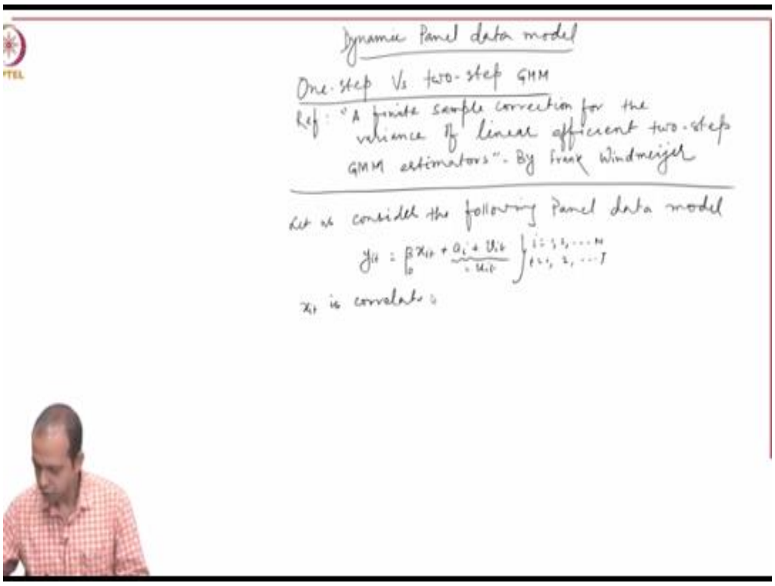
Applied Econometrics
Prof. Sabuj Kumar Mandal
Department of Humanities and Social Sciences
Indian Institute of Technology-Madras

Lecture - 54
Dynamic Panel Data Model - Part XVII

Welcome to our discussion on dynamic panel data model. And today what we are going to discuss is one-step versus two-step GMM estimation. These are important models. So let us learn little more on one-step versus two-step GMM.

(Refer Slide Time: 00:37)



This is a dynamic panel data model, one-step versus two-step GMM. This discussion is based on this important reference given by Frank Windmeijer and the title is "A finite sample correction for the variance of linear efficient two-step GMM estimators".

Let us consider the following econometric model, following panel data model wherein $y = \beta_0 \chi_{it} + a_i + v_{it}$ and $a_i + v_{it} = u_{it}$, and assume that $i = 1,2, \dots, N$ and $t = 1,2, \dots T$.

(Refer Slide Time: 04:30)

$y_{it} = \beta_0 x_{it} + u_{it}$ $t = 1, 2, \ldots, T$

x_{it} is correlated with the a_i but predetermined with v_{it}

$\Rightarrow E(x_{it}\, v_{it+s}) = 0 \quad \forall\, s = 0, 1, \ldots, T-t$
but $E(x_{it}\, v_{it-s}) \neq 0 \quad \forall\, s = 1, 2, \ldots, t-1$ } Predetermined

Arellano & Bond (1991) estimator is GMM estimator in the first difference,

$\Delta y_{it} = \beta_0 \Delta x_{it} + \Delta u_{it}, \quad t = 2, \ldots, T$

with $T(T-1)/2$ Sequential instrument

And χ_{it} is correlated with the past value a_i but predetermined, this is very important, predetermined with the v_{it}. What does it mean? It means that this implies $E(\chi_{it}\, v_{it+s}) = 0\ \forall\, s = 0, 1, \ldots, T - t$. Now if we put $s = 0$ that means $E(\chi_{it}\, v_{it}) = 0$.

If χ_{it} is not correlated with the present error term, but $E(\chi_{it}\, v_{it-r^o}) = 0$ where $r = 1, 2, .., t - 1$. Now if we put $r = 1$ it means $E(\chi_{it}\, v_{it-1}) \neq 0$, this means the explanatory variable is correlated with the past error but not the present and future. That is why it is called predetermined.

And in case if we combine these two and both are not equals to 0 type then that is a situation of endogeneity wherein, the explanatory variable is correlated with the present as well as the past error and if neither of them is nonzero that means in presence of $E(\chi_{it}\, v_{it+1}) = 0$ and $E(\chi_{it}\, v_{it-r}) = 0$ that is a situation called exogeneity, a strict exogeneity.

