Applied Econometrics - Policy Evaluation 2

Note on Panel Data

- Consider again the panel data model using the `jtrain` dataset (job-training grants given to firms):

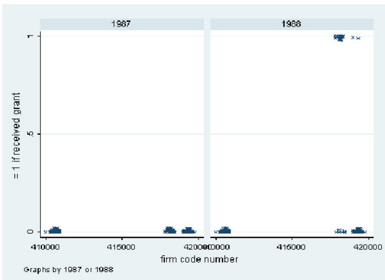
$$\texttt{lscrap}_{it} = \beta_0 + \delta_0 \texttt{y88}_t + \beta_1 \texttt{grant}_{it} + a_i + u_{it}, \quad t = 1, 2$$

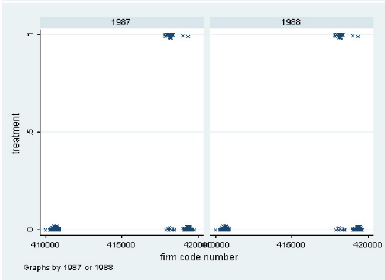
- Compare this to the model we used for repeated cross sections:

$$E(Y) = \beta_0 + \beta_1 D_{treatment} + \beta_2 D_{after} + \beta_3 (D_{after} \cdot D_{treatment})$$

- Consider replicating this with panel data by creating a `treatment` variable which is 1 in both periods if the firm is “treated”.

Note on Panel Data





Note on Panel Data

- If we run fixed effects panel on this we get:

```
. xtreg lscrap  d88 treatment interact,fe

Fixed-effects (within) regression               Number of obs      =       108
Group variable: fcode                           Number of groups   =        54

R-sq:  within  = 0.1392                         Obs per group: min =         2
       between = 0.0049                                        avg =       2.0
       overall = 0.0006                                        max =         2

                                                F(2,52)            =      4.20
corr(u_i, Xb)  = -0.0674                        Prob > F           =    0.0203

------------------------------------------------------------------------------
      lscrap |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         d88 |  -.0574357    .097206    -0.59   0.557    -.2524938    .1376224
   treatment |  (dropped)
    interact |  -.3170579   .1638751    -1.93   0.058    -.6458975    .0117816
       _cons |   .5974341   .0553369    10.80   0.000     .4863924    .7084757
-------------+----------------------------------------------------------------
     sigma_u |  1.4833025
     sigma_e |   .4066418
         rho |  .93009745   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(53, 52) =    26.22              Prob > F = 0.0000
```

Note on Panel Data

- Compare this to our original result:

```
. xtreg lscrap d88 grant,fe

Fixed-effects (within) regression               Number of obs      =       108
Group variable: fcode                           Number of groups   =        54

R-sq:  within  = 0.1392                         Obs per group: min =         2
       between = 0.0049                                        avg =       2.0
       overall = 0.0006                                        max =         2

                                                F(2,52)            =      4.20
corr(u_i, Xb)  = -0.0674                        Prob > F           =    0.0203

------------------------------------------------------------------------------
      lscrap |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         d88 |  -.0574357    .097206    -0.59   0.557    -.2524938    .1376224
       grant |  -.3170579   .1638751    -1.93   0.058    -.6458975    .0117816
       _cons |   .5974341   .0553369    10.80   0.000     .4863924    .7084757
-------------+----------------------------------------------------------------
     sigma_u |  1.4833025
     sigma_e |   .4066418
         rho |  .93009745   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(53, 52) =    26.42              Prob > F = 0.0000
```

Note on Panel Data

- We can compare the pooled regressions as well. This is the original result:

```
. regress lscrap d88 grant

      Source |       SS       df       MS              Number of obs =     108
-------------+------------------------------           F(  2,   105) =    0.18
       Model |  .810536068     2  .405268034           Prob > F      =  0.8378
    Residual |  240.098947   105  2.28665664           R-squared     =  0.0034
-------------+------------------------------           Adj R-squared = -0.0156
       Total |  240.909484   107   2.2514905           Root MSE      =  1.5122

------------------------------------------------------------------------------
      lscrap |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         d88 |  -.1889081   .3281441    -0.58   0.566    -.8395572     .461741
       grant |   .0566004     .43091     0.13   0.896    -.7978145    .9110152
       _cons |   .5974341   .2057802     2.90   0.005     .1894099    1.005458
------------------------------------------------------------------------------
```

