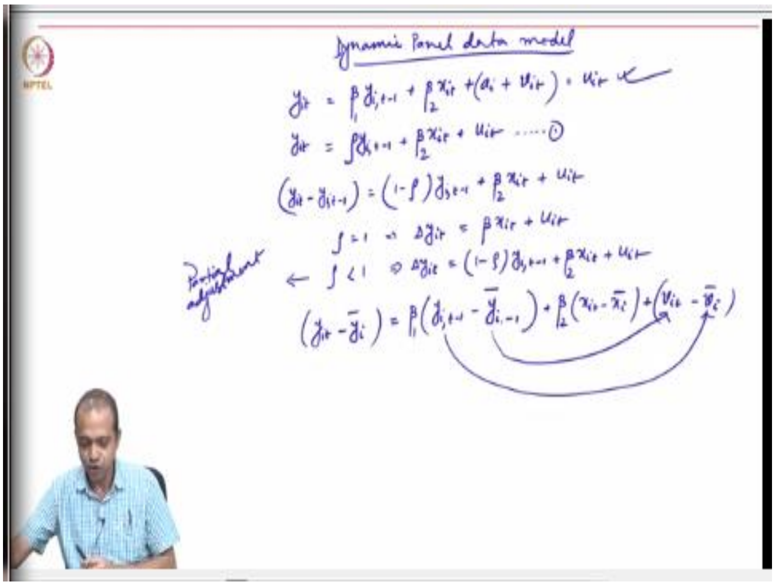
Applied Econometrics

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

Lecture-40
Dynamic Panel Data Model-Part III

(Refer Slide Time: 00:25)



To recap, why dynamic panel data is required? Because if you recall we say that many of the economic relationships are dynamic in nature. For example, in the context of agriculture when the farmer is deciding about how much land to be allocated for i^{th} crop in t^{th} period that basically depends on previous year's production. Because previous year's production would have resulted in some price followed by some amount of profit and that will motivate the farmer to decide about the land allocated for the same crop this year. So, a heavy profit or huge profit for the i^{th} crop last year may again motivate what the farmers to go for larger amount of land to be allocated for the same i^{th} crop in t^{th} period as well. And similarly, we have also discussed how in the context of a farm production, if the farm would have produced more in the last year and could not sell much in the market, so much of the production would have gone for inventory and that will reduce the amount of production this year. Similarly in the context of income for a country in India's income or GDP last year will basically determine what would be the GDP for this year as well.

Many a times when you apply for job market your employer will ask you what is your last scale that you used to draw in your previous organization. So, that means while fixing your grade or payment in the new organization your new employer will have a look on your previous salary.

So, these are all many examples that we can draw upon to show that how economic relationships are dynamic in nature. And how do you model this dynamic relationship? just by adding the lagged value of your dependent variable. So, that means we model dynamic panel data in this way: $y_{it} = \beta_1 y_{it-1} + \beta_2 x_i + a_i + \vartheta_{it}$ y , where $(a_i + \vartheta_{it})$ is basically the composite error term which is again is u_{it}. Sometimes instead of using the co-efficient β_1, sometimes we use a different coefficient just for the notation's sake to differentiate this variable with the other explanatory variables.

So, let us say this is: $\rho_1 y_{it-1} + \beta_2 x_{it} + u_{it}$

Then we said that this dynamic panel data model there is another name for this model which is called partial adjustment model. And to understand why this is called partial adjustment model? Then we can just subtract y_{it-1} from both sides of the equation. So, if we subtract y_{it-1} from both the sides of equation 1 what you are going to get is: $y_{it} - y_{it-1} = (1-\rho)y_{it-1} + \beta_2 x_{it} + u_{it}$

And see the coefficient (1- ρ) basically indicates when ρ equals 1, there is no lag dependent variable in this equation. So, simply we can say that $\Delta y_{it} = \beta \ \ x_{it} + u_{it}$

So that means there is no presence of lag dependent variable in the model. So, the dynamism goes off. So, if we want to have dynamism in the system then the value of ρ should be less than 1 and that implies this model is $\Delta y_{it} = (1-\rho)y_{it-1} + \beta_2 x_{it} + u_{it}$

So, one thing we have to clearly keep in mind since ρ equals 1 only gives the dynamic nature of this model. ρ less than 1 indicates partial adjustment. So it is not a fully adjustment model.

