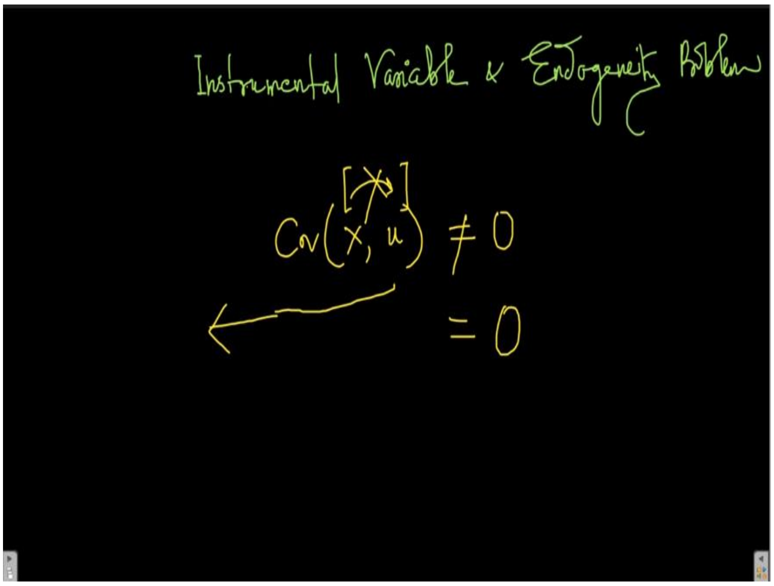
Applied Econometrics
Prof. Tutan Ahmed
Vinod Gupta School of Management
Indian Institute of Technology – Kharagpur

Lecture – 103
Instrumental Variable for Endogeneity Bias Problem

Hello and welcome back to the lecture on applied econometry.

(Refer Slide Time: 00:28)

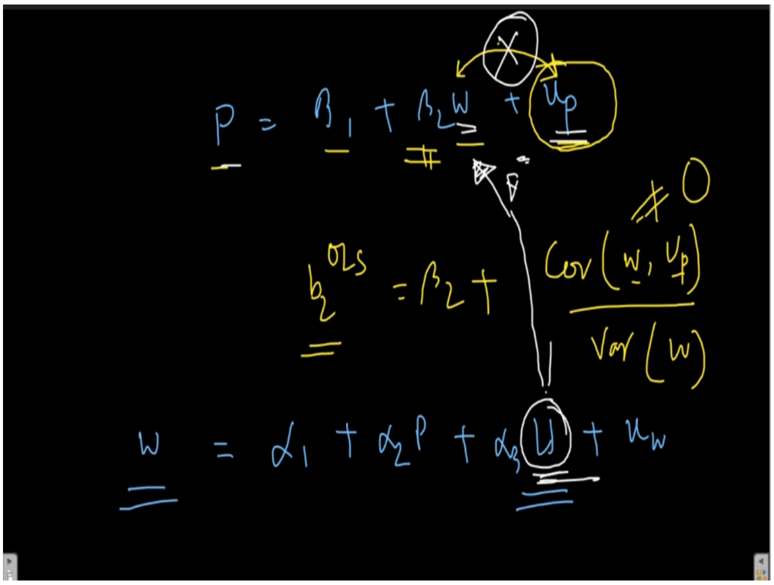


We have been talking about instrumental variable and in the last two lectures we spoke in detail about the problem of endogenity and specifically we spoke about the problem of simultaneous equation model. Now, in all these cases we have seen the heart of endogenity problem lies with the fact that whenever we have a situation where my covariance of X and u is not equal to 0 or in other words when X and u are not independent.

So then we have this problem of endogenity. Now, why exactly, it is a problem? So, to in very simple word we can say X we are basically interested to know in regression equation the influence of X and Y. So, we basically get a regression coefficient for X. Now, when X is influencing u, since u in this model is also determining Y. So, essentially what is happening because X is influencing U? So, we are getting a distorted coefficient beta for X.

