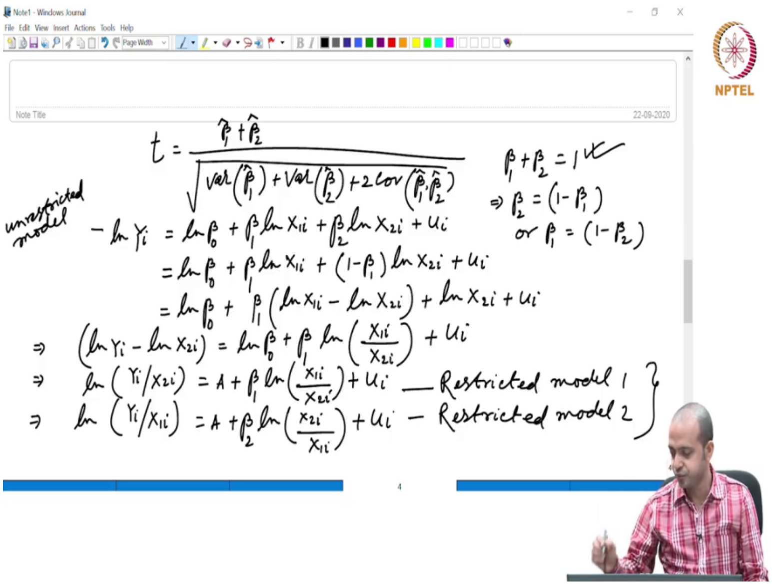
Introduction to Econometrics
Professor Sabuj Kumar Mandal
Department of Humanities and Social Sciences
Indian Institute of Technology, Madras
Lecture 28
Structural break analysis using Chow test Part - 1

(Refer Slide Time: 00:14)



So, t was defined as beta1 hat plus beta2 hat divided by variance of beta1 hat plus variance of beta2 hat plus 2 into covariance of beta1 hat beta2 hat, beta2 hat so, this we have a discuss a study but testing these linear restrictions this particular test you can also conduct that means, the test statistic can be alternatively formulated as using F statistic also that we have not discussed yesterday.

So, let me talk about the application of F statistic to test the same hypothesis of testing linear restriction, so that means are I will write our production function once again Yi basically log of Yi equals to log of beta 0 plus beta 1 log of X1i plus beta 2 log of X2i plus Ui this is what the production function after taking log.

Now, in this particular production function, we will impose the restriction and what is our restriction beta 1 plus beta 2 equals to 1. So, that means from there what I can write as beta 2 equals to 1 minus beta 1 or beta 1 equals to 1 minus beta 2 so you can either substitute the value of beta 1 in terms of 1 minus beta 2 or beta 2 in terms of 1 minus beta 1.

So, let me then, put the value in this way so, I will substitute beta 2 this would become log of beta 0 beta 0 plus beta 1 log of X1i plus in place of beta 2 I am putting 1 minus beta 1 log of

X2i plus ui so, that would become log of beta 0 plus here you have beta 1 log of X1i so, here you have minus of beta 1 log of X1i so, that means, you can take beta 1 common and this would become log of X1i minus log of X2i plus log of X2i plus ui.

Now, this log of X2i you can take left hand side so, that will make log of Yi minus log of X2i equals to log of beta 0 plus beta 1 log of X1i divided by X2i plus ui and this would become log of Yi by X2i equals 2, I can take this as a constant log of beta 0 let us say this is some constant A plus beta 1 log of X1i divided by X2i plus ui.

So, that means, this particular from this particular equation what I can say that output per unit of capital is actually a function of labour per of capital and this particular equation is called the restricted equation. Why this is restricted model?

Because, this model I have derived after imposing the restriction beta1 plus beta2 equals to 1 and our unrestricted model is basically this one this is our unrestricted model so I have not done anything new I have simply imposed the linear restriction beta1 plus beta2 in the model then ultimately, I have come up with this model which is the restricted version of this equation which basically says this is the output capital ratio which is a function of labour by capital.