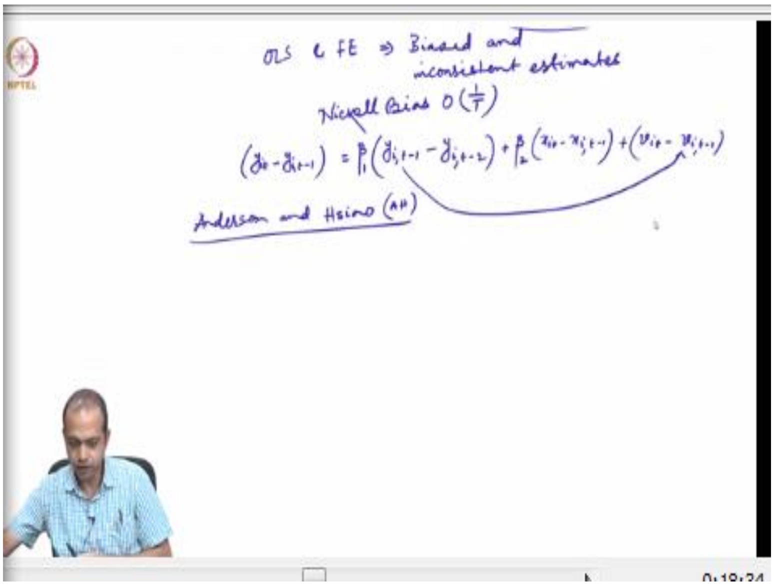
If anything happens to my x_{it}, that will result in a change in y which will be partially adjusted. Some part of the adjustment will happen in the next period also. That is the meaning of partial adjustment. ρ equals 1 is called fully adjustment model and when it is fully adjusted that means for any change in your explanatory variable y changes itself in the same period. So, this change is not getting carried over in the next period there is no dynamism in the system. So, to have a dynamic panel at a model I must have ρ less than 1 and that is why it is called partial adjustment

And then we have also discussed briefly in a model like $y_{it} = \beta_1 y_{it-1} + \beta_2 x_{it} + a_i + \vartheta_{it}$ there are 2 sources of persistence over a period of time.

The first persistence is due to the presence of this unobserved effect which is a_i and why it is giving persistence? Since y_{it} is actually related to a_i and a_i does not have a 't' subscript, that means we can say that y_{it-1} is also related to a_i. That means in each period one of my explanatory variables, the lag dependent variable is related to a_i. And that is one source of persistence, so that means in each period it is related.

And since a_i component is common in the error term we can say that in successive periods a_i is all your error terms are getting correlated and leading to autocorrelation problem. Autocorrelation arises for another reason in this dynamic panel data model since y_{it} is related to $y_{it-1}$and ϑ_{it}, then y_{it-1} is also related to ϑ_{it-1} . Same equation can be written for y_{it-1} .So, that means I can say that y_{it} is related to y_{it-1}as well as ϑ_{it-1} since y_{it} is correlated with ϑ_{it}. So, that means I can now easily say that ϑ_{it} is also getting correlated with ϑ_{it-1} through the channel of ϑ_{it}. And then we said that what is the solution, that means we cannot estimate this model by the OLS we must go for the fixed effect transformation to remove the a_i.

