Exploring Survey Data on Health Care

Prof. Pratap C. Mohanty
Department of Humanities and Social Sciences
Indian Institute of Technology, Roorkee

Lecture - 38
Difference-in-Difference Method (DID)

Welcome participants to my NPTEL MOOC module on Exploring Healthcare Survey Data. We are on the verge of the last week's contents, in last week's lectures. As I had already pointed out that healthcare without policy evaluation is incomplete and most of the research that is taking place these days is focusing on some forms or the other evaluation of policies or some impact evaluation techniques have been discussed.

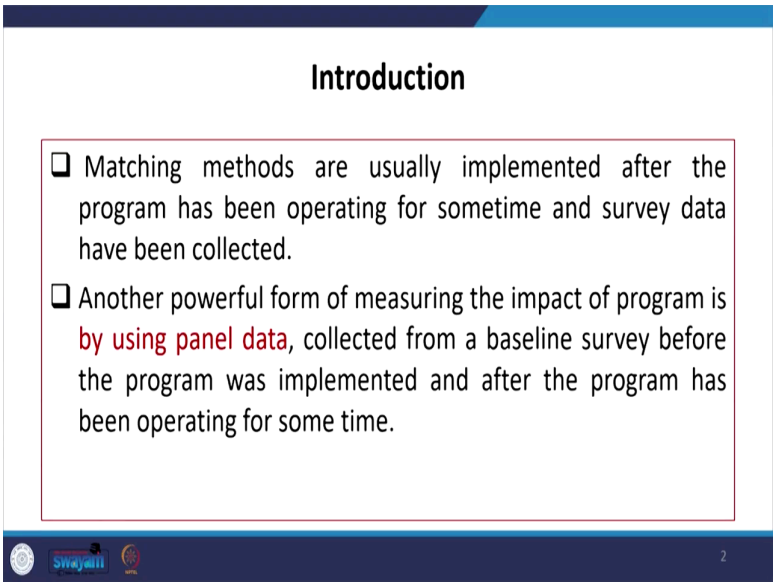
The recent Nobel prize that is being awarded are on this area and it's is quite interesting, that is why we have kept it for our module. In the last two lectures, we have tried to discuss the very basic understanding of experimental design and quasi-experimental design.

However, we will come up with all these practical handouts as well as this one in our next set of the module where we will be specifically focusing on its practical applications. At this moment we are trying to clarify with basic examples

Let us stick to this lecture on DID that is a difference in difference method. As from the word itself you can guess that the evaluation is emphasizing on the very marginal changes in the model.

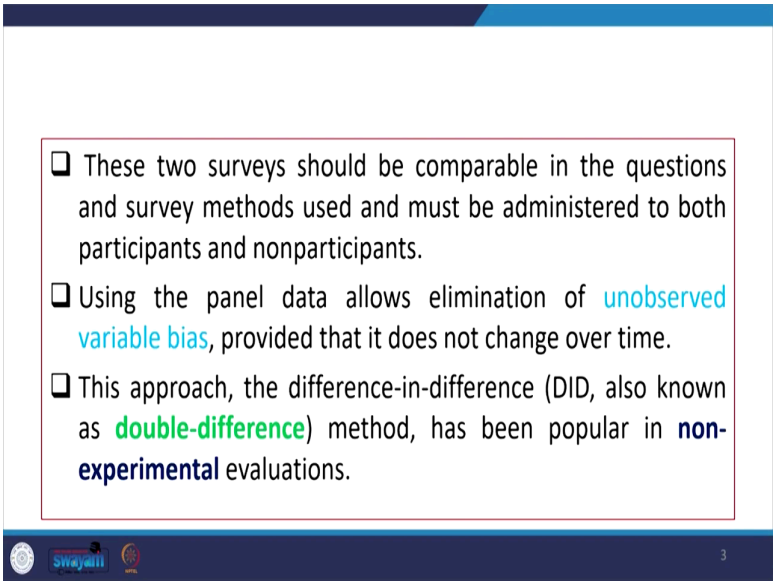
If the difference in the treatment as compared to the difference in the control is observed and if that is significant that means, there is an impact of the policy. This is what we are going to discuss in 15-20 minutes' time. So, let us start explaining it. The first one is here to emphasize on the matching methods that are usually implemented after the program has been operating for some time and survey data have been collected.

(Refer Slide Time: 03:07)



Another powerful form of measuring the impact of the program is by using the panel method those are collected from a baseline survey before the program as well as after the program operating operation of a particular focus.

(Refer Slide Time: 03:19)



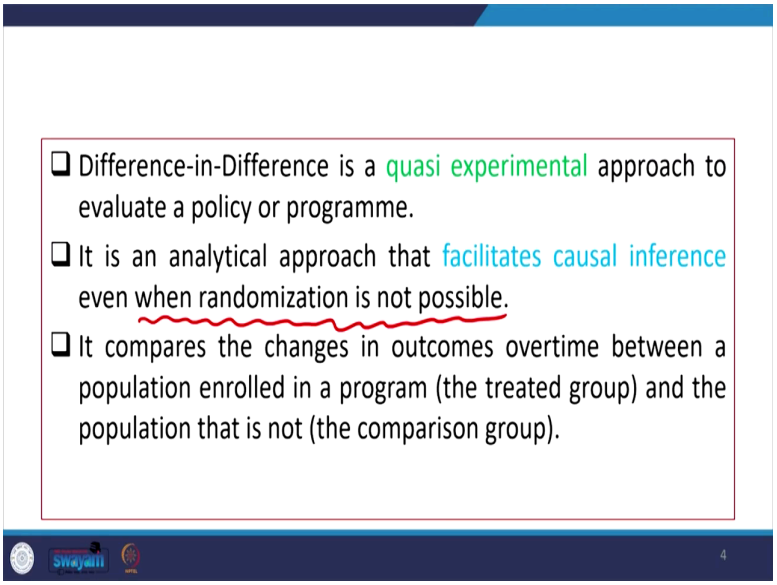
These two surveys would be comparable in the questions and survey methods used must be administered for both participants as well as the non-participants. Using the panel data allows the elimination of all observed variable bias; this provides that it does not change over time.

