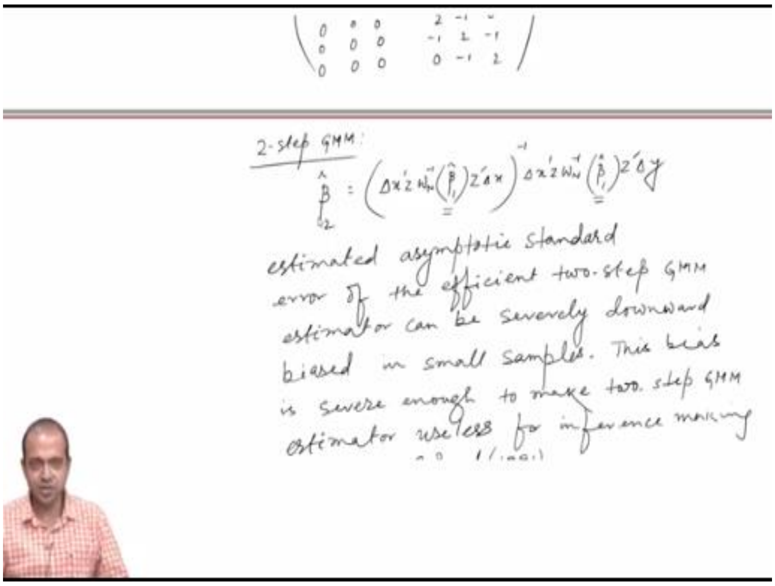
Applied Econometrics
Prof. Sabuj Kumar Mandal
Department of Humanities and Social Sciences
Indian Institute of Technology-Madras

Lecture - 55
Dynamic Panel Data Model - Part XVIII

(Refer Slide Time: 00:15)



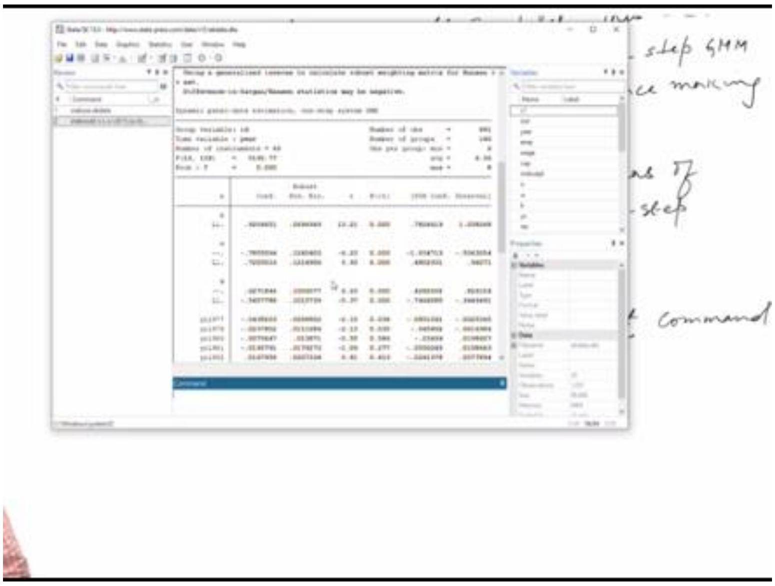
Now we will use the same data set to understand the severity of this problem. That means, how two-step estimation can make the standard error downward bias.

(Refer Slide Time: 00:41)

So let us go back to our statistical software and then we will be using both difference and system GMM to understand this downward bias. We will be again using the same data, AB data. The model that we are going to estimate is *xtabond2 n L.n L(0/1) . (w k) yr*, gmm(L.n) iv(yr*, equation(level)) robust small .*

Here in this model we have not specified any two steps so that is why we will be getting one step difference GMM because we have also taken equation level.