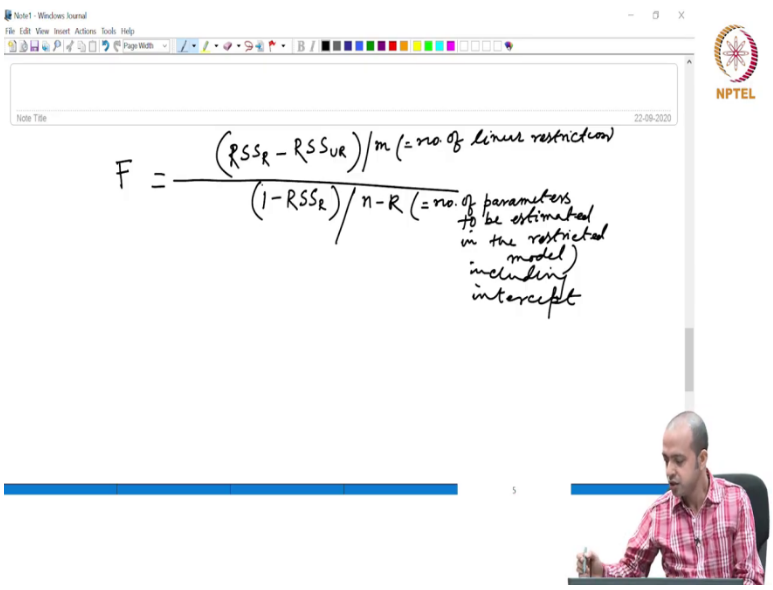
Similarly, if you put the other restriction see here I have imposed, I have substituted the value of beta 2 with beta 1 similarly, you can replace value of beta 1 with 1 minus beta 2 and in that case you will get another version of restricted equation which would become log of Yi divided by X1i would become A plus beta 2 log of X2i divided by X1i plus Ui here, this is another version of restricted model this is also restricted model instead of saying equation I will say this is the restricted model 1 this is the restricted model 2 so, both are restricted versions.

Now, for economist the restricted model 2 is more appealing and more used in the literature, where we assume that output level ratio output labour ratio is a function of labour capital ratio, capital labour ratio, how much capital per unit of labour is endowed with so, 1 unit of labour how much capital that particular labour is getting, that will determine a labour productivity. Because, this left-hand side output by labour which is called labour productivity, average productivity of the labour, average productivity of the labour is determined by how much capital is allocated per unit of labour.

So, equation 2 you can get both versions not a problem from econometric point of view, either you can substitute beta 2 for beta 1 or beta 1 for beta 2 what I am saying that, the restricted versions restricted model 2 is more appealing for the economist and more used in the literature compared to restricted version 1.

Now, what we will do from these 2, we will construct another test statistic to test the validity of this linear restriction so, what we are testing we are actually testing the validity of this linear restriction beta1 plus beta2 equals to 1 and the if you look at the way we are deriving, we are substituting that means, initially we have assumed the restriction is valid if the restriction is valid then, then I can substitute beta1 for beta2 or beta2 for beta1 then I can test I can get 2 restricted versions of the model 1 is restricted model 1 another restricted model 2 then, look at how I am constructing that test statistic.

(Refer Slide Time: 11:11)