So, the very focus of panel data is to actually emphasize the variable bias and those that are unobserved.

Unobserved variable usually does not change over time therefore, the panel data can easily capture in the error term and the error term is expected to be minimized as per the technique.

So, unobserved variable bias through the panel estimation is captured and DID is taking use of this. This approach is DID also known as double-difference because we are taking the difference of the difference in both the type of data and this has been popular over the time and this is also called one non-experimental design or evaluation.

(Refer Slide Time: 04:33)



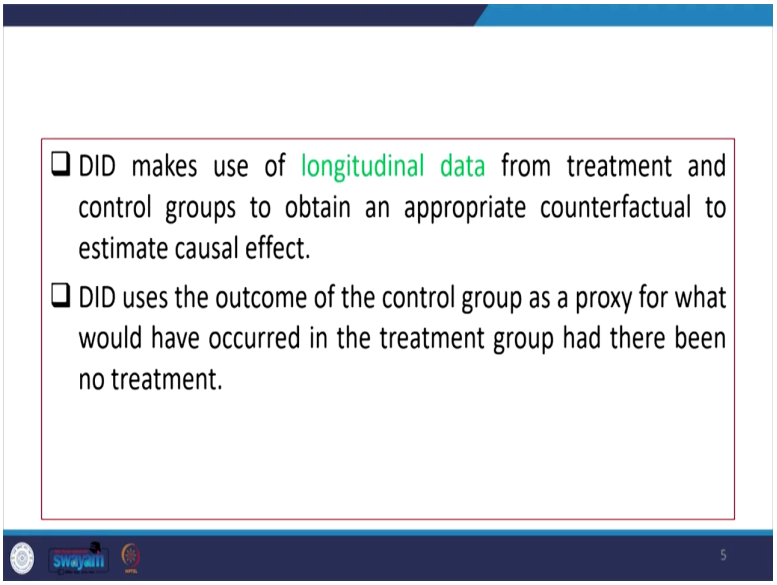
DID is also called quasi-experimental design to evaluate policy or the programs. This is an analytical approach that facilitates causal inferences, even when randomization is not feasible.

So, like in our cities, we said randomization is the most, but this is one of the important aspects because whenever your randomization is not feasible, we thought we are saying it has to be very random by assumption but making randomized with the data is not feasible in reality.

So, in that case, DID is the most appropriate and this gives better causal inferences without randomization of the data. This design compares the changes in outcomes over time between

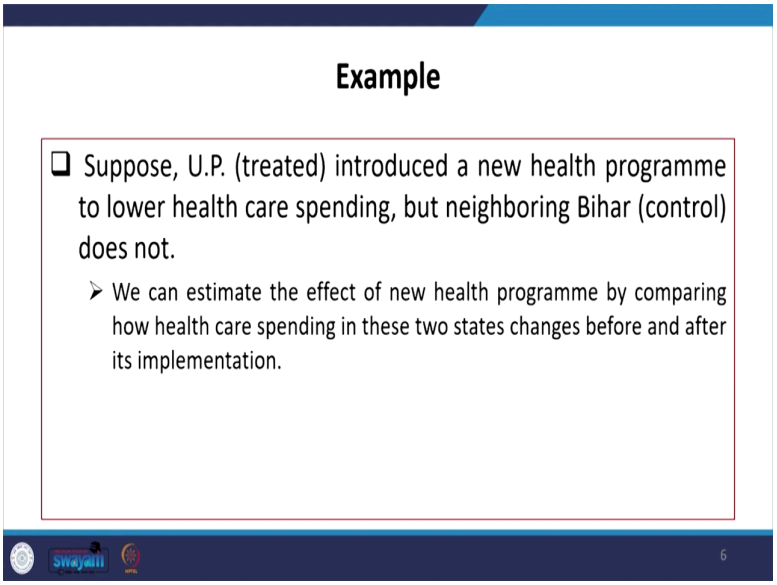
a population enrolled in a program that is in the treatment group and the population that is not enrolled in the program as that is the comparison group or called the control group.

(Refer Slide Time: 05:36)



DID makes the use of longitudinal data from treatment and control groups to obtain an appropriate counterfactual to estimate the causal effect. As I already mentioned this is useful in longitudinal data, the unobservable bias can be also dealt with correctly. DID uses the outcome of the control group as a proxy, for what would have occurred in the treatment group if there had been no treatment.

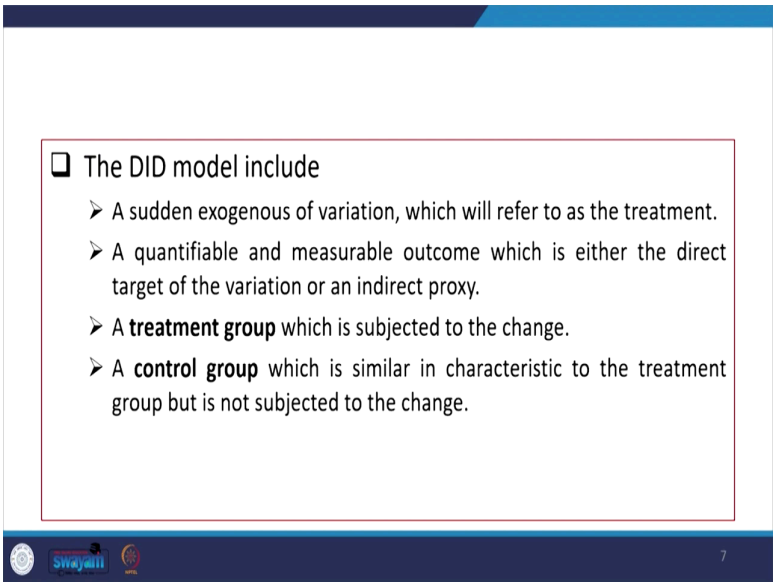
(Refer Slide Time: 06:12)



Here we are citing one example. The example is on two aspects of species specifically on health care spending in two states that are on UP and Bihar. Let us take UP as the treatment group and Bihar as the control group. Where the new health program to lower healthcare spending has been discussed.