Note on Panel Data

- And this is the result with the `treatment` dummy (which picks up the “group fixed effects” of those that are treated but is unable to control for firm specific fixed effects):

```
. regress lscrap d88 treatment interact

      Source |       SS       df       MS              Number of obs =     108
-------------+------------------------------           F(  3,   104) =    0.37
       Model |  2.52993689     3  .843312297           Prob > F      =  0.7763
    Residual |  238.379547   104  2.29211103           R-squared     =  0.0105
-------------+------------------------------           Adj R-squared = -0.0180
       Total |  240.909484   107   2.2514905           Root MSE      =   1.514

------------------------------------------------------------------------------
      lscrap |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         d88 |  -.0574357   .3619085    -0.16   0.874    -.7751139    .6602425
   treatment |   .3736583   .4314236     0.87   0.388    -.4818709    1.229187
    interact |  -.3170579   .6101251    -0.52   0.604    -1.526959     .892843
       _cons |   .4659617    .255908     1.82   0.072    -.0415134    .9734368
------------------------------------------------------------------------------
```

Repeated Cross Sections

Partial Compliance

- Suppose not all of the "treated" actually undergo treatment. For example, housing vouchers might not be used etc.
- Then we have a difference between the *intended* treatment group and the *actual* treatment group which may vary systematically.
- Let z be a dummy indicating that whether treatment was intended (i.e. a voucher was given)
- Let w be a dummy indicating whether the treatment was applied (i.e. voucher used)
- w may be correlated with the error (e.g. more motivated individuals more likely to use the voucher etc).

Repeated Cross Sections

Partial Compliance

- Then, if z was allocated randomly, we can use this as an instrument for w. Why is this a good instrument?
  - z will be correlated with w
  - z should not be correlated with the error (due to randomisation).
- Note that we will typically have data on both z and x.

Repeated Cross Sections

Partial Compliance

- We do IV in the usual manner e.g. for a simple case of one instrument:
  - Regress z on x and save fitted values
  - Use the fitted values in place of x in the main regression.

Repeated Cross Sections

Example

- Angrist (1990), *AER* 80(3). Looked at effect on earnings of enrolment in military.
- Problem: military enrolment is by choice and military only accepts certain types of individuals so entry is not random.
- Uses Vietnam War draft to circumvent these problems. Random lottery allocation (lowest numbers were drafted).
- Draft numbers used as an instrument since actual enrolment still dependent on physical requirements etc.
- Draft numbers correlated with enrolment but are random.

Interpreting Treatment Coefficients

Differing treatment effects

- The above methods assume that the effect of treatment is the same for all individuals (measured by the relevant coefficient).
- This might not be the case e.g. high cholesterol patients may benefit more from cholesterol reducing drugs etc.
- Two important cases:
  1. The treatment depends on a measurable attribute (a regressor)
  2. The treatment depends on some unobserved attribute (e.g. motivation)

Interpreting Treatment Coefficients

Measurable Heterogeneity

- The case of dependence on a measurable attribute is easy to handle.
- Create more interaction terms between the treatment variable and the regressors thought to influence treatment.
- Suppose `sex` is thought to influence the treatment. Then we have, for example,:

$$E(Y) = \beta_0 + \beta_1 D_{treatment} + \beta_2 D_{after} + \beta_3 (D_{after} \cdot D_{treatment}) + \beta_4 (\texttt{female} \cdot D_{treatment} \cdot D_{after})$$

- How can you interpret these coefficients?
- (Note: Can be done with panel data too.)

Interpreting Treatment Coefficients

Unmeasurable Heterogeneity

- The case of unmeasurable heterogeneity is a bit more involved. Each individual may now have their own intercept and slope coefficients.
- In this case, what we estimate (consistently) is the average treatment effect *provided the treatment is randomly allocated.*
- If not (e.g. partial compliance), then we need to use a valid instrument as noted above but the interpretation is complex.

Interpreting Treatment Coefficients

Unmeasurable Heterogeneity

- In this case, what we estimate is a local average treatment effect.
- This is the average treatment effect for those individuals who have been treated and those that are like them.
- It does not reflect the treatment effect of individuals that are not treated (e.g. if those with less motivation do not use their vouchers the coefficient says nothing about these people)