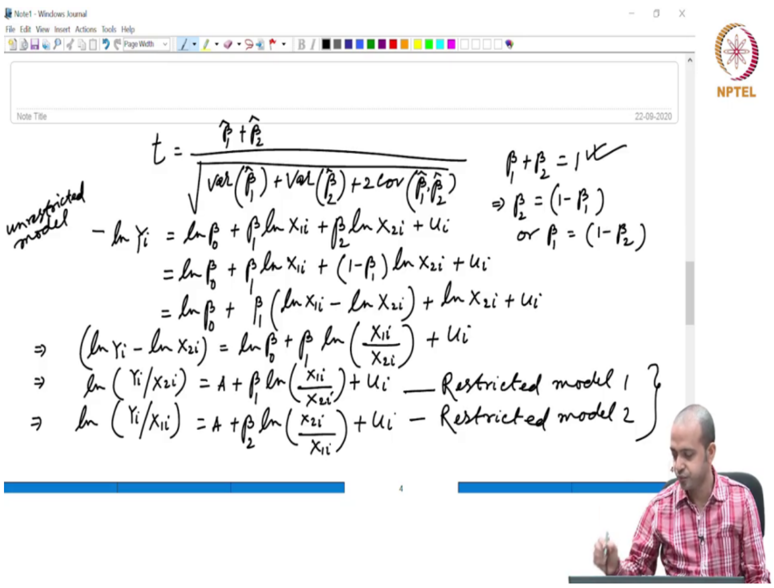


The T statistic is RSS and is defined in terms of RSS is not R square RSS is restricted minus RSS unrestricted divided by m and what is m please keep in mind m is number of linear restriction so I will write here m equals to number of linear restriction which is 1 in this particular context divided by 1 minus RSS R which is equals to n minus k where k is number of parameters to be estimated in the restricted model k equals to number of parameters to be estimated in a restricted model.

Now, first question if you look at the earlier F that means, when we are testing the equality between 2 regression coefficients, we are using F statistic but, that if was defined in terms of

R square, R square new and R square old then, why we are using, why we are not using R square to define the same test statistic here see, we are not defining R square in this module.

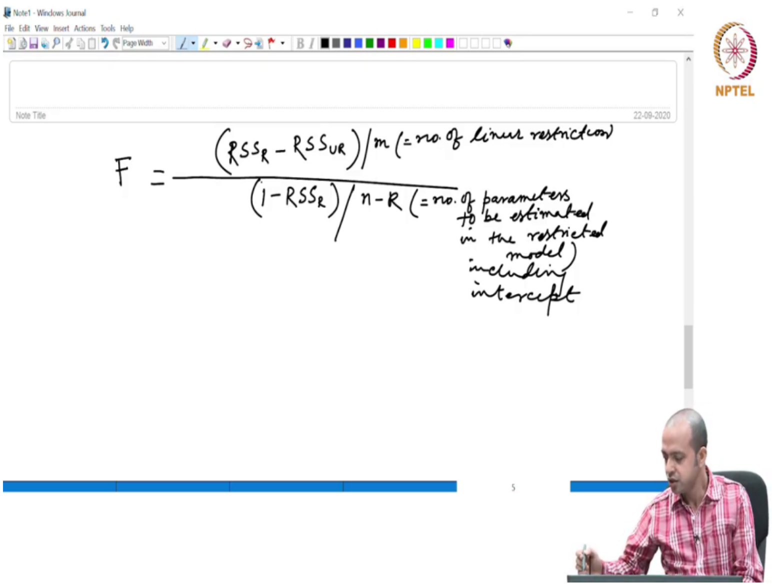
(Refer Slide Time: 13:19)



Because, if you look at what is our 2 models, if you compare the restricted and unrestricted model, see an unrestricted model, what is your dependent variable? Your dependent variable is log of Yi but, in the restricted model, what is your dependent variable?

Log of Yi by X1i so, that means, that dependent variable of the 2 models are not same and when the dependent variable of the 2 models are not same, please keep in mind that R square is not comparable because, what is R square? R square basically the proportion of the total variation in the dependent variable that is explained by your explanatory variables. Since, the dependent variable itself is different in 2 different model then, I cannot say that, these 2 R square are same, that is the reason the R square, what we get from the unrestricted model is different from R square, what we are getting from the restricted model that is the reason these 2 R square are not comparable.

