第八讲　货币经济学经验研究方法

张国雄

上海交通大学安泰经管学院

2015年11月17日

内容提要

- Propensity Score Matching
- Difference in Difference
- Regression Discontinuity
- 单元时间序列方法
- 向量自回归方法

- PSM的核心思想在于假设conditional on X_i, 个体的outcome与treatment无关

$$E(Y_{i0}|D_i = 1, X_i) = E(Y_{i0}|D_i = 0, X_i)$$

- 所以treatment effect可以表示为

$$ATT = E(Y_{i1}|D_i = 1, X_i) - E(Y_{i0}|D_i = 1, X_i) = E(Y_{i1}|D_i = 1, X_i) - E(Y_{i0}|D_i = 0, X_i)$$

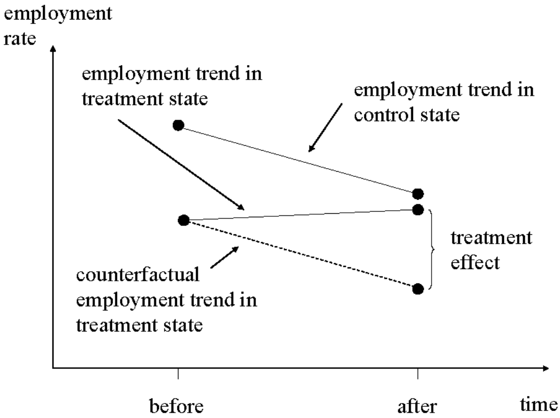
- 更practical地可以表示为

$$ATT = E(Y_{i1}|D_i = 1, p(X_i)) - E(Y_{i0}|D_i = 0, p(X_i))$$

Propensity Score Matching (PSM)

- 两个主要的问题
  - 如何选取X_i具有"随意性"
  - 如何决定两个个体的propensity score足够接近具有"随意性"

Difference in Difference (DID)



- DID的核心思想在于假设实验组(s=1)和对照组(s=2)有common trend:

$$Y_{ist} = \alpha_s + \lambda T + \beta D_{st} + \epsilon_{ist}$$

- 所以有

$$\beta = [E(Y_{i,s=1,t=1}) - E(Y_{i,s=1,t=0})] - [E(Y_{i,s=2,t=1}) - E(Y_{i,s=2,t=0})$$

Difference in Difference (DID)

- 或者更方便地可以创建dummy(D_s和D_t)直接用回归的形式

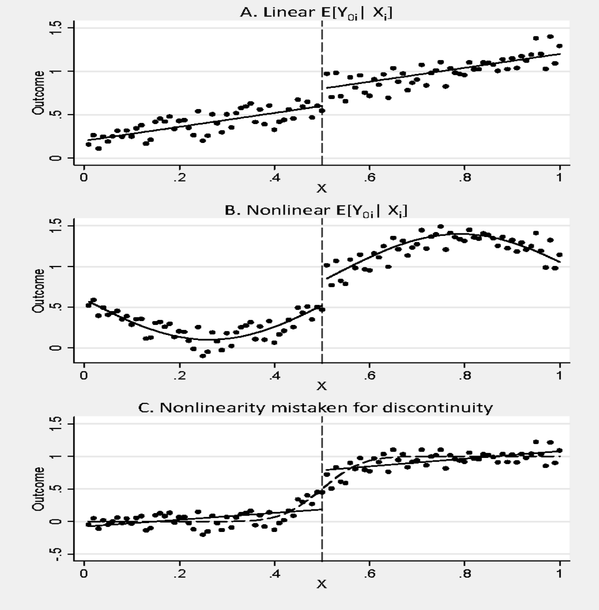
$$Y_{ist} = \alpha + \gamma D_s + \lambda D_t + \beta D_s * D_t + \epsilon_{ist}$$

- 如果两组明显没有common trend怎么办?

$$Y_{ist} = \alpha + \gamma D_s + \lambda D_t + \beta D_s * D_t + X_{ist}\delta + \epsilon_{ist}$$

- 如果没有现成地实验组和对照组怎么办?