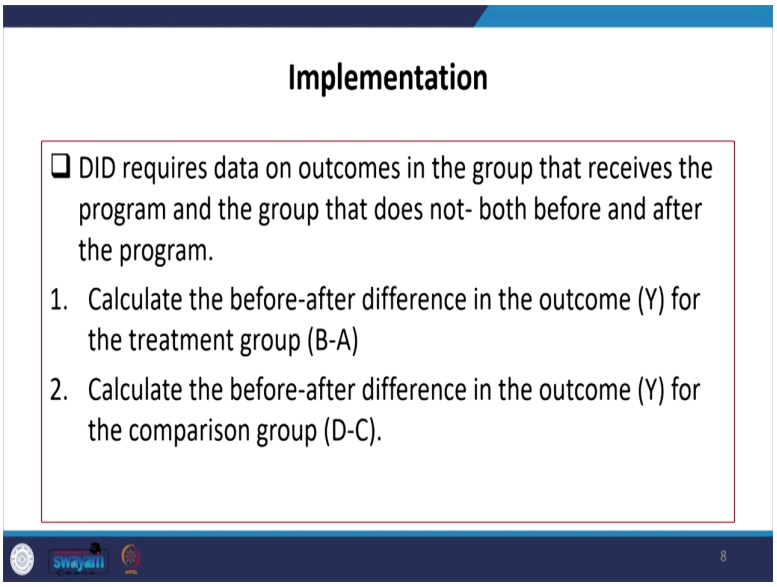
This estimates the effect of the new health program by comparing how healthcare spending in these two states actually impacted the people before and after its implementation.

(Refer Slide Time: 06:54)



The DID model includes a certain exogenous of variable which we will refer to as the treatment. A quantifiable and measurable outcome is either the direct target of the variation or an indirect proxy. A treatment group is subjected to the change, whereas a control group is similar in characteristics to the treatment group, but is not subjected to change which we have already discussed earlier.

(Refer Slide Time: 07:23)

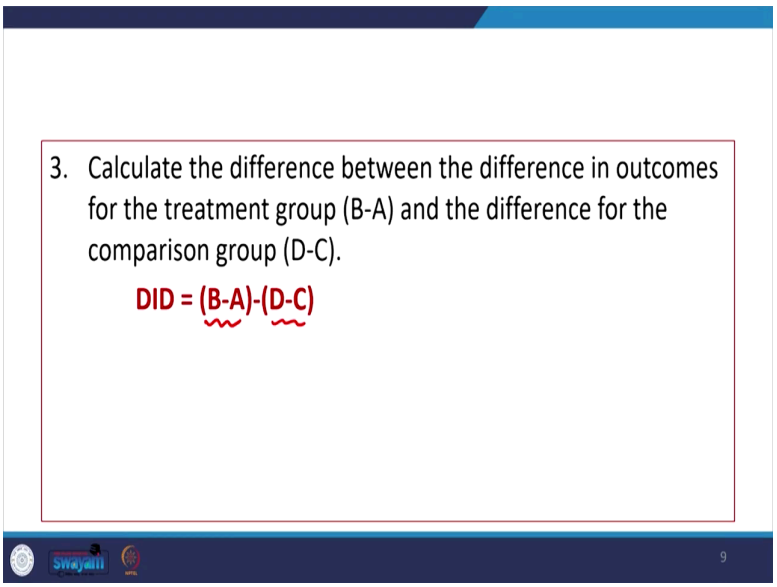


Regarding the implementation, DID requires data on outcomes in the treatment group that receives the true program and the group that does not have both before as well as after the program. So, in that case, we are supposed to calculate the before effect, before as well as after-effect difference in the outcome for the treatment group.

In that case, we are saying before after; that means, B-A in the outcome and also in the control group that is we have considered as before after for the DC, DC here we have taken the code for the control group for the comparison group. So, calculate the before and after difference in the outcome for the comparison group as D to C.

So, in that case, the 1st one is to calculate before and after difference in the outcome for the treatment group and the 2nd one is to calculate the before-after difference in the outcome for the comparison group as mentioned here.

(Refer Slide Time: 08:30)



So, basically, in the DID what it takes, it takes the difference. Difference between this B-A before and after and this is D-C before and after in another group. So, in that case, if there are any changes observed and if those changes are significant that means, our program has certain implications and a significant impact can be derived.

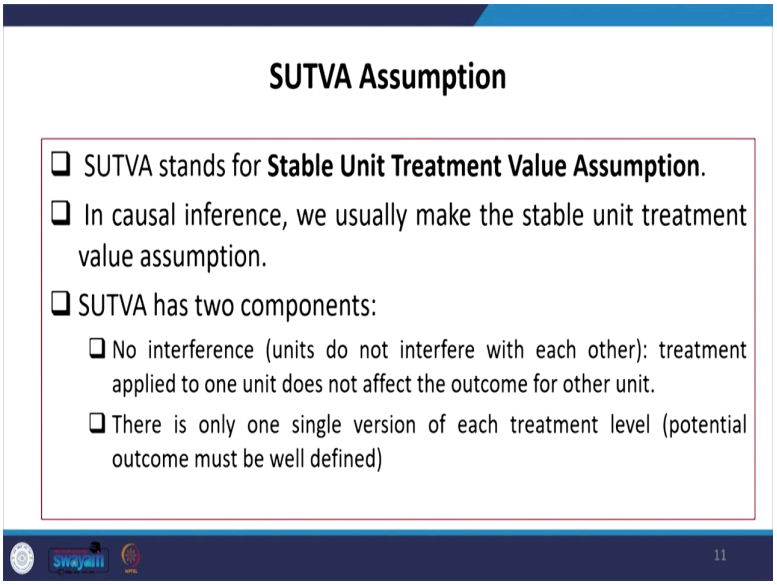
(Refer Slide Time: 08:54)

This is a diagram chart that gives pre and post-treatment of any program. So, pre and post where the treatment is given. So, there is a change that can be noted. Now, you can see the constant difference in outcome if it is there; that means, changes are throughout the same.