(Refer Slide Time: 14:43)

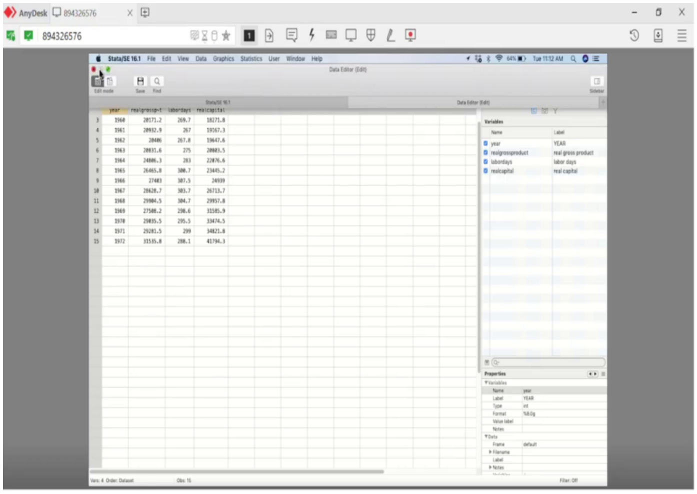


So, alternatively we are defining our test statistic F in terms of RSS, RSS R that means RSS from the restricted model RSS from the unrestricted model 1 minus RSS restricted and k equals to number of parameters to be estimated in the restricted model including intercept that also you should write, including intercept including intercept.

Now, here if you forget whether the numerator should be RSS R minus RSS UR or RSS is UR minus RSS R you just think out of these 2 RSS which is more obviously, when you put restriction in the model you have less number of extramarital variable that is the reason RSS should be higher quite simple, quite understandable in a restricted version, you have less number of explanatory variable that is why, when the number of explanatory variable is less, obviously, the residual part would be more.

So, that is why RSS are is always greater than or at least equals to RSS UR it should not be less than RSS UR that is the reason if you remember that point in mind then, you will never get confused between whether the numerators should be RSS R minus RSS UR or RSS UR minus RSS. So, you can put, you can estimate the 2 models and then, you can collect RSS R RSS UR and then, you can put this value and you can always get the value of F statistic.

(Refer Slide Time: 16:44)



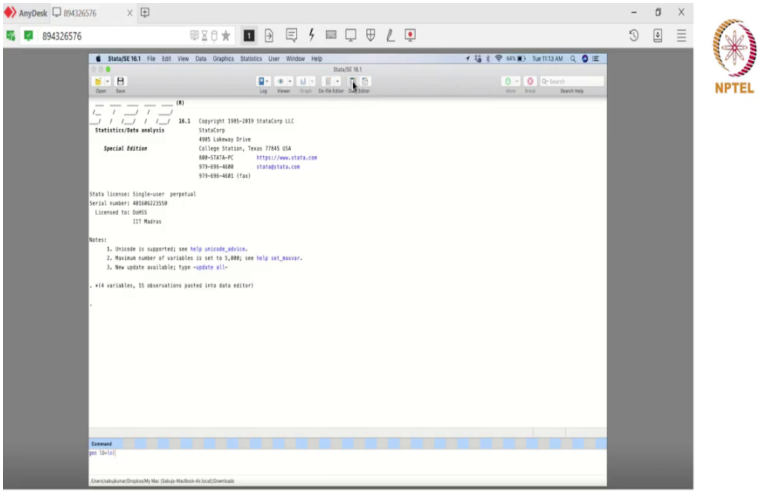
So, now what I will do, I will go back to again, the data set so, I lost for you how to estimate this model so, for that, I am just exiting from this child mortality data opening a new data set, which is given here, this capital labour so, this is our data this is a time series data on real cross product and then I mean in putting this into data.

(())(17:58)

So, I will put that into data editor here, so this is how so, here I will now, what you have to do, you have to estimate the model now, see our variable if you look at our variable is in nominal term that means you have variables are in level you have output in level form, you have capital and labour both your level form and how we have defined a variable? So, we have defined the variable as this we have defined as, see look at real gross output then, labour days gross output then, this is real gross product real this then, labour days and real capital.