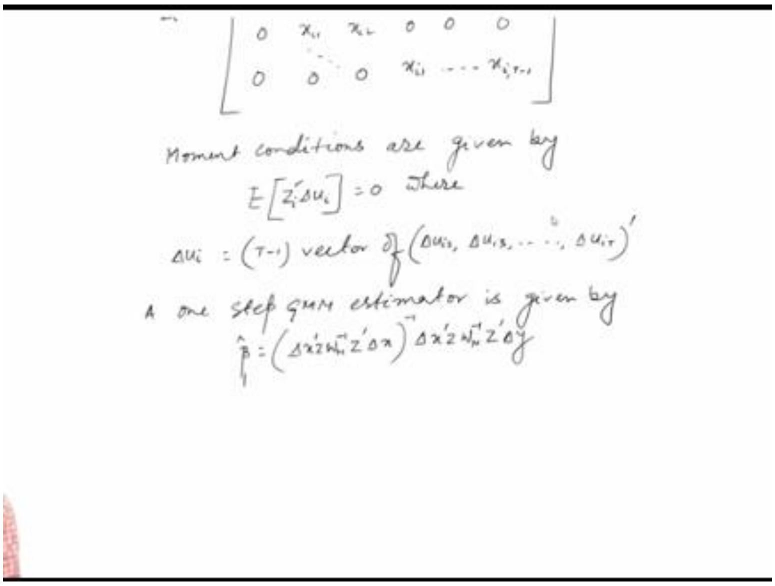
This means in a variable which is strictly exogenous it does not depend on the error in past, present and future. That is called strictly exogenous variable. An explanatory variable which is not at all correlated with the past, present and future error. When it is predetermined, it is not related with the present, but it is related with the past, this is called predetermined.

And endogenous variable is related with the present as well as the past. That is the difference between predetermined, endogenous and exogenous variable. If that is the case if you recall, we started our discussion of dynamic panel data model assuming that we have a right-hand side explanatory variable which is not strictly exogenous. That means, it is predetermined or endogenous.

So, if that is the case, in place of χ_{it} you can very well assume Υ_{it-1} which is an Arellano and Bond 1991 estimator, GMM estimator in the first difference. That we have already discussed, in the first difference, that is, $\Delta\Upsilon_{it} = \beta_0 \Delta\chi_{it} + \Delta u_{it}$, and $t = 2, .., T$.

So, how many moment conditions with $T(T-1)/2$ sequential instruments? One condition is $T(T-1)/2$ sequential estimates instruments that we are using.

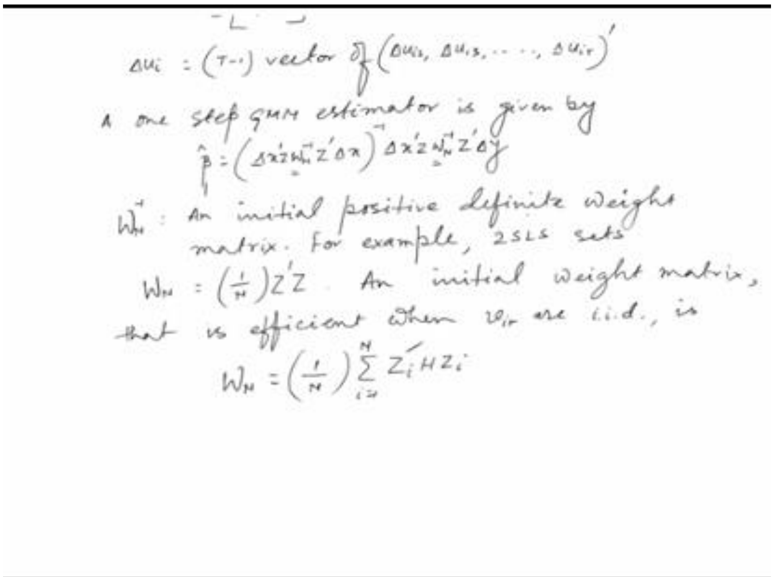
(Refer Slide Time: 10:40)



And then if you recall our Z matrix that means, the instrument matrix was this $\begin{bmatrix} 0 & x_{i1} & x_{i2} & 0 & 0 & 0 \\ 0 & 0 & 0 & x_{i1} & \ldots & x_{i,t-1} \end{bmatrix}$. In our instrument matrix we assumed moment conditions are given by expectation. What was the moment condition? Moment condition is $E[Z_i' \Delta u_i] = 0$,this means, the instruments are not correlated with the differenced error structure. That was the moment condition where $\Delta u_i = (T-1)$ vector of this $(\Delta u_{i2}, \Delta u_{i3}, \ldots, \Delta u_{it})'$.

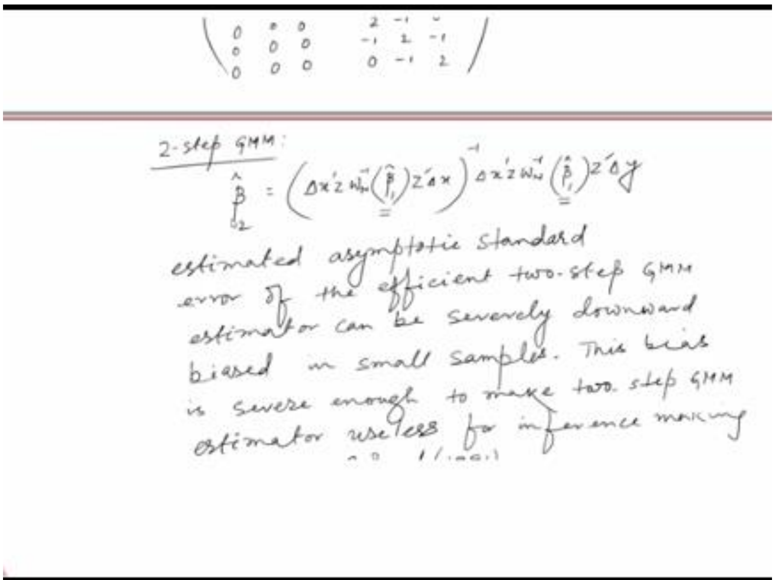
In this setup a one-step GMM estimator is given by the following $\widehat{\beta_1} = (\Delta\chi' \mathrm{Z} W_N^{-1} Z' \Delta\chi)^{-1} \Delta\chi' \mathrm{Z} W_N^{-1} Z' \Delta\gamma$. This is the one-step GMM estimator and what is Δ, W_N^{-1} here? This is an important element in this GMM estimator which we need to pay special attention to.