So that we do not want. So, essentially I want a variable instead of X which is actually free from u. So that I can have covariance of that specific variable with the error term is 0. So, if I can find a particular variable that I can suitably substitute that which can be suitable substitute for X. So then that can actually play a role to actually remove the problem of endogenity and I claim that instrumental variable or a good instrumental variable can actually do this job. So, let us actually look at our example.

(Refer Slide Time: 02:01)



So, previously, in our previous class we talked about this equation where price rate is basically dependent on the wage rate and the error term. So, we said that the wage rate in the previous equation, we said that the wage rate was actually endogenous. So, in a way wage rate is also getting determined by the price rate. So, what is happening here is wage rate and the error term u p are not independent.

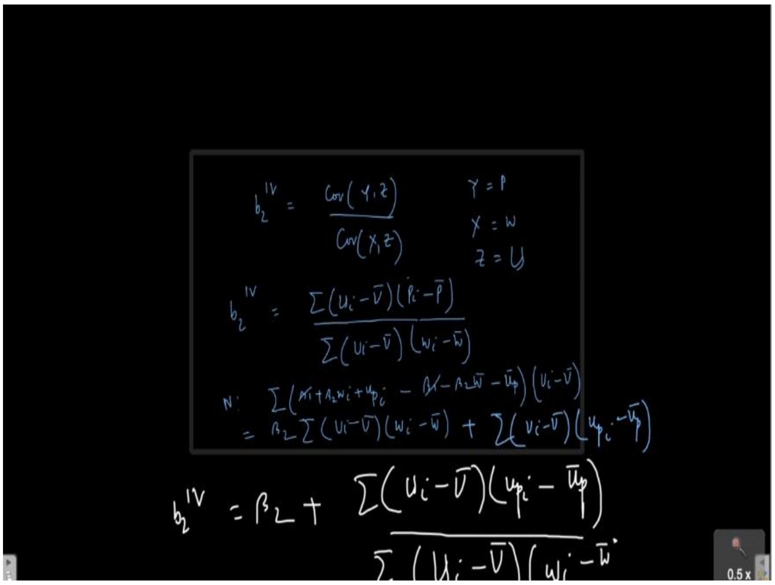
So, when I use, when I actually obtained the reduced form equation for price rate or wedge rate, we saw that wedge rate is having a component in this expression that actually have this u p term. So, when I take a covariance, when I actually estimate the beta ols for the price for basically for this beta 2. So, the b 2 ols is going to have a term where my w so, I will have a covariance of w u p.

So, I will have a beta 2 plus covariance of w u p by variance of w. Now, because w and u p are not independent so, what I am going to have is this term is not going to be 0 because they are not independent. And even asymptotically that is the case. So, essentially I get a b 2 ols which is biased and inconsistent. So, I want to basically get rid of this. And that is basically the source of endogenity.

So, I want to get rid of this and to get rid of this there is a way to actually get rid of this. And if you remember the second equation we had previously is w = alpha 1 + alpha 2 into price rate plus alpha 3 into unemployment rate this is a U + u w. Now, this variable, this particular variable U is coming not from the equation, it is not determined by the equation, it is not an endogenous variable but rather, it is an exogenous variable, it is coming from outside.

And this U is not there in this equation. So, if I can actually replace w with U. So, what I will have? The first thing that I will have is that this correlation between error term and error and the X variable w will go away. And that basically would solve the problem of this endogenous. Now, only point is whether I can suitably substitute w with U. So that is basically the criteria we have to see but let us say we can substitute w with U.

(Refer Slide Time: 05:01)