Whereas after especially in the post-treatment you can note from 4 onwards post-treatments in the intervention effect can be also noted.

In both the case there are differences in the outcome you can find out the difference, but again you can also find out from which particular effect point you can have a major difference. So, that is basically, called the intervention effect.

(Refer Slide Time: 09:47)

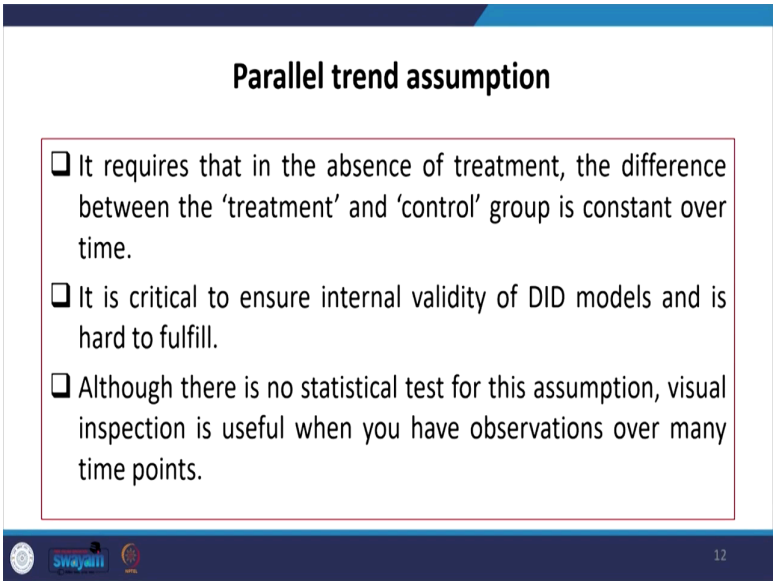


And now we are discussing the stable unit treatment value assumption. There is a certain assumption on which DID is based, in short it is called a SUTVA. SUTVA stands for Stable Unit Treatment Value Assumption. In causal inference, we usually make the stable unit treatment value assumption.

This has again two components: the first one is called no interference and that means units do not interfere with each other, the treatment applied to one unit does not affect the outcome of another unit. So, that means it is part of the SUTVA that is having no interference.

And the second one there is only one single version of each treatment level that is otherwise called potential outcome that must be well defined. The potential outcome should also be defined prior to it, then only we can understand whether it has actually created a difference or not.

(Refer Slide Time: 10:53)

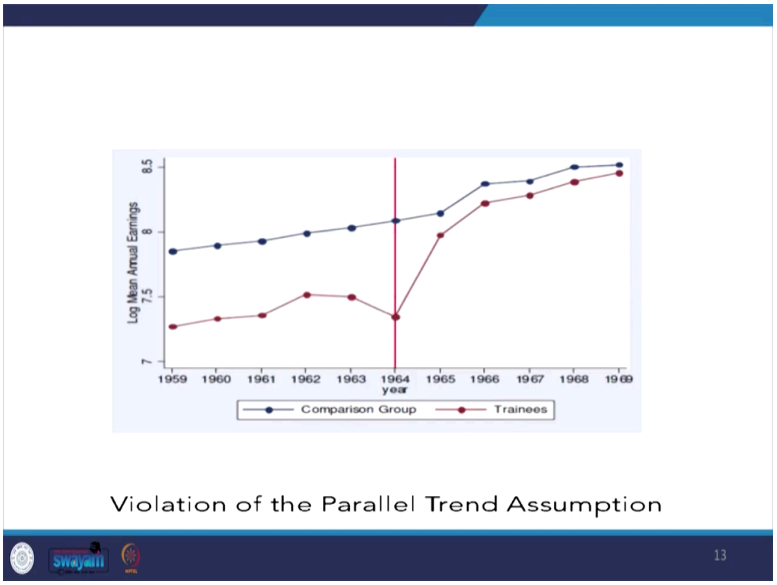


Another assumption is called the parallel trend assumption. This requires in the absence of the treatment, the difference between the treatment and control group is constant over time.

So, that is why it is called parallel. So, a parallel assumption is very required while comparing the intervention effect. So, it is critical to ensure the internal validity of the DID models and is hard to fulfil. Although there is no statistical test for this assumption, visual inspection is useful when you have observations over many time periods.

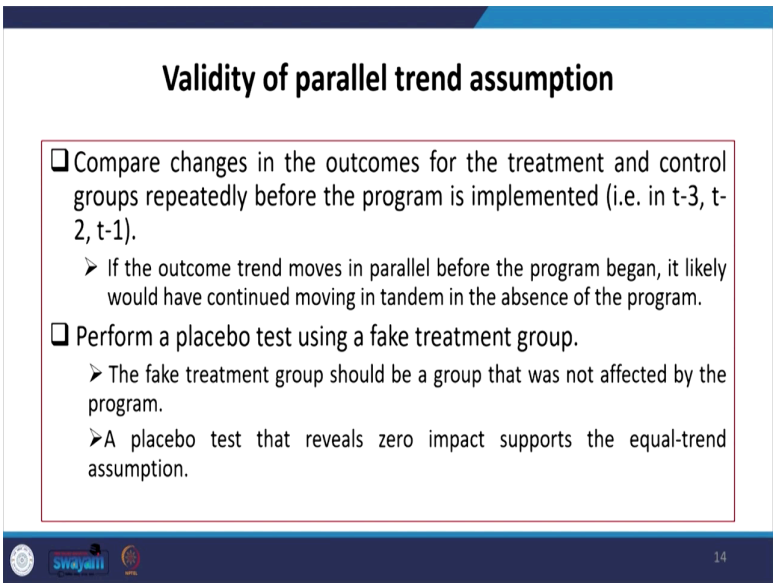
So, if you have so many frequencies point's we can observe and find out, whether the trends are parallel between these two-time periods or two groups and accordingly we can proceed with the DID.