Regression Discontinuity (RD)



Regression Discontinuity (RD)

- 假设ρ是我们关心的treatment effect

$$Y_i = f_0(x_i)(1 - D_i) + f_1(x_i)D_i + \rho D_i + \epsilon_i$$

- 我们用polinomial来拟合非线性方程$f(x_i)$

$$E(Y_{0i}|x_i) = f_0(x_i) = \alpha + \sum_{j=1}^{p} \beta_{0j}(x_i - x_0)^j$$

$$E(Y_{1i}|x_i) = f_1(x_i) = \alpha + \sum_{j=1}^{p} \beta_{1j}(x_i - x_0)^j + \rho$$

Regression Discontinuity (RD)

- 写成回归方程的形式

$$Y_i = \alpha + \sum_{j=1}^{p} \beta_{0j}(x_i - x_0)^j + \rho D_i + \sum_{j=1}^{p} (\beta_{1j} - \beta_{0j}) D_i (x_i - x_0)^j + \epsilon_i$$

- 这种参数方法的风险在于我们的拟合方程可能不够准确，这个时候可以采取非参数的方法

$$\lim_{\delta \to 0} [E(Y_i | x_0 < x_i < x_0 + \delta) - E(Y_i | x_0 - \delta < x_i < x_0)] = E(Y_{1i} - Y_{0i} | x_i = x_0) = \rho$$

- Discontinuity还可能以非deterministic的方式出现

$$P(D_i = 1|x_i) = \begin{cases} g_0(x_i), & x_i \leq x_0 \\ g_1(x_i), & x_i > x_0 \end{cases}, \quad g_1(x_i) > g_0(x_i)$$

- 这时我们可以用2SLS来进行估计

$$D_i = \gamma_0 + \sum_{j=1}^{p} \gamma_j x_i^j + \xi T_i + u_i, \quad T_i = 1(x_i = x_0) \qquad \text{First Stage}$$

$$Y_i = \mu + \sum_{j=1}^{p} \beta_j x_i^j + \rho \hat{D}_i + \nu_i \qquad \text{Second Stage}$$

- 当Discontinuity以非deterministic的方式出现时,我们也可以使用非参数的方法进行估计

$$\lim_{\delta\to 0}\frac{E(Y_i|x_0<x_i<x_0+\delta)-E(Y_i|x_0-\delta<x_i<x_0)}{E(D_i|x_0<x_i<x_0+\delta)-E(D_i|x_0-\delta<x_i<x_0)}$$
$$=E(Y_{1i}-Y_{0i}|x_i=x_0)=\rho$$

时间序列数据(time series data)

- 无论用那种时间序列方法，第一步都是看时间序列数据；
- 单时间序列数据按照稳定性可以分为三种
  - 稳定的(stationary)
  - 带有非随机的趋势的(deterministic trends)
  - 带有随机趋势的(stochastic trends)=单位根(unit roots)
- 多个时间序列数据可能还存在协整(co-integration)的问题

Stationarity: example

$$y_t = at + u_t$$
$$u_t = \rho u_{t-1} + \epsilon_t, \quad 0 < \rho \leq 1$$

- $a = 0, \rho < 1$: stationary
- $a \neq 0, \rho < 1$: deterministic trends
- $a \neq 0, \rho = 1$: unit root

协整关系(cointegration)

$$x_t = x_{t-1} + \epsilon_t$$
$$z_t = z_{t-1} + \nu_t$$

- 如果存在θ使得$x_t - \theta z_t$是stationary的，那么我们就说x_t和z_t具有协整关系

- deterministic trends： detrend
- unit root: first difference
- cointegrated: do nothing

自回归滑动平均模型(ARMA)

- 一个ARMA(p,q)模型可以表示为

$$z_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i z_{t-i} + + \sum_{j=1}^{q} \beta_j \epsilon_{t-j}.$$

- ARMA的优点在于灵活性：基本上可以模拟(fit)所有的时间序列数据，因此适合进行预测
- 但是ARMA模型在经济学分析中的用处非常有限

$$y_t = \gamma_0 + X_t\gamma + u_t.$$

其中$var(u_t) = \sigma_t^2$.

