Advanced Methods in Impact Assessment Workshop

Day 3: The Effect of the Treatment: Difference-in-Difference and the Average Treatment Effect
Today we will apply the information you learned this morning regarding Difference-in-Difference (DiD) and the multitude of treatment effects, such as Average Treatment Effect (ATE), Local Average Treatment Effect (LATE), and Marginal Treatment Effect (MTE).

There are two objectives for today's exercises:
1. Calculate the difference and then the Difference-in-Difference.
2. Measure a variety of treatment effects

Difference-in-Difference (DiD)
To gain some sense of how the DiD is calculated we will start by looking at simple differences. This will be similar to what we did at the end of Day 1 when we looked at the Within/Without and the Before/After comparisons. For the DiD load the data set `India_Prod_Data_exer3.dta`. As you did before, open a `.log` file and write your DiD and treatment effect code in a `.do` file so you can reference it later.

1. Do a simple t-test to examine whether parcels in the irrigation program had higher log yield than parcels not in the program, using observations from 2011. In the command, we will tell Stata to only use observations from 2011 by using an `if` statement.

   ```
   ttest logyield if sur_yr==2011, by (irr)
   ```

   What does the t-test measure?

In order to measure the difference in outcomes across time between the treatment and control groups we need to manipulate our data set a bit more. Below is code for how to set up the measure.

```
gen lny0 = logyield if sur_yr == 2010
**with the egen command, we are adding the 2010 yield observation
to the 2011 yield observation so we can calculate the difference.
egen will assign the maximum value of lny0 to all observations
that have the same prcl_id.**
egen lny10 = max(lny0), by(prcl_id)
gen lny11 = logyield
gen lny1011 = lny11 - lny10 ** calculate the first difference **
```

2. Do a t-test to compare the differences in the 2010 and 2011 log yield (`ln1011`) between parcels in the irrigation program and those not in the program. In other words, do a differences-in-differences estimate. Compare your results with those in Question 1. If they are different, explain why. Which estimate do you think is closer to the truth, and why?

Now, do the difference-in-differences estimation in a regression context. Use `egen` to construct a dummy variable indicating that the parcel was in the treatment group in 2011 (note that for each parcel, the value of this variable will be the same in 2010 and 2011). Also construct a dummy variable for 2011 and for the interaction of being in the treatment group and the year being 2011.

```
gen treat=1 if irr==1
egen treatgroup=max(treat), by(prcl_id)
replace treatgroup=0 if treatgroup==.
gen y2011=1 if sur_yr==2011
replace y2011=0 if y2011==.
gen treat2011=treatgroup*y2011
```

Let's also create a balanced panel, which is one where each plot has two observations. We can do this by using the duplicates command and dropping all the observations where there is no match.

```
duplicates tag prcl_id, generate(dup1011)
tab dup1011
keep if dup1011==1
```

3. What do the coefficients in your regression mean?
4. Add the following control variables and re-estimate the equation: `loglabor logfert logmech logpest ageH genderH sizehh logaindex loglindex logtot_acre logdist`. What happens to your estimates? Why?
5. You can also do DiD using fixed effects. The command is `xtreg`.

    To use `xtreg`, you first have to tell Stata that you have loaded a panel data set and what variable represents a unit of observation and what represents time.

    ```
    xtset prcl_id sur_yr
    xtreg logyield y2011 treat2011, fe
    ```

    Redo Questions 3 and 4 using the `xtreg` command. The fixed effect will be parcel fixed effect. Do your fixed effects differ from the results in Questions 3 and 4? Explain.

    What happens if you add your variable `treatgroup` to the regression? Why?

As a comparison to our DiD and FE regressions, let's try to apply propensity score matching to this data set. To conduct the PSM we need to generate a new data set with only data from 2010 but with an indicator variable equal to one if the parcel was treated in 2011. You can use the variable `treatgroup` to do this.

Save the data set with the 2010 and 2011 observations. Then keep only the observations from 2010.