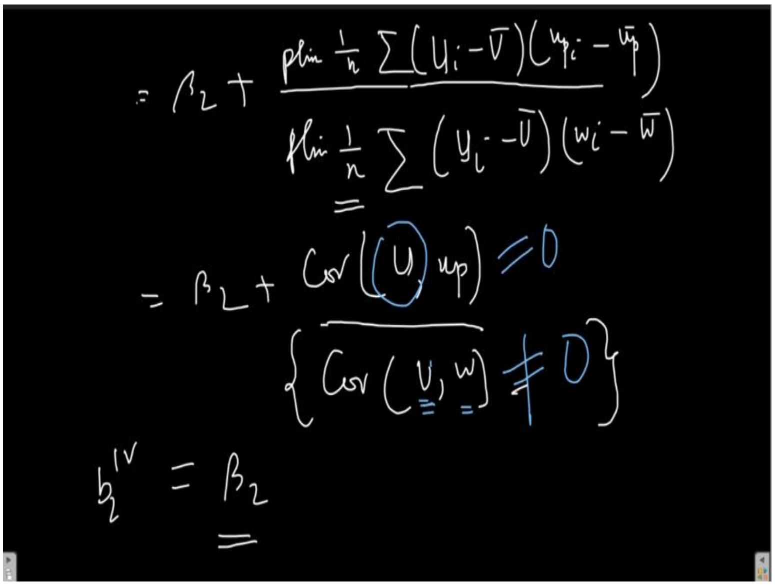
So, if we can do that so then that will be an I V and we have seen previously for I V estimator we can write b 2 I V is going to be beta 2 plus, I have in this case now, my I V let me actually, write down in covariance term. Covariance of X covariance of Y z divided by covariance of X z. In this equation, in our equation Y is the price rate, X is the wage rate and my z is q. Now, if I use this formula and if I actually sort of expand it.

So, what I will get is b 2 I V is equal to summation of U i – U bar into P i minus P bar divided by divided by summation U i, these are all capital U, U i – U r unemployment rate and then I have w i – w bar. Now, we can expand it and if I substitute the value of this P i P bar. So, what I will get is this? Let me write down the numerator only. Numerator is going to be summation of so, basically let us take the P i – P bar term first.

So, beta 1 + beta 2 w i + u p i. So, for **(()) (06:48)** u p i – beta 1 – beta 2 w bar – u p bar. And then I have U i – U bar. So, we are familiar with this expansion and what we will get is I will cancel this beta 1, beta 1. I will get a beta 2 and then I will have a beta 2 here and I will have U i – U bar into w i – w bar and let me reduce the size and then I will have the same term here. But I will have U i – U bar and u p i – u p bar.

So that is what I get. Now, if I divide by the denominator so, this is going to be simply beta 2. So, let me write down the final expression. So, b 2 I V is going to be beta 2 plus summation U i – U bar into small u p i – u p bar divided by the denominator term which is capital U i – U bar and w i – w bar. So, let me use a new page. Let us start flatter it further.

(Refer Slide Time: 08:24)



And if I actually take P lim on both the sides what I will get? For P lim beta 2 ols, beta 2 I V is going to be here beta 2 this is constant term. And here I will have P lim on both numerator as well as the denominator. So, let me take that P lim and let us divide each of these with 1 by n. So, essentially I will have capital U i – U bar into u p i – u p bar. And in the denominator I will have P lim 1 by n summation U i – U bar into w i – w bar.

Now, this term above so, essentially what I will get is beta 2 is equal to covariance of U and u p let us say. And the denominator I have covariance of U and w. Now, U and u p, they are independent. In the previous structural equations we have seen that these terms are independent because U is exogenously given and even when I substitute that in this equation, it does not really make a difference in terms of when you consider the relationship between U and u p.

u p is something that is determined by the price rate equation. So, it has nothing to do with U. So, they are independent. So that I can convincingly argue. So, because we can have this they are independent the U and u p, I can claim that this term is equal to 0. And the denominated term the w and U. Again they are not 0 because U is a important regressor for the second equation here. For wedge rate I have my u phich is my regressor for wedge rate.

So, if I, if there is a regressor so, then it cannot be the covariance cannot be 0. So, they and I am assuming they have a strong correlation and that is why I have included U in the regression term. So, it is not equal to 0 and that is basically with this formulation I will get b 2 I V is equal to beta 2 b. So, it is basically unbiased and it is consistent because for large n, I am getting b 2 I V is equal to beta 2.

Now, in the next lecture we are going to talk about how this relationships like here we see that they are not this is not equal to 0. But on what conditions this U can be a good substitute of X? So that is something we need to explore further. We have assumed that U is actually a good substitute of X. But is it always the case? And that is what we are going to see in the next lecture when we talk about a good I V or a bad I V or a weak I V. With this, we end this lecture here. Thank you.