(Refer Slide Time: 11:51)



Like here in the comparison group in trainees, you can see where the log mean annual earnings are given. In the comparison group here you can see that the violation of the parallel trend assumption is made because these two are not in fact parallel. If they are in parallel, then there would have been any sort of intervention that could have been tested.

(Refer Slide Time: 12:16)



Now, validity of parallel trend assumption is presented here this compares changes in the outcomes for the treatment and control groups repeatedly before the program is implemented that is in t 3 or t 1 or t 2 or t 1 time 3 time 2 and time 1 period. If the outcome trend moves in

parallel before the program began, it likely would have continued moving in tandem in absence of the program.

Similarly, another test that is required for understanding the parallel trend assumption is called placebo effect. Placebo test that usually considered to be a kind of fake treatment; a kind of a scenario is given with certain in false information.

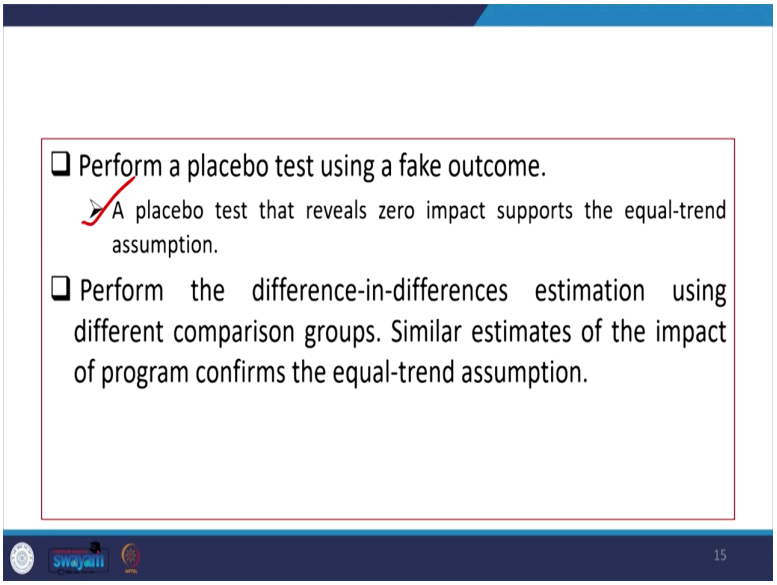
Like a person at the old age used to take if having certain problem with him or her. He has one attention throughout the family that please take me to the doctor.

Now, they give him medicines. So, if you give any sort of fake medicines as well, sometimes those fake medicine also works. So, the parallel trend assumption also goes by this placebo test.

These kinds of test are called a kind of pseudo experiment with those patients. The fake treatment group should be a group that was not affected by the group program, placebo test that reveals zero impact suppose the equal trend assumptions.

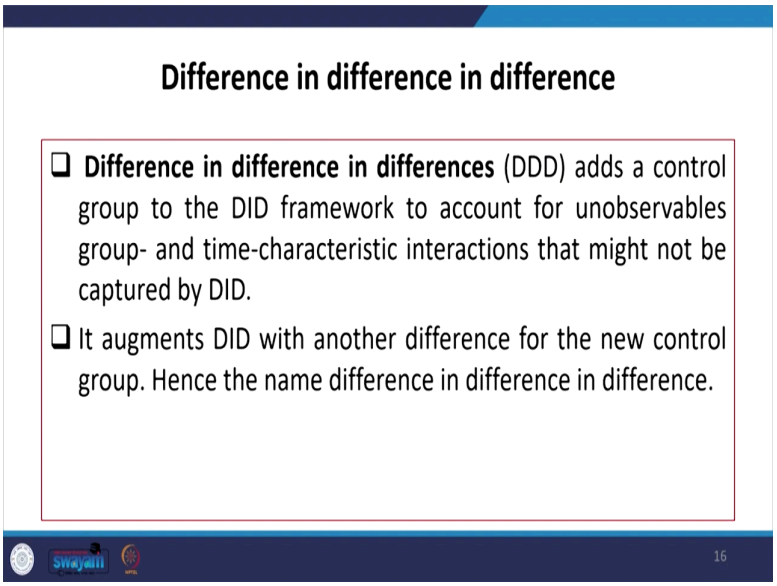
So, like if you just simply test by another intersection the trend should have been diluted, trend should have been disturbed. If it is disturbed that means, it is not working it is not parallel.

(Refer Slide Time: 14:12)



To perform a placebo test using a fake outcome. A placebo test that reveals zero impact supports the equal trend assumption, this is what we just said. Perform the difference in differences estimation using different comparison groups. Similar estimates of the impact of the program confirms the equal trend assumptions. So, equal trend assumption should have been followed.