Now run the `pscore` command on the irrigation data matching on our control variables, except `logdist`. Be sure to select the common support option with `comsup` and create a pscore id called `pscore1`.

6. Is the balancing principle satisfied?

Now we need to merge the matched households in the baseline year back into the panel data:

```
keep if pscore1~=. **only those with nonmissing pscore1 were good
matches **
keep prcl_id pscore1 **this keeps only the prcl_id & pscore1
variables **
sort prcl_id **Stata requires you to sort data sets by the
variable you will use to match before matching **
```

Then merge this data with the data set you that we saved prior to Question 1. Finally, in this merged data set, keep only the merged observations. With this data set carry out the DiD method as before using the FE implementation. You will need to select only those observations that fall within the common support. Your Stata output will give you the minimum and maximum values of pscore1 that fall within the common support. So keep only those cases with pscores within the common support.

The command for merging is

```
merge 1:m prcl_id using "<</Users/MaryArends-
Kuenning/Downloads/India_Prod_Data_exer3a.dta>>"
keep if _merge==3
```

7. How do the number of observations change across the DiD, FE, and FE using PSM approaches? Does this have an effect on the external validity of our impact assessment?
8. Comparing results between DiD, FE, and FE using PSM. Does one have stronger internal validity? Why?

Treatment Effects

We covered a variety of treatment effects in the session this morning. It turns out that we have already implemented several of the treatment effect estimates in the last two days. When we measured the impact of the index insurance intervention, we were measuring the Average Treatment Effect (ATE). When we measured the impact of the index insurance using propensity score matching, we were measuring the Average Treatment Effect on the Treated (ATT). The section which we just concluded, in which used the difference-in-difference estimator, was calculating the ATE.

Now, just because we measured the ATE or ATT does not mean those measured effects are unbiased. This is where the real difficulty of impact assessment lays. We can implement the methods correctly but if our assumptions are not valid our methods will correctly measure a biased effect. The DiD estimator is a perfect example. While we calculated the ATE using DiD in Question 3 that ATE is only accurate under the assumption of parallel trends.

9. Comparing results between DiD, FE, and FE using PSM. Which method do you think is the most appropriate one to estimate the impact of parcel participation in the irrigation program on yield? Are any of these estimates biased compared to the true effect? Why?

As we have seen, when we have experimental data we can be pretty sure that the ATE we estimate is the unbiased or true effect of the treatment. In this case we have strong internal validity but our results are only externally valid to the extent that our randomized sample is representative of some larger population. When we use DiD or FE we require the assumption of parallel trends, which in most non-experimental data will not be a valid assumption. Using PSM to construct a matched control group and then conducting DiD relaxes the parallel trend assumption, assuming that parcels with similar characteristics experience similar trends. When we combined PSM and DiD we can be pretty sure that the ATT we estimate is the unbiased or true effect of the treatment on the treated. In this case, like with RCTs, our results will be internally valid but external validity will be limited depending on how much data (and why that data) was excluded from the region of common support.

An alternative to the ATE and ATT is the Local Average Treatment Effect (LATE). As we saw this morning the LATE is the same as the instrumental variable (IV) estimator. We will spend several hours discussing and estimating IV regressions on Day 4. So, we will put off estimating the LATE until tomorrow. Today, we will just note that, under the necessary set of identification assumptions, an IV allows us to estimate the unbiased or true effect of the treatment on the subpopulation of those who participated in the program due to a change in the value of the IV. In this case, like with RCTs, our results will be internally valid. Additionally, since, in general, an IV's effect on the population should be random and if our data comes from a large, randomly selected, observational study (like the VDSA) we will also have strong external validity.

The final treatment effect, and the one we will consider in detail today, is the Marginal Treatment Effect (MTE). As we said this morning, the MTE measures the unbiased or true effect of the treatment for participants at the margin of entering treatment, given a set of observed characteristics. It is similar to both PSM and IV in that we calculate the probability of assignment to treatment (the propensity score) based on an instrument (IV). In other words, we measure the effect of the treatment on two similar or "matched" individuals where the only difference between them is that one of the individual participated in the treatment due to a change in the value of the IV.

