

Applied Econometrics - Policy Evaluation

Introduction

- Do poor children do better educationally if they live in the suburbs rather than inner cities?
- Suppose we could *randomly* assign poor children to either suburbs or inner cities.
- Assignment to suburbs will be called *treatment*.
- The inner city assigned children are the *control*.
- In general, this is not possible for a variety of reasons.

Introduction

- If it was possible, we could assign a dummy variable D_i which is 1 if a child is assigned to a suburb and 0 otherwise.
- Then we could estimate:

$$y_i = \beta_0 + \beta_1 D_i + u_i$$

and interpret the coefficient β_1 as the *treatment effect*.

Introduction

- If the sample is large enough, unobserved traits will be relatively unimportant (i.e. little correlation between treatment and the disturbance).
- In practice, β_1 will be biased for various reasons because such randomised experiments cannot be performed.

Randomised Experiments

- Consider the “Moving to Opportunity” program in the US
- Low-income families offer to take part in program.
- They are randomly assigned to treatment or control groups, either:
 - a voucher for living in a low (< 10%) poverty area
 - participation in control group

Randomised Experiments

Possible sources of bias:

- Non-response bias (failure of individuals to report outcomes)
- Attrition bias (individuals left the experiment)
- Sample-selection bias (if the population sampled from is not the one we want e.g. some poor households are not in public housing)
- Self-selection bias (not all voucher recipients succeed in finding a place to live in low poverty areas - motivation, connections etc)

Correlation problem

- These biases introduce correlation between the regressor and disturbance leading to inconsistency.
- This link can be broken by *instrumental variables* (see previous course). This breaks the link *ex post* by isolating part of the variation in the regressor that is uncorrelated with u_i .
- Experiments break the link *ex ante* (i.e. by “design”).

Natural Experiments

- Natural experiments occur when “randomisation” happens by chance.
- This occurs when some exogenous event (e.g. a government policy) changes the environment in which some individuals, families, firms operate (treatment group) and leaves others unchanged (control group).
- For example, an educational change that takes place in some cities and not in others.
- But how can we control for the differences between cities?

Difference-in-differences

- We need *two* periods to control for systematic differences of this kind.
- We then have four groups: control group (before and after) and treatment group (before and after).
- The aim is to remove the systematic difference between the cities and estimate the difference between the treatment and control groups.
- Called the “difference-in-difference” approach.

Simple Example

- Simple example - suppose that average outcomes are:

	Before	After
Treatment Group	40	80
Control Group	20	30

- Just looking at the “after” averages, we would estimate the effect of the treatment as $80 - 30 = 50$.
- But this ignores the fact that the treatment group began at a higher level (systematic differences).
- Thus we would want to estimate the treatment effect as $(80 - 40) - (30 - 20) = 30$
- This is a “difference in differences” estimate.

Kiel and McLain example

- Example: Kiel and McLain (1995), “House Prices during Siting Decision Stages”, *Jnl of Environ. Econ and Mgmt* 28, 241-287.
- Look at effect of incinerator on nearby house prices.
- Like a natural experiment - some houses are “near” and some are not. The change (incinerator) is exogenous (households do not demand it).
- Data on pre-incinerator (sold) prices (1978) and post-incinerator (sold) prices (1981). This is a repeated cross-section.

Kiel and McLain example

- A naive estimation would be to just look at 1981 and regress:

$$\ln price_{81} = \beta_0 + \beta_1 nearinc_{81} + u$$

- Stata commands:

```
use KIELMC;  
regress lnprice nearinc if y81;
```

Kiel and McLain example

- But - try running the regression for 1978 as well.
- Stata commands:
`regress lprice nearinc if y81==0;`
- Thus, even before incinerator, house prices were lower in the vicinity.

Kiel and McLain example

- The difference-in-differences estimate is ($nr = \text{near}$ and $fr = \text{far}$):

$$\delta_1 = (\overline{\text{lprice}}_{81,nr} - \overline{\text{lprice}}_{81,fr}) - (\overline{\text{lprice}}_{78,nr} - \overline{\text{lprice}}_{78,fr})$$

- This gives (use Stata command 'mean <var> if <cond>' to calculate these means):

$$\delta_1 = (11.3399 - 11.7424) - (10.9455 - 11.2854) = -0.0626$$

Kiel and McLain example

- Regression version:

$$\ln \text{price} = \beta_0 + \delta_0 \text{y81} + \beta_1 \text{nearinc} + \delta_1 (\text{nearinc} \cdot \text{y81}) + u$$

- The coefficient δ_1 is the difference-in-differences estimate we found earlier.
- Using regression, we can also calculate a standard error and do t-tests etc.

Kiel and McLain example

- Stata commands

```
regress lprice y81 nearinc y81nrinc
```

```
regress lprice y81 nearinc y81nrinc age agesq
```

```
lintst lland larea rooms baths
```

Kiel and McLain example

- Note that in this example, δ_1 is insignificant.
- We can also add other regressors to improve efficiency of the estimates. In this example, δ_1 then becomes significant.
- A shortcoming of these estimators is that other factors could have changed in the two areas which are influencing our results.

Difference-in-differences: general approach

- In general terms we have:

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

- β_1 captures the underlying difference between treatment and control groups.
- β_2 captures the underlying difference between the two time periods.
- β_3 captures the effect of the treatment.
- We can also add other variables X .

Two-period Panel Data

- Using panel data rather than a pooled cross section allows controlling for unobserved individual effects (e.g. some individuals may be more able etc). Not just aggregate/group effects as above.
- Consider the Michigan job training program (data in Stata file JTRAIN) - the cross-sectional unit is the firm.
- Some firms received job training grants (variable grant).
- The dependent variable is `lscrap` (log no of items scrapped per 100).
- Unobserved effects can be: average employee ability, managerial skill etc. Firms may be more likely to apply if they have lower average ability etc. Ignoring this effect will lead to inconsistency.

Fixed Effects

- We use the standard fixed effects model:

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

(Note that a grant is only equal to 1 in the second period (1988) unlike `nearinc` in the previous example. Thus, there are only two dummies here.)

- Use the standard methods to eradicate the fixed effects (e.g. differencing) and calculate the within estimator.
- In stata, estimate this using the `xtreg` command.
- The coefficient on `grant` is -0.317 which translates to a percentage decrease of about 27% (i.e. $100[\exp(-0.317) - 1] \approx 27$).

Fixed Effects

- Using pooled data (i.e. don't use the 'xt' regression commands in Stata - just use 'regress') we get an insignificant coefficient.
- Suggests that firm fixed effects are very important here.
- It is therefore likely that firms with lower-ability workers were more likely to get grants (since these would be the ones where the improvement would be highest).