(Refer Slide Time: 14:35)

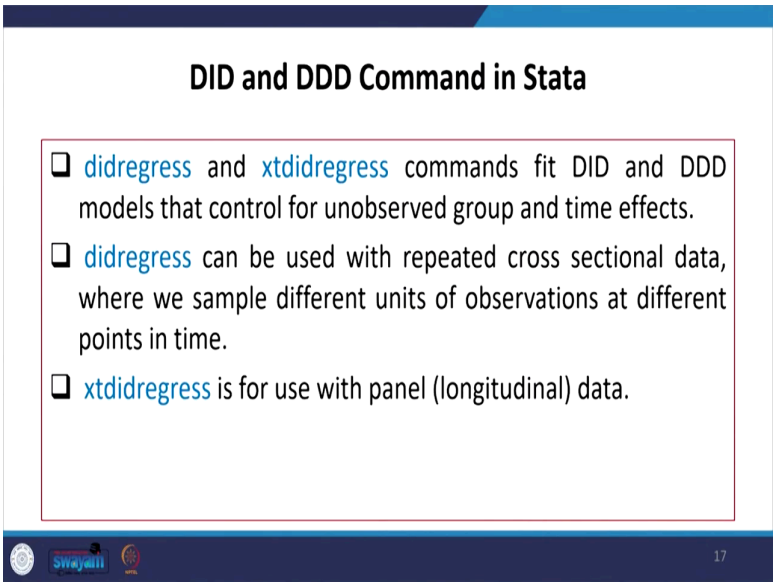


Now, difference in difference in difference. So, triple D is important not just double difference. So, difference in difference and it is difference has to be counted. This adds a control group to the DID framework to account for unobservables group and time characteristic interaction that might not be captured by a DID.

When you have some time component is not just a cross sectional and it changes that might be a third component in the usual in the panel. When time component is there the third difference could also be taken to understand the differences correctly.

So, usually you refer to the panel content or longitudinal data with certain time. Its arguments, it argues DID with another difference for the new control group. Hence the name difference in difference in difference is referred or discussed.

(Refer Slide Time: 15:36)

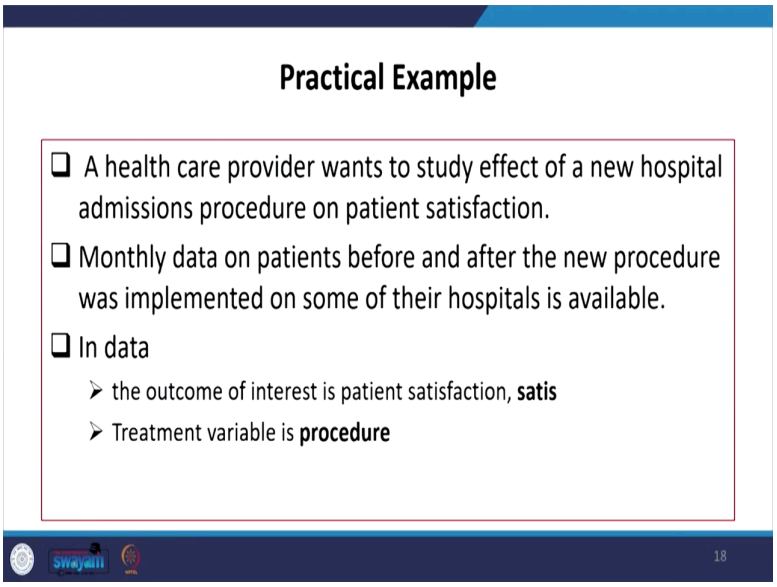


Here are some of the important commands that is essential in running the DID difference indifference or difference in difference indifference. So, triple D like we are not running at this moment but you may follow the latest software. It has already contained all those details in the previous softwares, you would not find that you have to give some proxy or other direction.

At this moment we are actually simply guiding based on the latest software strata that has been published. So, the first command is called didregress. So, DID based regression. So, didregress is given for the repeated cross-sectional data that it is a repeated cross-sectional data where we sample different units of observation at different points in time.

But, in case of panel data that is though repeated cross sectional and panel where longitudinal content is given more time period is given. So, in that case you have to use the xtdidregress. So, xtdidregress is most fitted in case of triple D models and that control for the unobserved group and the time effect as well.

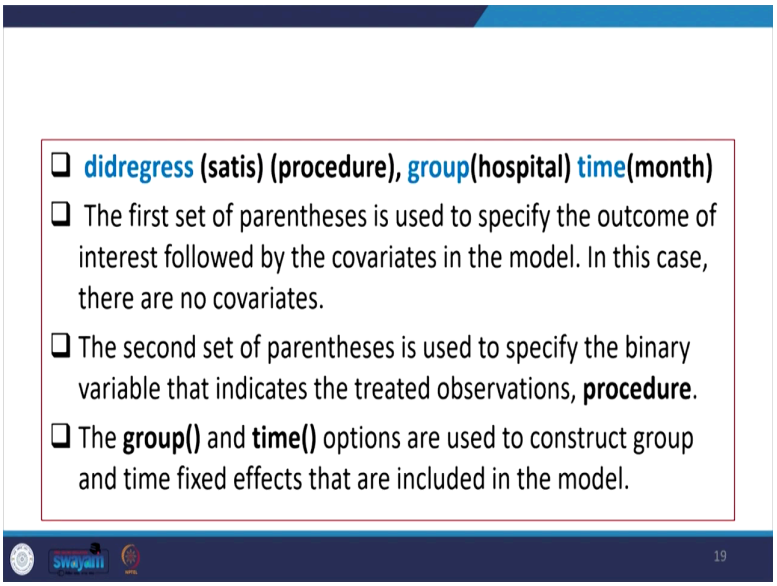
(Refer Slide Time: 17:01)



Practical example we are citing here some results will also show and their interpretation can also be explained. A healthcare provider wants to study effect of a new hospital admissions procedure on patient's satisfaction. So, they have collected monthly data on patients before and after the new procedure that was implemented on some of their hospitals those are available.

In the data, the outcome of interest is that, it is obviously patient satisfaction in short. We are writing as satis as the variable name whereas, the treatment variable is called procedure.