To do measure to MTE, we will need the `margte` command. You can find it by typing `findit margte` into your command window. When the results come up, click on the newest update for the `margte` package: `st0331_1`.

```
search for margte

        Keywords:  margte
          Search:  (1) Official help files, FAQs, Examples, SJs, and STBs
                   (2) Web resources from Stata and from other users


Search of official help files, FAQs, Examples, SJs, and STBs

Web resources from Stata and other users

(contacting http://www.stata.com)

2 packages found (Stata Journal and STB listed first)
--------------------------------------------------

st0331_1 from http://www.stata-journal.com/software/sj14-3
    SJ14-3 st0331_1. Update: Marginal treatment... / Update: Marginal
    treatment effects / by Scott Brave, Federal Reserve Bank of Chicago, /
    Chicago, IL / Thomas Walstrum, University of Illinois at / Chicago and
    Federal Reserve Bank of / Chicago, Chicago, IL / Support:

st0331 from http://www.stata-journal.com/software/sj14-1
    SJ14-1 st0331. Marginal treatment effects / Marginal treatment effects /
    by Scott Brave, Federal Reserve Bank of Chicago, / Chicago, IL / Thomas
    Walstrum, University of Illinois at / Chicago and Federal Reserve Bank of
    / Chicago, Chicago, IL / Support:  sbrave@frbchi.org, /

(click here to return to the previous screen)
```

Then install the package by clicking install.

```
ESCRIPTION/AUTHOR(S)
    Update: Marginal treatment effects
    by Scott Brave, Federal Reserve Bank of Chicago,
        Chicago, IL
      Thomas Walstrum, University of Illinois at
        Chicago and Federal Reserve Bank of
        Chicago, Chicago, IL
    Support:  sbrave@frbchi.org,
              twalstrum@frbchi.org
    After installation, type help margte,
      margte_dgps, and locpoly2

NSTALLATION FILES                               (click here to install)
    st0331_1/margte.ado
    st0331_1/margte.sthlp
```

Once the package is installed we want to estimate the MTE of irrigation on a parcel for households with a similar set of characteristics. Use the same data set that we have used throughout the day, the set that only contains observations from 2010 and 2011 in season 1. The format of the MTE regression is as follows:

```
margte depvar_o indepvars_o, treatment(depvar_t indepvars_t)
       semiparametric first common savepr
```

In this equation `depvar_o` is our outcome of interest, `indepvars_o` are covariates that we want to control for because they might affect the outcome, `depvar_t` is our treatment variable, and

`indepvars_t` are covariates that affect treatment, i.e., our matching variables. The other terms set options that we want to include in the regression. Type `help margte` for a full set of options.

Set up the following MTE regression

```
margte  logyield  loglabor  logfert  logmech  logpest  ageH  genderH
    sizehh logaindex loglindex, treatment(irr average ageH genderH
    sizehh   logaindex   loglindex   logtot_acre   logdist   )
    semiparametric first common savepr
```

10. Examine the variables included in the first stage equation as well as the second stage equation. Why have we excluded some variables in the first stage from the second and vice-versa? Why are these two equations not exactly the same nor completely different?

Run the MTE regression as defined above. When `margte` runs it will put out two sets of results. The second set will take a few minutes to generate since Stata needs to bootstrap the standard errors.

11. What do you learn from the first stage estimates? How do these estimates differ from the estimates generate by the `pscore` command?
12. Examine the graph of the propensity scores. What does this graph tell you?

Once the bootstrapping has finished, Stata will print the second stage results and a second graph.

13. What do you learn from the second stage estimates? How does these compare to the DiD, FE, and FE using PSM estimates.
14. Examine the graph of the MTE and ATE. What does this graph tell you? Are returns to irrigation heterogeneous?