(Refer Slide Time: 14:38)



This W_N^{-1} is an initial positive definite weight matrix. For example, 2 SLS two stage list square sets $W_n = \left(\frac{1}{N}\right) Z'Z$ is an initial weight matrix that is efficient when v_{it} *i.i.d.*, is $W_n = \left(\frac{1}{N}\right) \Sigma_{i=1}^{N} Z'_i H Z_i$. Now, what is this H over here?

(Refer Slide Time: 17:06)



Here, H $= \begin{bmatrix} 0 & 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 \end{bmatrix}$. This is very interesting if H is 2, -1, 0, -1, 2 ,-1, 0,-1, 2. So that means 2 is there on the diagonal element and then these are all 0 and this is also 0 and this is again 2, -1, 0, -1, 2,-1, 0, -1, 2. Now how to get the two-step GMM? 2-step GMM $\widehat{\beta_2} = \left(\Delta\chi' Z W_N^{-1}(\widehat{\beta_1}) Z'\Delta\chi\right)^{-1} \Delta\chi' Z W_N^{-1}(\widehat{\beta_1}) Z'\Delta\gamma.$

Here if you compare the $\widehat{\beta_2}$ and $\widehat{\beta_1}$, everything else is same. Only thing is that while estimating $\widehat{\beta_2}$, I am incorporating this $\widehat{\beta_1}$estimated earlier. That is why it is called 2-step GMM.

Now why do we need this two-step GMM? To gain on the consistency front. So, two-step GMM is more consistent than the one-step GMM. But in the estimation of two-step GMM since we are using $\widehat{\beta_1}$estimated through the first step as a result of which estimated asymptotic standard errors of $\widehat{\beta_2}$, it is still efficient, but the standard errors are severely downward bias that is the problem.

So, to gain consistency what we are losing is basically efficiency. That means the standard errors are actually downward bias. We are getting a bias in the standard errors of this $\widehat{\beta_2}$. So, what is happening? The estimated asymptotic standard errors of the efficient two-step GMM estimator can be severely downward bias okay, severely downward biased in small samples.

And this bias is severe enough to make two step GMM estimator useless for inference making, this is told by Arellano and Bond 1991? So, what is happening here compared to one step GMM estimator, two-step GMM estimators are more consistent, but to achieve consistency, what we are losing here is we are getting a downward bias in the asymptotic standard error of this two-step GMM estimators.

So asymptotic standard errors of the efficient two-step GMM estimators are severely downward biased and the bias is enough to make it useless for inference making, as pointed out by Arellano and Bond. If that is the case, then what is required is to correct this downward bias or the two-step GMM estimator. So that is made by Windmeijer 2005 and that we can implement in STATA by a simple Robust command.

(Refer Slide Time: 24:37)

biased in small samples. This ...
is severe enough to make two. step GMM
estimator useless for inference making
Arellano & Bond (1991)

— Correction of this downward bias of
the standard errors of 2-step
GMM estimator.

$\Downarrow$

Windmeijer (2005) $\Rightarrow$ Robust command

What is required is a correction of this downward bias of the two step GMM, downward bias of the sorry of the standard errors of two step GMM estimators. And this is done by Windmeijer 2005 by the Robust command.

Therefore, there are two types of GMM estimator, both in the context of difference as well as in the system. We can get two types, either one step or two step. And actually, the difference between one step and two step is basically on a particular weight matrix. If you look at the difference is in the way we are estimating this in one step, everything is same.

In two step, we are using the $\widehat{\beta_1}$estimated in the first step and then we are using it here. So just because $\widehat{\beta_1}$is used here, so this is making two step GMM estimator consistent, more consistent than the first step but the cost is actually in terms of the downward bias that we are getting in the asymptotic standard error of $\widehat{\beta_2}$, which should be corrected by the Windmeijer correction.

So that is why two step GMM estimators are preferable compared to one step because of this consistency gain, but this is the cost. That is why some researchers, they prefer to report both one step as well as two-step GMM estimator to show how much bias we are getting in the standard error of $\widehat{\beta_1}$. And then we try to implement that Windmeijer correction to correct for the downward bias.