But if you go for fixed effect transformation then as we have discussed a_i will get canceled out and then what will happen I will get let us say $\beta_1(y_{it-1} - \overline{y_{\iota.-1}}) + \beta_2(x_{it} - \overline{x_\iota}) + (\vartheta_{it} - \overline{\vartheta_\iota})$
Then we said that this fixed effect transformation also does not work because this $\overline{\vartheta}_\iota$ is nothing but $\sum_{t=1}^{T} (\vartheta_{it})/T$
So, this $\overline{\vartheta}_\iota$ includes $\vartheta_{it-1}\, which$ will be related to y_{it-1}. That is why this fixed effect transformation cannot rule out the correlation between your one of your explanatory variables with this transformed error. At the same time when you calculate the mean of $y_{it-1} - \overline{y_{\iota.-1}}$ which is nothing but $\overline{y_{\iota.-1}}$that also includes y_{it} and that y_{it} would be correlated to this ϑ_{it}.
So, in both the cases we have established that $\overline{\vartheta}_\iota$ is correlated with y_{it-1}. because $\overline{\vartheta}_\iota$ includes ϑ_{it} that is why that is correlated with y_{it-1}. Then we have discussed, so when fixed effect transformation is not applicable to estimate this type of dynamic panel data model, we can actually go forward. So, neither OLS nor fixed effect actually is applicable, straightforward OLS application is not possible, fixed effect transformation is also not applicable; both of them will lead to a bias and inconsistent estimate.
(Refer Slide Time: 14:33)

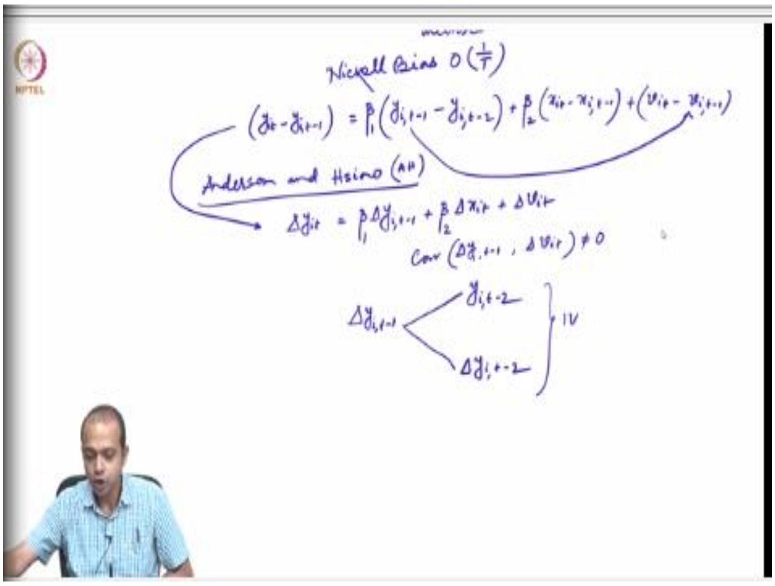


So, OLS application in the original equation and FE transformation will lead to biased and inconsistent estimates. And that bias we discussed that is called Nickell Bias which is of order 1 by T. So, we cannot apply OLS and fixed effect transformation because both of them will lead to some kind of bias and that bias is known as Nickell bias because Nickell was the first econometrician to introduce this type of bias which is of order T. That means for a small sample this bias would be a great amount and we also said that even when T equals 30 that bias is amounted to almost 20% of your original estimates. So, when fixed effect transformation and OLS does not work then what is the solution available? So, we may go for the first differencing, so if we go for first differencing basically that means $y_{it} - y_{it-1} = \beta_1(y_{it-1} - y_{it-2}) + \beta_2(x_{it} - x_{it-1)} + \vartheta_{it} - \vartheta_{it-1}$

And once again we can say that this $y_{it-1}$is correlated with ϑ_{it-1} so that means this is correlated with the error term. So, you can first differencing is also not able to solve the problem of endogeneity because one of my explanatory variables is getting correlated with the error term. Now that much we have discussed in your previous discussion, so what is the solution then?