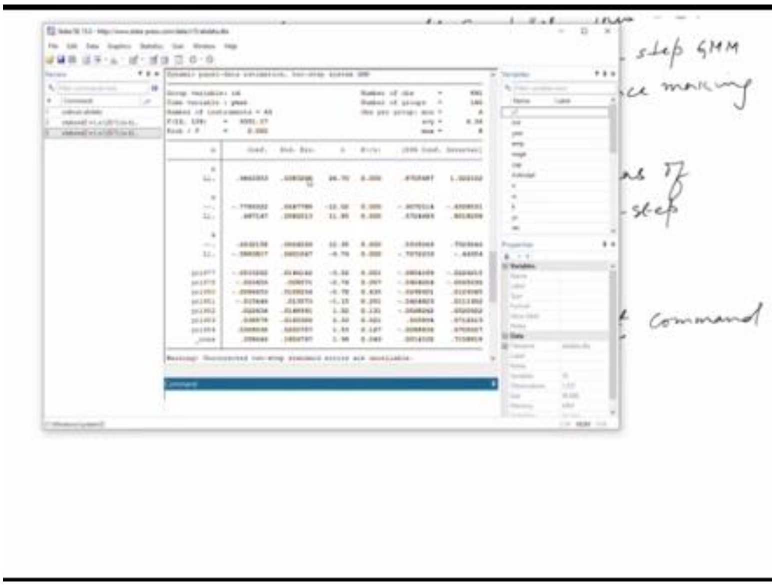
(Refer Slide Time: 03:07)



The moment we input equation level that means, both the difference as well as level equation are considered. This is one step and the coefficient is 0.92, this means it is lying well within the interval of 0.74 and 1.04.

Therefore, the criteria is fulfilled. Now if I go for two-step without any Windmeijer correction, I would not need to put any *robust* command just put *twostep*. This can be rewritten as *xtabond2 n L.n L(0/1) . (w k) yr*, gmm(L.n) iv(yr*, equation(level)) twostep small .*

(Refer Slide Time: 04:06)

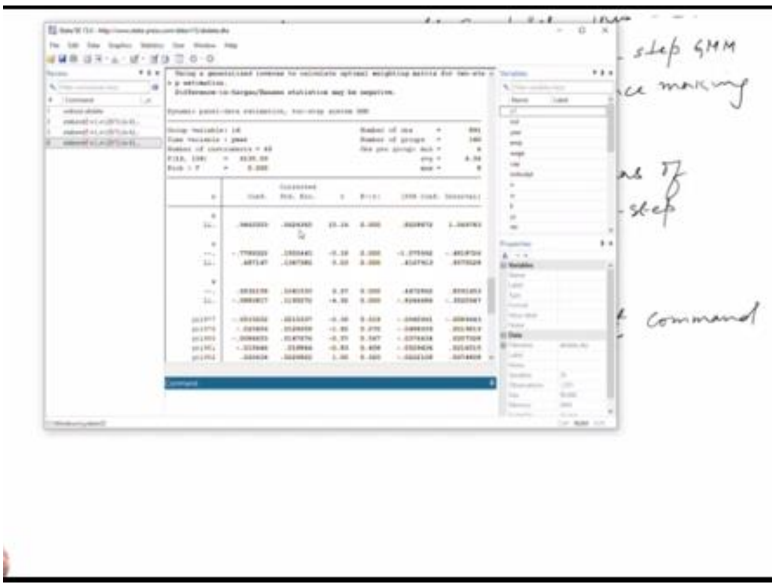


Therefore the coefficient is 0.94 which was earlier 0.92. Now compare it with the standard error. This is 0.06. But look at this is 0.03. Therefore, when I put two steps without any correction then I am getting 03.

Thus, this two step estimator 0.94 is more consistent compared to the one step. Both of them though lying within the interval, but the two step standard error is downward bias. You can see this is 0.038 while the earlier this was 06. So compared to this, this is less. So what we will do now, we will put *twostep* and also *robust* and then you see what is happening *xtabond2 n L.n L(0/1) . (w k) yr*, gmm(L.n) iv(yr*, equation(level)) twostep robust small .*

.