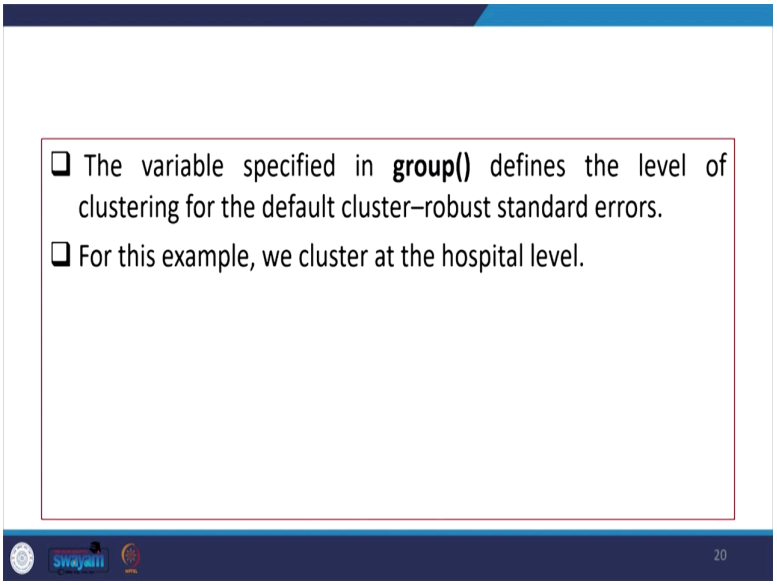
(Refer Slide Time: 17:42)



So, the procedure how procedures have been changed, here as the command didregress. The First one is your satisfaction variable that is the outcome variable, then procedures are the treatment variable; then it is referred to other factors like group as the hospitals, in the group, then time component if any has to be specified.

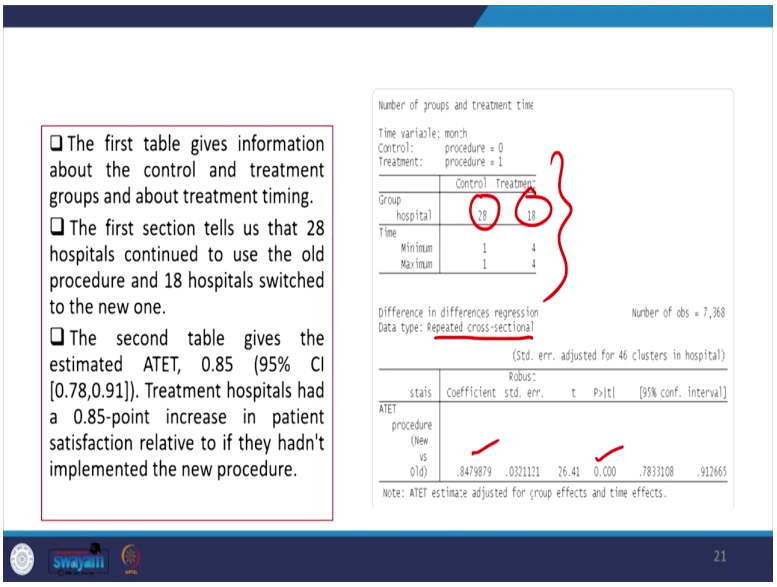
The first set of parentheses is used to specify the outcome of the interest followed by the covariates in the model. So, they are other covariates basically, then, in this case there are no specific covariates. The second set of parentheses is used to specify the binary variable that indicates the treated observations that is called a procedure and the others are like group or time are used to construct group and time effect that are included in also the model.

(Refer Slide Time: 18:43)



The variable specified in the group as defined as the level of clustering for the default cluster robust standard errors, basically, group is given to compare which set of hospitals are actually giving much better trends in terms of satisfaction. So, for this example, we cluster at the hospital level only. So, the hospitals has been considered through the group.

(Refer Slide Time: 19:10)



And here is some of the sample results that was taken from the latest strata software. The information given here is compared with the number of groups and the treatment time.

So, the first table that is here gives better information about the control and treatment groups and a simple description is given. This description provides an idea about how many hospitals in each group are there in control groups and treatment groups.

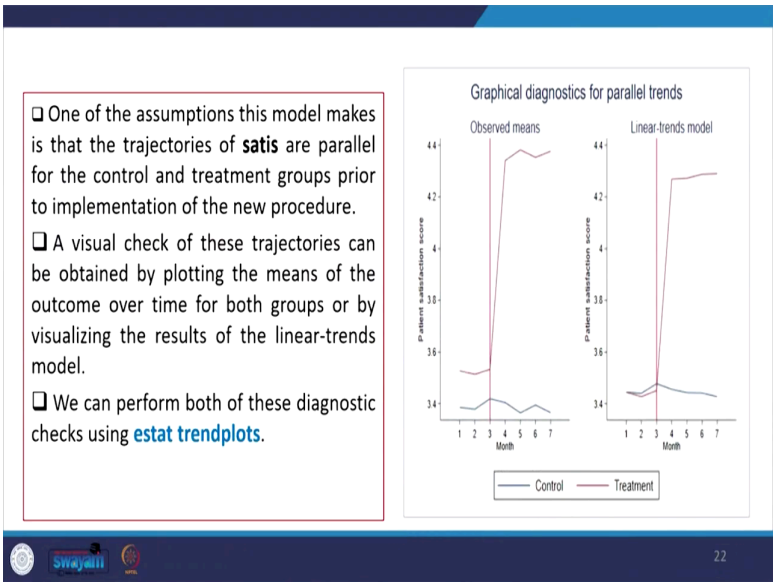
And their maximum and minimum observations frequencies are also mentioned in that particular table. Now, procedure as we already said as a 0 and 1, it is treatment and no treatment, then if it is with no treatment then, of course, they are the control group, is not it?

