Lecture 9: Empirical Methods in Corporate Finance II

Adam Hal Spencer

The University of Nottingham

Advanced Financial Economics 2019

Roadmap

1. Introduction
2. Instrumental Variables
3. Panel Data Methods
4. Difference in Differences Estimator
5. Conclusion

Motivation

- Last time we talked about a simple method for studying empirical relationships between variables in corporate finance: OLS and regression analysis generally.
- Endogeneity is a problem that's pervasive throughout economics though, that plagues our regression estimates with asymptotic bias.
- Is there any hope for statistical analysis in economics?

Motivation

- Yes!
- Before you can find the remedy, you need to ask the questions: what variable is endogenous? Why? What are the implications with regard to bias?
- What are the alternative hypotheses, about which one should be concerned?

Roadmap

1 Introduction

2 Instrumental Variables

3 Panel Data Methods

4 Difference in Differences Estimator

5 Conclusion

General setting

- A common way to deal with endogeneity.
- Say that we're considering a framework of

$$y_i = \beta_0 + \beta_1 x_{1,i} + ... + \beta_M x_{M,i} + u_i \tag{1}$$

where $Cov(x_{k,i}, u_i) \neq 0$ for some $k \in \{1, 2, ..., M\}$.

- Generally means that all of our coefficient estimates will be biased.
- Unless $x_{k,i}$ happens to be uncorrelated with the other regressors.

General setting

- An instrumental variable (denoted z_i) for $x_{k,i}$ satisfies two conditions

  (1) Relevance condition.

  (2) Exclusion condition.

Relevance Condition

- Requires that the partial correlation between the instrument and endogenous variable not be zero.
- Means formally that the coefficient γ in the regression

$$x_{k,i} = \alpha_0 + \alpha_1 x_{1,i} + ... + \alpha_{k-1} x_{k-1,i} + \alpha_{k+1} x_{k+1,i} + ... + \alpha_M x_{M,i} + \gamma z_i + v_i$$

be non-zero.

- Says that the endogenous variable and the instrument are correlated after netting-out the effects of all other exogenous variables.

Exclusion Condition

- Says that z_i's only influence on the outcome variable of interest is through the endogenous regressor.
- That is: $Cov(z_i, u_i) = 0$ where u_i is the residual from (1) above.
- What does it mean if this condition is not satisfied?
- Would mean that the instrument is also endogenous! This is the same problem that we're trying to fix.

Multiple variables

- One can use multiple instruments for an endogenous variable.
- Both conditions need to be satisfied for each instrument though.
- Relevance condition can be done with joint test of statistical significance.
- May also have multiple endogenous variables.
- In this case, need at least as many instruments as we have endogenous variables.

Examples in corporate finance

- Bennedsen et al. (2007): does replacing an outgoing CEO with a family member hurt firm performance in family firms?
- CEO succession is likely correlated with things that affect performance also.
- E.g. non-family CEO may come in during bad times, family CEO during good times.
- Need exogenous variation in the CEO succession decision.
- They use gender of the first-born child of a departing CEO.
- Show that CEOs with boy-first families are significantly more likely to appoint a family CEO.
- Gender is a biological thing, likely uncorrelated with the firm's performance.

Estimation

- Common approach is to use two-stage least squares (2SLS).
- **1st stage:** regress the endogenous variable ($x_{k,i}$) on the exogenous variables and instruments. Gives fitted values $\hat{x}_{k,i}$: those predicted by the regression.

$$\hat{x}_{k,i} = \hat{\alpha}_0 + \hat{\alpha}_1 x_{1,i} + ... + \hat{\alpha}_{k-1} x_{k-1,i} + \hat{\alpha}_{k+1} x_{k+1,i} + ... + \hat{\alpha}_M x_{M,i} + \hat{\gamma} z_i$$

- **2nd stage:** use the fitted values ($\hat{x}_{k,i}$) to stand-in for the endogenous variable ($x_{k,i}$) in the regression of y_i (outcome) on all the right-side variables.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1,i} + ... + \hat{\beta}_k \hat{x}_{k,i} + ... \hat{\beta}_M x_{M,i}$$

Estimation

- The residual in the first stage contains all the junk that's correlated with the outcome variable.
- We disregard that and just keep the exogenous component, (since recall z_i is doesn't affect the outcome variable directly).
- The idea is that the fitted values $\hat{x}_{k,i}$ contain only variation in the endogenous variable that is exogenous to the regression system.
- We've removed the "evil" endogenous part.
- The regression coefficient will be consistent.
- Introduces more noise into the estimation though: need to correct the standard error estimates.

Estimation

- Easy to run two stage least squares in Stata.

- If we're dealing with a model

$$y_i = \beta x_i + u_i$$

where z_i is an instrument for x_i then use the Stata command

```
ivregress 2sls y (x = z)
```

Roadmap

1. Introduction
2. Instrumental Variables
3. Panel Data Methods
4. Difference in Differences Estimator
5. Conclusion

Panel setup

- A panel dataset for firms follows the same firm for multiple periods of time.
- Are there unobservable attributes at the firm-level that are time-invariant?
- We could control for these with fixed effects.

Fixed effects

- A fixed effects panel regression takes the form

$$y_{it} = \beta x_{it} + \alpha_i + u_{it}$$

  where parameter α_i is a firm-specific, unobservable and time-invariant fixed effect.

- We can't observe this fixed effect.

- I.e. it's likely an omitted variable (absorbed into the residual term).

- That is: there might be omitted variables that are correlated with the firm itself.

- So how is it helpful conceptually?