- ARCH(p)模型

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i u_{t-i}^2.$$

- GARCH(p,q)模型

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i u_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2.$$

- 同样，ARCH和GARCH比较适合做数据拟合进而做forecast，而进行经济学分析的作用不大.

- VAR分为structural form和reduced form;
- 前者包含经济学含义，用来进行经济学分析(but still not quite structural)
- 后者用来做计量估计(estimation)

- $$y_t = \alpha_y + \alpha_{y,0} i_t + \sum_{j=1}^{p} \alpha^{y}_{y,j} y_{t-j} + \sum_{j=1}^{p} \alpha^{i}_{y,j} i_{t-j} + \epsilon_{y,t}$$
  $$i_t = \alpha_i + \alpha_{i,0} y_t + \sum_{j=1}^{p} \alpha^{y}_{i,j} y_{t-j} + \sum_{j=1}^{p} \alpha^{i}_{i,j} i_{t-j} + \epsilon_{i,t}$$
- $$E\begin{pmatrix} \epsilon_{y,t} \\ \epsilon_{i,t} \end{pmatrix} = 0, \qquad cov\begin{pmatrix} \epsilon_{y,t} \\ \epsilon_{i,t} \end{pmatrix} = \begin{pmatrix} \sigma_y^2, 0 \\ 0, \sigma_i^2 \end{pmatrix}$$
- 因为这两个方程都有明显的内生性所以不能直接用OLS估计

从structural form到reduced form

- 矩阵的形式为

$$\begin{pmatrix} 1, -\alpha_{y,0} \\ -\alpha_{i,0}, 1 \end{pmatrix} \begin{pmatrix} y_t \\ i_t \end{pmatrix} = \begin{pmatrix} \alpha^y_{y,1}, \alpha^i_{y,1} \\ \alpha^y_{i,1}, \alpha^i_{i,1} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ i_{t-1} \end{pmatrix} + \dots$$

$$+ \begin{pmatrix} \alpha^y_{y,p}, \alpha^i_{y,p} \\ \alpha^y_{i,p}, \alpha^i_{i,p} \end{pmatrix} \begin{pmatrix} y_{t-p} \\ i_{t-p} \end{pmatrix} + \begin{pmatrix} \epsilon_{y,t} \\ \epsilon_{i,t} \end{pmatrix}$$

- 或者可以表述成

$$A_0 \begin{pmatrix} y_t \\ i_t \end{pmatrix} = A_1 \begin{pmatrix} y_{t-1} \\ i_{t-1} \end{pmatrix} + \dots + A_p \begin{pmatrix} y_{t-p} \\ i_{t-p} \end{pmatrix} + \begin{pmatrix} \epsilon_{y,t} \\ \epsilon_{i,t} \end{pmatrix}$$

从structural form到reduced form

- 两边同乘以A_0^{-1}:

$$\begin{pmatrix} y_t \\ i_t \end{pmatrix} = A_0^{-1}A_1 \begin{pmatrix} y_{t-1} \\ i_{t-1} \end{pmatrix} + \ldots + A_0^{-1}A_p \begin{pmatrix} y_{t-p} \\ i_{t-p} \end{pmatrix} + A_0^{-1} \begin{pmatrix} \epsilon_{y,t} \\ \epsilon_{i,t} \end{pmatrix}$$

- 假设$\sigma_y^2 = \sigma_i^2 = 1$, 那么

$$u_t = A_0^{-1} \begin{pmatrix} \epsilon_{y,t} \\ \epsilon_{i,t} \end{pmatrix} \qquad \Sigma_u = A_0^{-1} {A_0^{-1}}'.$$