And now, the first section tells us the 28-hospitals continued to use the old procedure and the 18 hospitals switched to the new one. The second table that is mentioned here gives the standard error and its p values, this application is the result based on the repeated cross-sectional data.

The number of observations are 7368, the average treatment effect on the treatment ATET is calculated and this gives 0.85 as the result with its significance level as 0.00. So, here is the 0.85. This suggests that the treatment had a 0.85-point increase in patient satisfaction relative to if they had not implemented the new procedure.

So, if their new procedures are not adopted, then in that case it could have been lesser, but since those who have adopted this has created a positive change and it is significant.

(Refer Slide Time: 21:29)

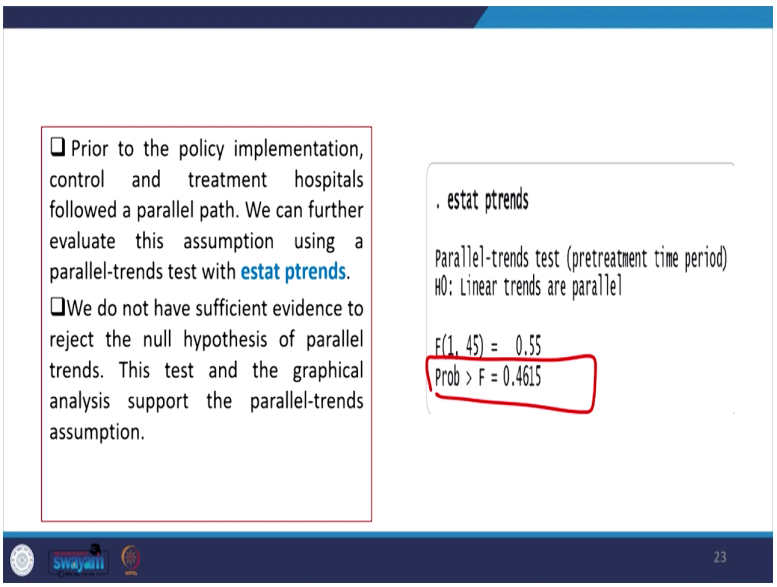


The parallel trend assumption as we already mentioned it is like we discuss satis as the satis for satisfactory as the variable, we try to check some post estimation to draw trend plots whether the assumptions were actually correctly followed or not.

So, the estat trend plots actually give the right direction to it. So, this gives like the blue line is the control variable whereas the red the one is the treatment one, this has actually corrected the changes or not. Whether it has actually changed the trajectory or not, the visual check of this trajectory can be obtained by plotting the means of the outcome over time for both groups or by visualizing the results of the linear trends model.

In that case, it seems that the parallel trend has not been disturbed. So, the parallel trend has been followed in the observed means as well as in the linear trend model it seems almost the same the changes are not significant, so therefore it is following the parallel trend.

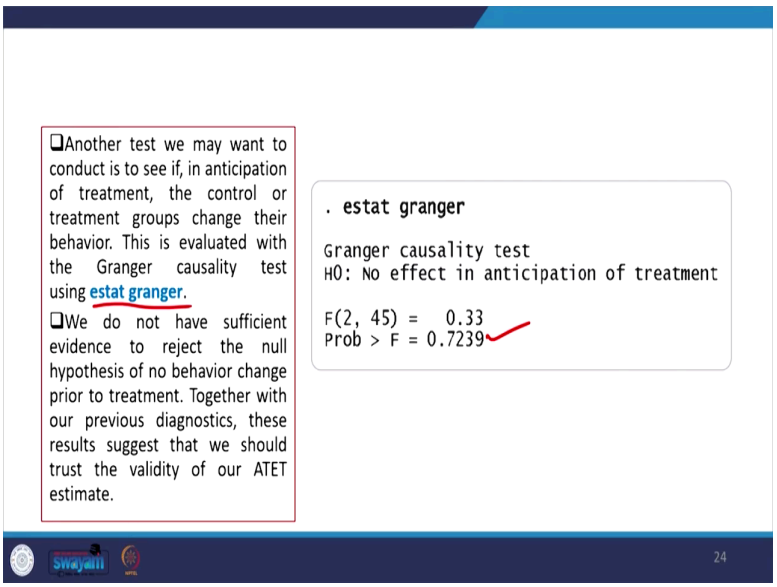
(Refer Slide Time: 22:35)



Now, you can command whether you are going to reject the parallel trend assumption to see this is based or not based on the significance level. The command here is called estat ptrend. Prior to the policy implementation control and treatment, hospitals followed a parallel path.

We can further evaluate this assumption using a parallel trends test with estat ptrends. We do not have sufficient evidence to reject this null hypothesis about the parallel trend which means, this test and the graphical analysis supports the parallel trend assumption. Since this is not significantly deviating, not rejecting our null hypothesis, is alright.

(Refer Slide Time: 23:23)



Now, another test that is also important to give you a certain idea about the causality through the granger causality, no basically, the hypothesis is there; there is no effect in anticipation of the treatment.

with the expectations, the treatment might change on the outcome on the satisfaction level. So, that has also been checked through the granger causality test that is estat granger we have taken that will clarify further that, yes the parallel trend hypothesis has been followed correctly.

So, here the test about all this is not significant. So, we consider that it is not anticipating much difference in the treatment or there is no effect on the treatment. Here in anticipation of treatment the control or the treatment group changes their behaviour, this is evaluated with the granger casualty test.

So, estat granger is going to give you the result. We do not have sufficient evidence to reject the null hypothesis of no behaviour change prior to treatment together with our previous diagnostics this result suggests that we should trust the validity of our ATAT estimate, alright.

So, these are all the details we wanted to explain, further details like its direct state applications, we will come up with a detailed module in our revised module of this.

At this moment, we are trying to clarify the concept and how people are using it differently have also been emphasized. I hope you have gone through it and you will come up with certain ideas and we will discuss them. With this, let me stop here.

Thank you.