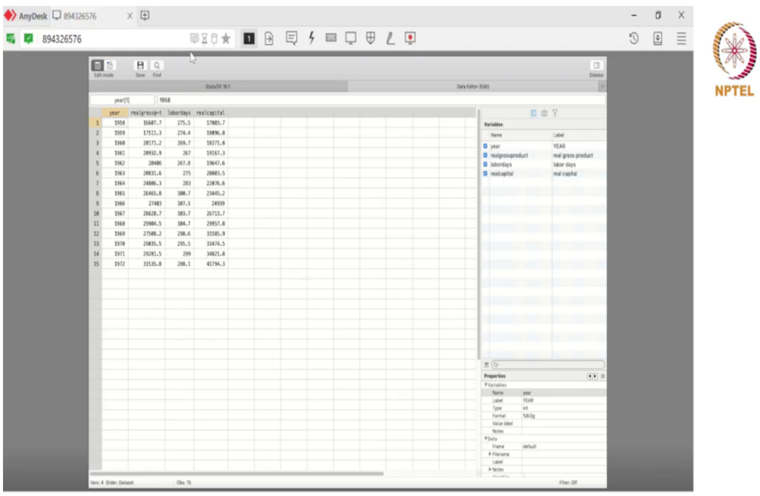
So, first thing, what we have to do, if we want to estimate this model, we have to convert all these variables into log form right and how to do that, how to do that I will just give a, I will just show you, how to convert the variable in level form in log form using data command. So, that is also something to learn so, I will now show you that.

(Refer Slide Time: 23:23)



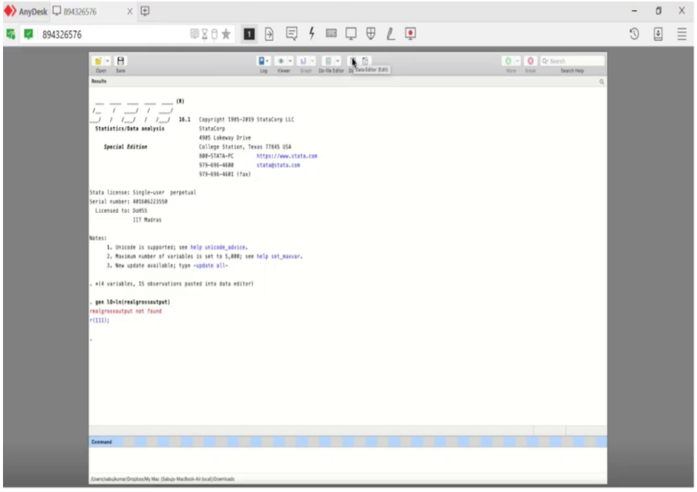
So, here the command would be please keep, in mind command would be gen, I am giving a name for the variable that I am going to generate let us say gen L and then I am putting output. So, this is become, this how we, I am defining the variable log of in the model then, you have to go back and see how you have defined the variable, how you have defined the variable, this you have to see how you have defined the variable.

(Refer Slide Time: 24:15)



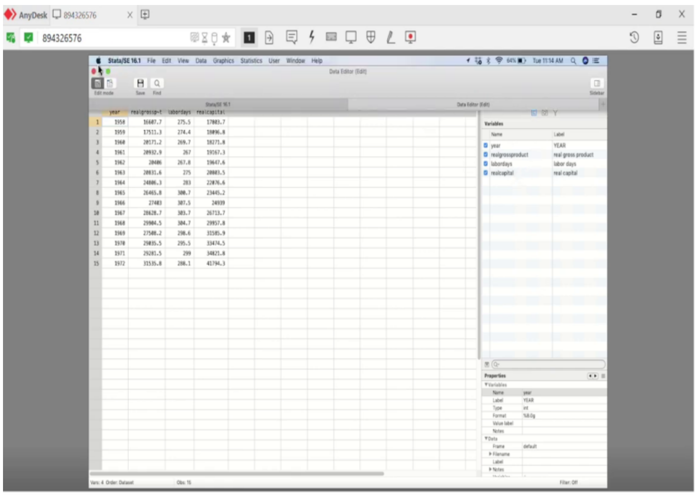
So, there you have to just go to data editor and just you see this is the real gross output, this is how I have defined real gross output so, when you are converting this variable into log form, you have to just put this variable name after that log function then, it will work.