(Refer Slide Time: 05:52)

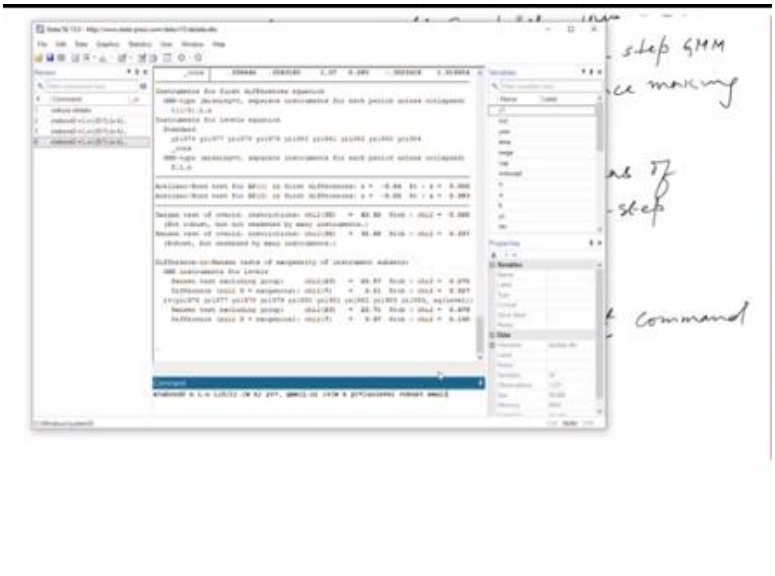


Now the standard error becomes 0.0624, which was earlier 0.03. So again, it has come back and if you compare the one-step estimation standard error is 0.0696, it is almost like 0.0624. So, we have gained consistency but at the same time, we have corrected the standard error, asymptotic standard error of the two-step estimator. This was suggested by Windmeijer in a 2005 paper Windmeijer correction.

So, within the system GMM estimation, we showed that if two-step GMM if applied, standard errors are downward biased. To correct for that, we need to put the robust command even though the coefficients are actually more or less the same, compared to the one step. Therefore, whenever we are applying two-step GMM, we must put that robust command to correct for that bias.

Now we will apply the same model in the context of difference GMM and then we will see what is happening.

(Refer Slide Time: 07:32)

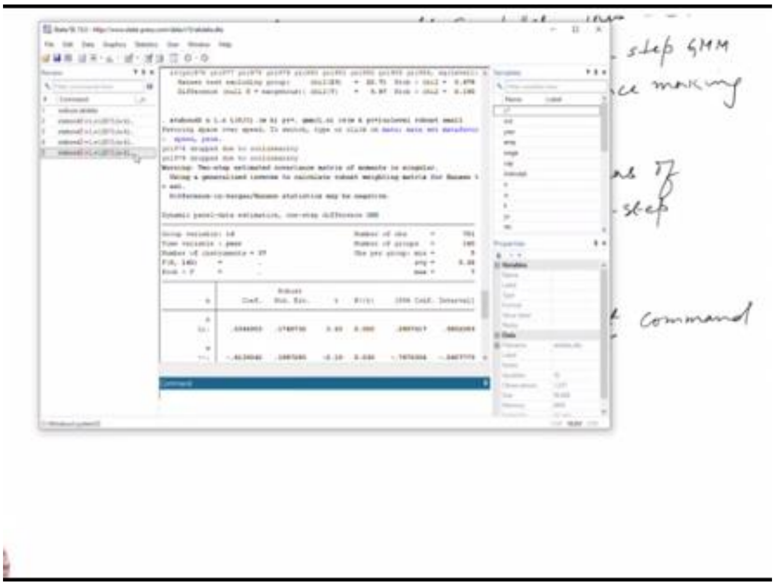


So let us put *xtabond2 n L.n L(0/1) . (w k) yr*, gmm(L.n) iv(w k yr*) nolevel robust small* . Here we are estimating a one-step difference GMM.

(Refer Slide Time: 09:50)

So, when estimating a one-step difference GMM the coefficient is 0.63. That means the coefficient does not lie within the specified interval. So, a one-step difference GMM could not ensure that the estimate lies within the theoretically defined bound. There might be many reasons for that. Therefore, we should not blame only on one step or two-step.