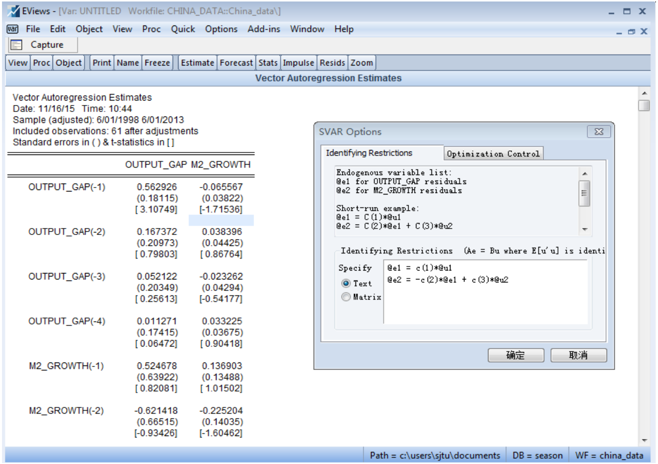
从reduced form到structural form

- reduced form已经没有内生性，可以用OLS直接估计
- 估计出来的reduced form可以直接拿来做forecast，但是如果要进行经济学分析还必须解出structural form
- 所以问题就是如何从$\hat{\Sigma}_u$推出$\hat{A}_0^{-1}$
- 所以在这里我们一般需要加入identification assumption (restriction)
- 最常见的restriction是A_0^{-1}中部分元素为0: Choleski Decomposition, 背后的经济学含义是某些变量不受当期内其他变量的影响(所以ordering很重要)

做structural var的几个步骤

- 选定变量：关键是一定要有经济学含义
- 确定变量的形式(level, log level, log difference, HP filter): 用log level一般比较保险，除非有明显的trend像中国的GDP就用HP filter
- 确定滞后项的个数:一般以一年为单位，所以季度数据就用四个，当然也可以用BIC或者AIC来做选择
- 回归估计，得到参数
- 做impulse response和variance decomposition
- 向量误差修正模型(vector error correction model)?

Structural VAR in Eviews



Structural VAR in Eviews

Var: UNTITLED Workfile: CHINA_DATA::China_data\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Impulse | Resids | Zoom

Structural VAR Estimates

Structural VAR Estimates
Date: 11/16/15 Time: 10:44
Sample (adjusted): 6/01/1998 6/01/2013
Included observations: 61 after adjustments
Estimation method: method of scoring (analytic derivatives)
Convergence achieved after 1 iterations
Structural VAR is just-identified

Model: Ae = Bu where E[uu']=I
Restriction Type: short-run text form
@e1 = c(1)*@u1
@e2 = -c(2)*@e1 + c(3)*@u2
where
@e1 represents OUTPUT_GAP residuals
@e2 represents M2_GROWTH residuals

	Coefficient	Std. Error	z-Statistic	Prob.
C(2)	0.008572	0.026994	0.317562	0.7508
C(1)	0.031018	0.002808	11.04536	0.0000
C(3)	0.006540	0.000592	11.04536	0.0000

Log likelihood 345.5761

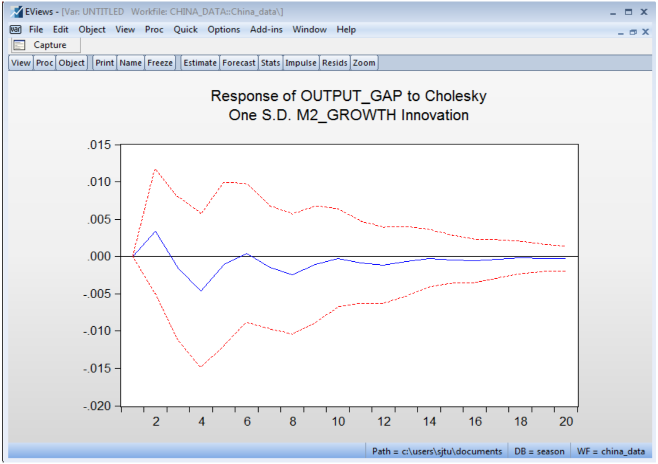
Estimated A matrix:

1.000000	0.000000
0.008572	1.000000

Estimated B matrix:

0.031018	0.000000
0.000000	0.006540

Structural VAR in Eviews



Structural VAR in Eviews