(Refer Slide Time: 24:39)



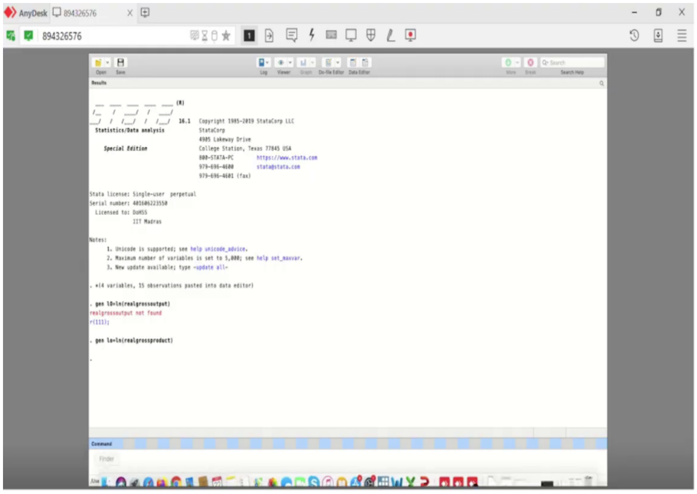
So, I will show you how to convert 1 variable so, log of real then, cross out output then bracket close and if you put enter so that means I have made some mistake there in defining the variable let me see once again how I have defined the variable. So, I have said real cross output how it is defined there.

(Refer Slide Time: 25:14)



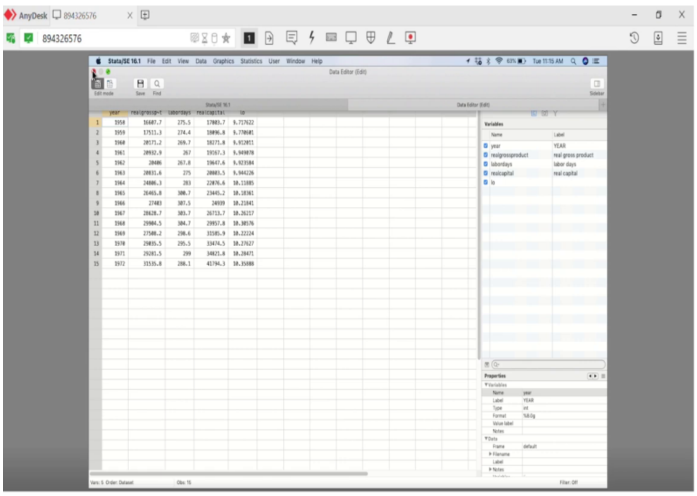
Real gross product, see you have to exactly type this variable name, if you commit mistake there then, it will not come real gross product.

(Refer Slide Time: 25:25)



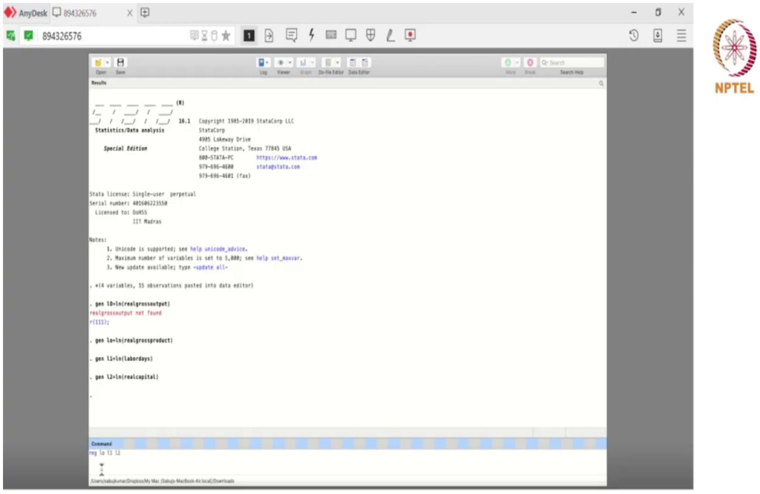
So, instead of output what I will say this is I can this is gen L 0 equals to Ln real gross product then, if you put enter it will work. So, that means now, I have converted the level variable of output into log form and the variable is defined as L0.