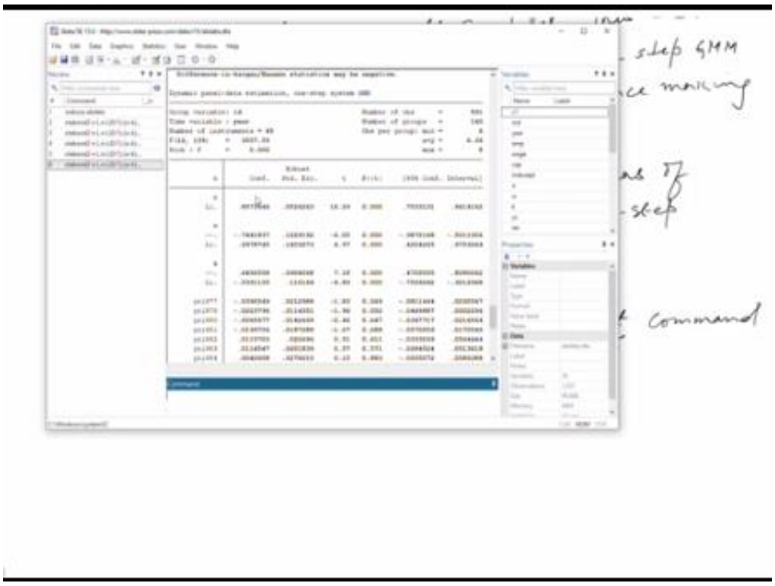
(Refer Slide Time: 10:27)



Firstly, what we are doing here in the specification is that I have considered only one period lag of that employment variable as endogenous. I have not considered second or third-order lags. Here, for the w and k I have considered one period lag and then I have not considered industrial output proxy for demand. Here in the iv accordingly, I have not considered the w L 1, W L 2, k L 1, k L 2 and all those as instruments.

So, this could be due to the specification of those variables. For the time being, I assume that this entire reduction in the coefficient is due to the one-step difference GMM.

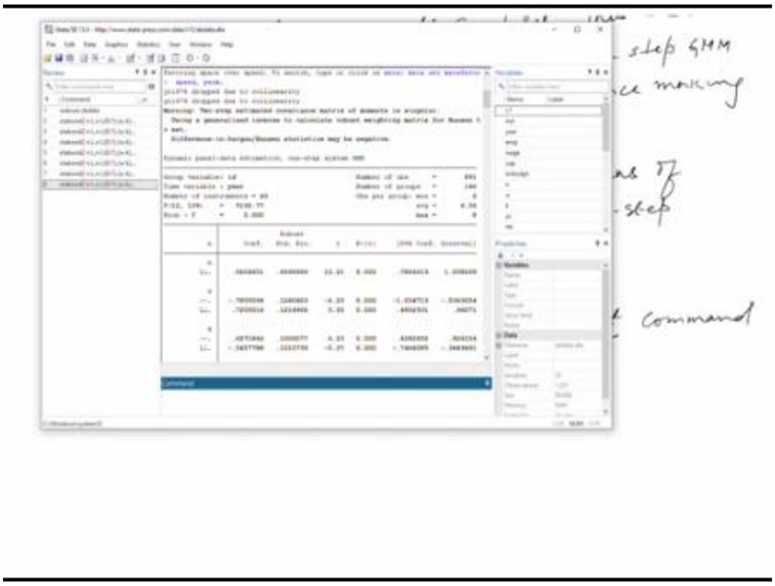
(Refer Slide Time: 11:30)



Suppose I take the same specification with the level equation, Will it give me an estimate which will lie within the interval? Let us see *xtabond2 n L.n L(0/1) . (w k) yr*, gmm(L.n) iv(w k yr*) robust small*. The one-step system ensures the coefficient is lying there. Therefore, while one-step difference GMM could not ensure the estimate to lie within the interval, one-step system GMM ensure.

Now, we will take the same difference GMM only, but we will try to put two step. Same specification, but for twostep difference GMM *xtabond2 n L.n L(0/1) . (w k) yr*, gmm(L.n) iv(w k yr*) nolevel twostep small* to understand how much difference it makes to the GMM estimates as well as its asymptotic standard error if we go for two step discussed earlier.