The standard solution is basically the utilization of instruments. As we all know when to solve endogeneity problem -when you have some of your explanatory variable is correlated with the error term then use the instrument. And this particular idea first was given by 2 famous econometrician Anderson and Hsiao in short AH approach.

(Refer Slide Time: 18:35)



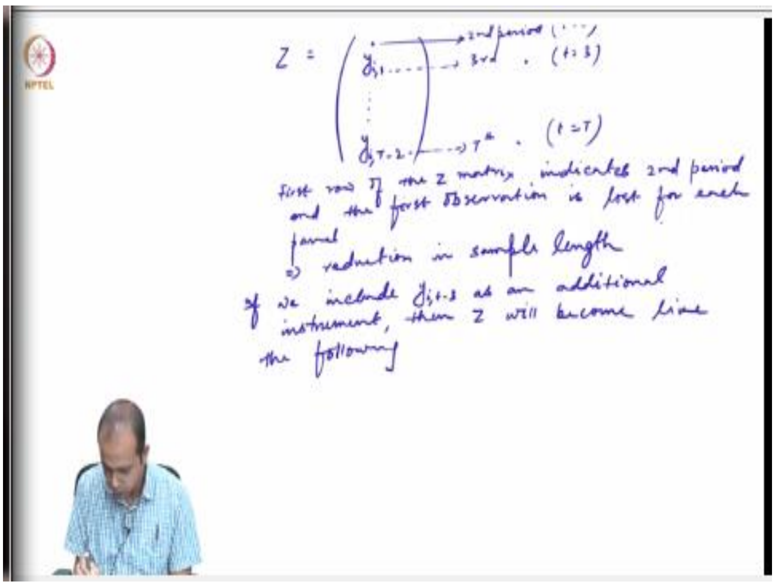
Basically, that means this model turns out to be

$$\Delta y_{it} = \beta_1 \Delta y_{it-1} + \beta_2 \Delta x_{it} + \Delta\vartheta_{it}$$

So, when correlation between Δy_{it-1} and $\Delta\vartheta_{it}$ is not equal to 0, then the idea is we need to use instruments for this Δy_{it-1}. We can use 2 types of instruments basically based on an Anderson Hsiao.

They say that you can use either second lag of the difference variable that means y_{it-2} or Δy_{it-2} Second lag of the level or second lag of the difference both of them can be used as IV. That is the idea given by Anderson and Hsiao.

(Refer Slide Time: 20:32)

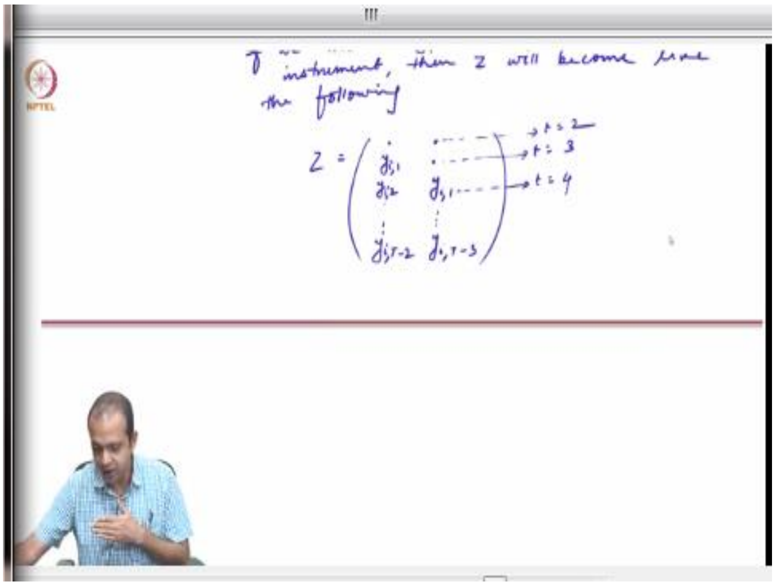


So, in this case if we use let us assume that we are using y_{it-2}. Which is the second lag of the level variable we are using as instrument. And then in the Anderson and Hsiao approach my instrument matrix will look like this.