(Refer Slide Time: 26:17)



Similarly, I can look at what is the variable name for other one which is labour days and real capital, see here L0 is there so, that means their variable is converted into log so, this is labour days and real capital so, same way you have to convert these two variables also labour days and real capital.

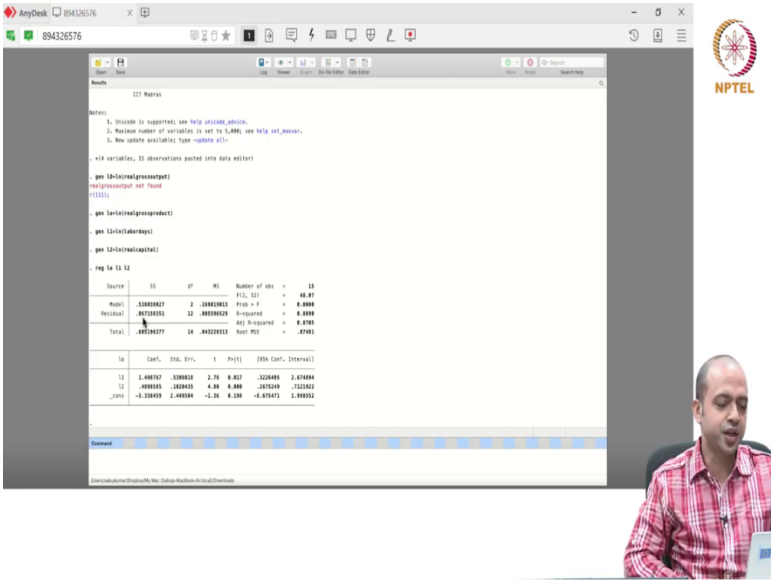
(Refer Slide Time: 26:37)



So, I will put another gen command gen let us say that is L1 equals to Ln labour days and this variable is now converted into log and then, the third on gen L2 defined as Ln in the bracket real capital so, all the theory variables are now converted into log. So, my dependent variable which was real gross product that is also now converted into long form and defined as L0 then.

Labour is measured by labour days which is also now converted into log and defined as L1 and capital variable is also now converted into log of capital and that is defined as L2. So, now all your variables are converted into log and you can easily run the regression which is reg L0 so, L0 then, this is Lo I have given equals to now, equals 2 then, L1 and then L2 this is how I have now converted the linear model into a log linear 1 and you can put enter.

(Refer Slide Time: 28:30)



So, this is your unrestricted model, this is your unrestricted model and from here, what you have to do you have to collect your RSS look at, what is your RSS? Residual sum of square which is 0.0671 so, this you have to collect which is RSS. this is RSS UR because, I have not yet put any restriction in the model that is why it is called RSS unrestricted. what is the value? 0.0671 this would become your RSS UR then, you have to estimate the RSS restricted 1 and how will you do that?

Again, you have to convert your dependent variable into labour output ratio in dependent variable as capital output ratio and how will you do that? simple, again using the gen command. So, if you use gen command, then you have to simply put 1 name let us say gen L output that would be defined as output by labour.

Similarly, you can put a name for labour capital ratio, capital labour ratio and then, that would be defined as capital divided by labour and once you do that, you can estimate the model of unrestricted one and then, restricted one and then you have to collect the RSS from that restricted version. So, we will discuss about the remaining portion of this estimation in tomorrow's class, we will estimate the restricted model in tomorrow's class. Thank you.