Least Squares with Dummy Variables

- Why don't we just use a dummy for each firm (each i)?
- Run the following regression

$$y_{it} = \beta x_{it} + \sum_{j=1}^{N} \alpha_j d_j + u_{it}$$

where $d_j = 1$ if $j = i$ and 0 otherwise.

- In Stata, use the command

```
xi: regress y x i.firm
```

Demeaning

- Another approach is to demean the variables across time.

- Consider

$$y_{it} - \bar{y}_i = \beta(x_{it} - \bar{x}_i) + (u_{it} - \bar{u}_{it})$$
$$\tilde{y}_{it} = \beta \tilde{x}_{it} + \tilde{u}_{it}$$

  where notice that the mean variables (with bars over the top) are firm-dependent (note $\bar{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it}$).

- Why does this work? Because α_i is time invariant — its mean is the same as itself. Drops-out in the differencing.

- In Stata, use the command

```
xtreg y x, fe
```

  should offer the same results as the dummy approach.

Roadmap

1. Introduction
2. Instrumental Variables
3. Panel Data Methods
4. Difference in Differences Estimator
5. Conclusion

Treatment Effects

- Difference in differences is a method used to estimate the impact of a treatment effect.
- Recall that simply taking a difference in the averages for treated and untreated samples leads to a selection bias term.
- The key is to take another difference!

Treatment Effects

- Consider a two-period example.
- In 1987 Arizona passed anti-takerover legislation, while Connecticut had not.
- Assume we have data pre and post reform: start of 1986 and end of 1987.
- Arizona firms are the treatment group and Connecticut firms are the control group.

Treatment Effects

- The regression model for the DD estimator is given by

$$y = \beta_0 + \beta_1 d \times p + \beta_2 d + \beta_3 p + u$$

  where

  - d is the treatment assignment variable equal to 1 if an Arizona firm and 0 if a Connecticut firm.
  - p is the post-treatment indicator equal to 1 if datum is from 1987 (post-reform) and 0 if from 1986 (pre-reform).

- β_2 captures differences across the two states, while β_3 captures differences across time.

- β_1 is our object of interest here: how did the policy change affect the Arizona firms?

Treatment Effects

- What are the possible combinations of outcomes?
  - $(d = 0) \wedge (p = 0) \Rightarrow y = \beta_0 + u$.
  - $(d = 1) \wedge (p = 0) \Rightarrow y = \beta_0 + \beta_2 + u$.
  - $(d = 0) \wedge (p = 1) \Rightarrow y = \beta_0 + \beta_3 + u$.
  - $(d = 1) \wedge (p = 1) \Rightarrow y = \beta_0 + \beta_1 + \beta_2 + \beta_3 + u$.

where $\wedge$ is shorthand for "and".

Treatment Effects

- Assume that $\mathbb{E}[u|p,d] = 0$: that is, the expectation of the residual is zero, irrespective of the values of d and p.

- Conditional expectations are all given then by
  - $\mathbb{E}[y|d=0,p=0] = \beta_0$
  - $\mathbb{E}[y|d=1,p=0] = \beta_0 + \beta_2$
  - $\mathbb{E}[y|d=0,p=1] = \beta_0 + \beta_3$
  - $\mathbb{E}[y|d=1,p=1] = \beta_0 + \beta_1 + \beta_2 + \beta_3$

- So our DD estimator is given by

$$(\mathbb{E}[y|d=1,p=1] - \mathbb{E}[y|d=0,p=1])$$
$$-(\mathbb{E}[y|d=1,p=0] - \mathbb{E}[y|d=0,p=0]) = (\beta_1+\beta_2) - (\beta_2) = \beta_1$$

Treatment Effects

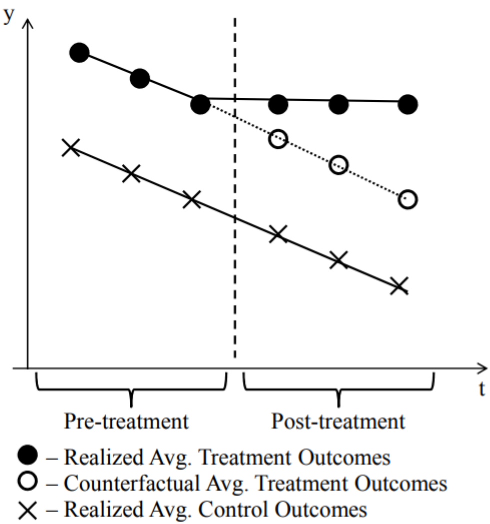
- Approximate this difference using sample averages!

Identifying assumption

- For this to work, we need the parallel trends assumption to hold.
- Means that the time-trend the two groups would have followed would be the same in the absence of the treatment.
- In the counter-factual, the trend would have been the same for the treatment and control groups.
- Then we can attribute the difference to the treatment effect.

Identifying assumption

Figure 1: **Difference-in-Differences Intuition**



Identifying assumption: robustness tests

- We don't observe the counterfactual outcome for the treatment group.
- Test: repeat the DD estimation for previous years (where the treatment was not present).
- Estimated treatment effect should be no different from zero statistically.
- Other tests possible as well.

Roadmap

1. Introduction
2. Instrumental Variables
3. Panel Data Methods
4. Difference in Differences Estimator
5. Conclusion

Summary

- Linear regressions are easy to implement in Stata and can be very informative.
- Selection bias and endogeneity bias can be issues though.
- To deal with endogeneity: IV or panel regressions are two possible tools for fixing the issue.
- Difference in difference regressions can be used to identify treatment effects under some fairly strong assumptions.