(Refer Slide Time: 12:45)

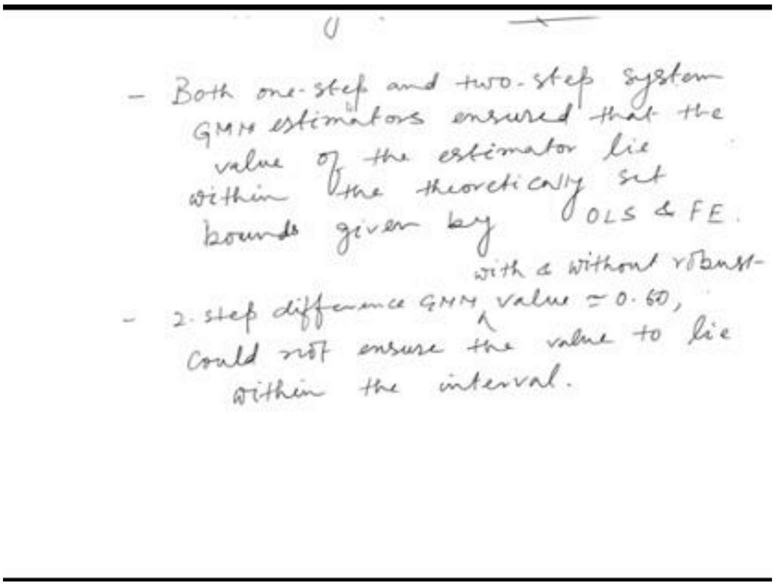


When I put twostep, then that difference GMM is not actually able to ensure my estimate to lie between the theoretically defined bounds. Whatever might be its standard error, I do not bother. First, my estimate must lie within the interval. Since it is not, this means that if you compare the standard error, it is 0.08, earlier when I estimated this model, it was 0.17.

One thing is very clear that two-step estimator standard errors are severely downward bias. It is almost 50 per cent. Look at here, 0.17 and here it is 0.08. That means almost a 50 per cent reduction in the standard error. So, what should we do now? Here in twostep let us put the robust command *xtabond2 n L.n L(0/1) . (w k) yr*, gmm(L.n) iv(w k yr*) nolevel twostep robust small*. Robust triggers the Windmeijer correction in the standard error.

Here even though my estimates are still not lying within the interval, at least my standard error is increased. Therefore, two-step GMM estimators could ensure almost both the estimates of system GMM. That means when we have estimated system GMM the two-step system GMM coefficient is 0.94 lying within the interval. Before that, we have estimated one step GMM, the coefficient lying within the interval. So that means both one-step and two-step system GMM ensure the coefficient lying within the interval.

(Refer Slide Time: 15:10)



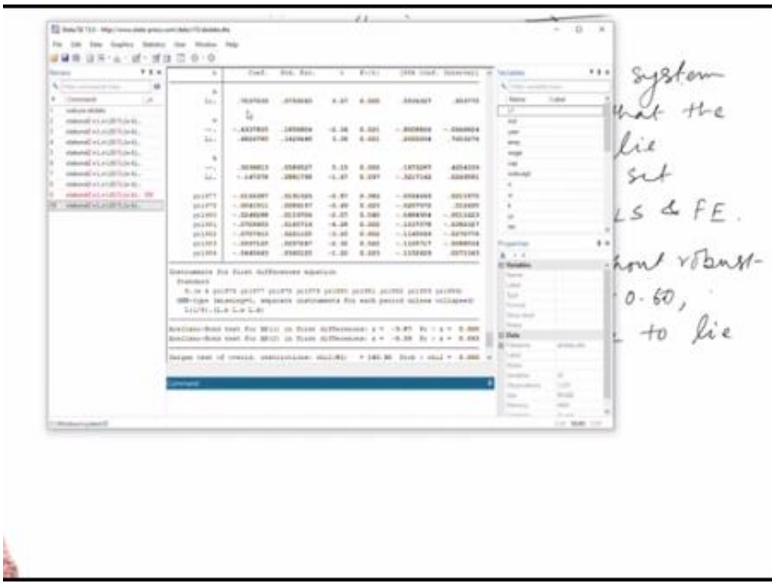
Therefore, we can write that our finding in both one-step and two-step system GMM estimators ensured that the value of the estimator lies within the theoretically set bounds given by OLS and FE. So, both one-step and system GMM ensure the value lies within the interval. We have also estimated two-step GMM with robust command, to ensure that two-step GMM without robust command which shows a downward bias in the standard error.

