So, one thing we have learned here we can write that fixed effect estimates are biased downward.

(Refer Slide Time: 24:56)



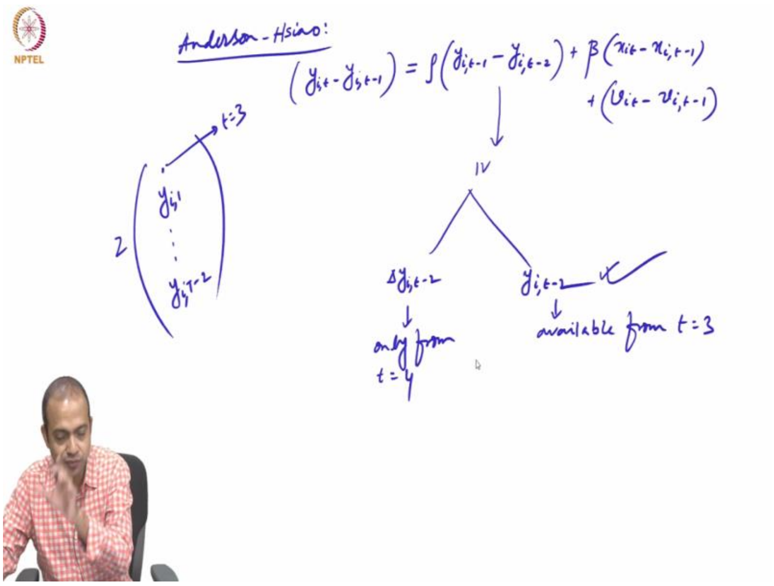
So, what we observe that you have OLS here, you have FE over here. FE you can get by three alternative ways either by estimating LSDV or by estimating direct FE or by removing the firm specific effect from all these explanatory variables and then running this model. So, this is upward bias, this is downward bias so that means the true estimates should be actually between since it is upward it should be less than this, but it should be greater than this.

So, true estimates then must lie between OLS and FE estimates. So, OLS gives upper bound, FE estimates gives downward or lower bound of the estimates in a dynamic panel later model and we can always check whether my true estimates are actually lying between OLS and this or not. So, this is a very good post estimation checkup whatever dynamic panel data model we estimate after estimation we can very well check whether the estimates are lying between OLS and FE it is lying.

Then only we can say that at least theoretically my estimates are satisfying the upper and lower bound that is what I am talking about. So, when neither a FE nor OLS cannot be applied in the context of dynamic panel data model then what is the solution? Solution is if you recall we need to go for Anderson and Hsiao approach and once again what is Anderson and Hsiao approach?

So, he suggested first you take the first difference of the model so y i t - y i t - 1 equals to y i t - 1 - y i t - 2 + beta x i t - x i t - 1 + v i t - v i t – 1. So, since y i t - 1 is correlated with v y t - 1 we need to use IV that is what Anderson and Hsiao suggested an IV can be of two types.

(Refer Slide Time: 28:39)



So, what should be the IV? First one is delta y i t - 2 or y i t - 2. Now advantage of using y i t - 2 is that it will give you enough sample length because if you take y i t - 2 that observation would be available only from only from t equals to 4, but when we are taking y i t - 2 then it is available from t equals to 3. If you go back and see our instrument matrix the first observation is available from the third period y i 1.

Anderson and Hsiao approach the Z matrix if you recall that first observation is lost and then this is y i 1 then dot, dot, do y i t - 2. So, this is for t equals to 3, but if you take delta y i t - 2 that means my first observation would be available only from t equals to 4. So, at least one more period of observation we can have if we use the second order lag of the untransformed or level variable y i t – 2.

And when your sample size is small then one extra period also matters in the efficiency of your estimates. So, that is why we will be estimating our first dynamic panel data model using Anderson and Hsiao approach by taking y i t - 2 as instrument for y i t - 1 - y i t – 2 that is the first approach we will estimate and we will see following Anderson and Hsiao approach whether my estimates are lying between OLS and FE.

If it lies between this interval at least the first criteria is satisfied that my estimates my dynamic panel data model estimates are lying within the theoretically defined upper and lower bound. If it does not then we cannot use those estimate because first of all they are not lying within the theoretically defined bound. It must lie between the two. So, we will be using this Anderson and Hsiao approach the first model of dynamic panel data model in our next class.

And we will check whether that is satisfying the upper and lower bound or not. With this, we are closing our discussion today. In our next class once again we will be using the same data set and we will try to estimate this Anderson and Hsiao approach basically instrumenting y i t - 2 for y i t - 1 - y i t – 2 that is what we will do. We will not be using delta y i t - 2 to start with because as I said that is available only when t equals to 4.

This is available from t equals to 3 one extra period also matter. With this, we are closing our discussion today. Thank you.