$$Z = \begin{pmatrix} y_{i1} \\ \vdots \\ y_{i,T-2} \end{pmatrix}.$$

So, the first row of this instrument matrix indicates second period and that observation is lost. So, we need to remove that observation for each panel and that will lead to, so what I will say that first row of the Z matrix indicates second period and the first observation is lost for each panel. And this implies a reduction in sample length. So, first observation is lost for each panel when your Z matrix is like this, what we are doing? Based on the idea given by Anderson and Hsiao we are taking y_{it-2} as the instrument. So, that means the first observation is available only from 3rd period to be used in the regression and if we include third period also if we include y_{it-3} as an additional instrument then Z will become like the following.

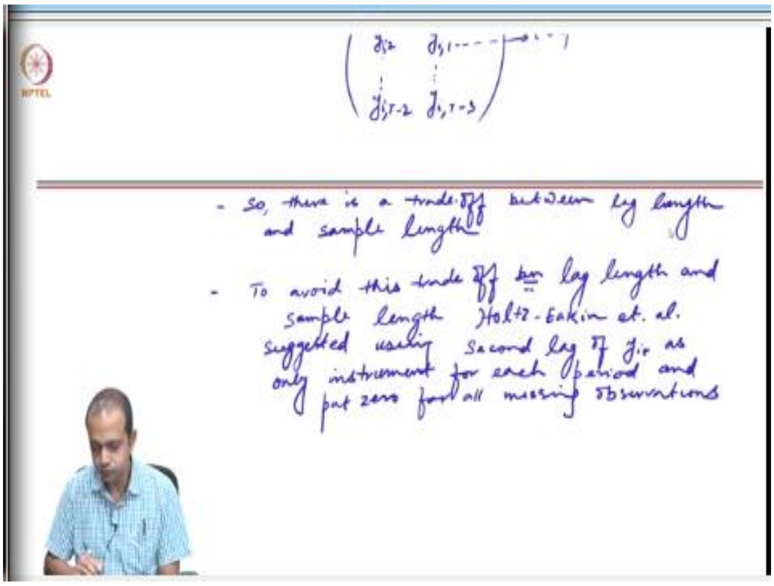
(Refer Slide Time: 25:04)



When you use the 3rd lag also as additional instrument we will lose one more observation and then this will become y_{it-3} . So, that means this is for t equals 2 so this is for t equals 3 and this is called t equals 4. So, that means I lost 2 observations if we include third lag also as an additional instrument. So, that means here what is happening?

In the previous Z matrix I was using only second lag. So, that means using second lag means I am using only one orthogonality conditions. And when we are using a 3rd period lag also as an additional instrument my Z matrix looks like this. And as a result of which what is happening? I am getting more instruments, more orthogonality conditions but at the same time I am having less number of observations available for my estimation. So, that means there is a trade-off between what you can write that means which implies.

(Refer Slide Time: 27:40)



So, there is a trade-off between lag length and sampling length**.** If we use only y_{it-2} as instrument then my lag length is only t equals 2, so that means I have preserved sample length. sample is not getting reduced but lag length is only 2. so that means I am using one single orthogonality condition, I am not utilizing all the available information available in the system. And the moment I include additional instrument as y_{it-3} then I am losing 2 observations for each panel. So, this basically says that there is a trade-off between lag length and sample length. And to avoid this trade-off between lag length and sample length Holtz-Eakin et al. suggested that instead of using many more lag as instruments why not using only one instrument y_{it-2}. For each period, second lag of the untransformed variable as instrument for each period and replace all these missing observations by 0. So, that means suggested that using second lag of y_{it} as only instrument for each period and put 0 for all missing observations.