Then we estimated the same two-step system GMM with two-step correction and we got back the standard error in its previous value. We also did a difference GMM one step and the value is 0.63, which is not within the interval. Then we estimated GMM, again one step but with robust command.

Lastly, the two-step difference GMM with and without robust almost equal to .60, this means could not ensure the value lie within the interval. The value is not within the interval. So, none of these differences GMM model could ensure that my value lies within the interval.

Now what you can do, is actually change the specification. That means here I have considered only L.n as my endogenous variable.

(Refer Slide Time: 21:32)



Now if I change that, the endogenous variable *xtabond2 n L.n L(0/1) . (w k) yr*, gmm(L.(n w k)) iv(w k yr*) nolevel twostep robust small* then the estimate almost lying towards 0.70. So that means, it is approaching towards the limit. Therefore, while estimating the model, we have to check every possibility of whether this coefficient that we are getting is due to one step, or due to two-step. If it is two-step and then it satisfies the criteria lying within the interval, and to correct the standard error just put robust command.

Secondly, even after doing all this if nothing is working out what we need to check, is whether the specifications of GMM style and IV style instruments are correct or not. If it is not, then we can change the variables which were considered exogenous earlier and consider some of those variables as endogenous. However, all these modifications should not be done randomly.

It should be based on some theoretical argument. Here I have a strong ground to consider w and k as endogenous because some theory says that the labor supply and labor demand curve interact and that interaction gives an equilibrium level of wage and labor. Similarly, labor employment and capital also there is simultaneity because of the reverse causality running from employment to capital.

That is why I can use those variables as endogenous. But if you do not have such a theoretical background, you cannot simply add those variables in the GMM style. Then the reviewer of your paper will ask the question why you have considered those variables as endogenous. So exogeneity and endogeneity should come from a theory or your own logic, they should not be randomly selected.

Summarizing today's discussion, what we learned is that there are two variants of GMM estimates, one step and twostep, both in the context of difference and system.

The basic difference between of one step and two step GMM is that we go for two step because two-step is more consistent, and its standard error actually gives the heteroscedasticity and autocorrelation adjusted standard error that hack standard error, because in the dynamic panel data model, we have already assumed there might be heteroscedasticity and autocorrelation within the individuals but not across.

If that is the case, we need to have standard error, which is consistent with this heteroscedasticity and autocorrelation. And if that is the case, two-step estimators are preferable to one-step GMM estimators. But two-step estimators while giving you extra consistency come with some cost. The cost is some kind of bias in their asymptotic standard error.

Therefore, whenever we go for two-step GMM estimation, we must correct that bias in that asymptotic standard error. And the bias in that asymptotic standard error is corrected given the suggestion by Windmeijer 2005, which is the paper I referred to in to this discussion. Appling that, twe will get a two-step GMM and automatically it will correct your downward bias.

Because the biases might be severe as Arellano and Bond said, therefore, this bias could be so severe that it cannot be used for any inference-making. Therefore, we simulated different types of models to understand the sensitivity of our estimate towards one-step and two-step in the context of both difference and system. While system GMM when we applied two-step, the finding shows that both one-step and two-step system GMM ensured that the value lies within the interval. But when we have applied the two steps in the context of difference, then with or without Windmeijer correction, it could not ensure that. So that means two-step GMM estimation is always good. That type of perception we should not carry over, because it is not in the case of difference GMM.

Downward bias is automatically corrected by robust command. But firstly, what we need to check is whether the estimated value is lying within the interval. If that is not lying within the theoretically defined bound, then it is of no use. That is why we could not take this two-step difference GMM.

But at the same time, what we said that we cannot blame everything on the difference GMM method, because the moment I change the specification, some of the variables, which were earlier mentioned as IV style, the moment I brought them into GMM style, the quality of the estimator improved a lot. Therefore, we have to play around with all these specifications until and unless the value lies within the interval.

Once it lies within the interval, then we will correct for its standard error by the Windmeijer correction given by the robust command in STATA.