Additional topics

Paul Schrimpf

UBC
Economics 326

April 3, 2018

Section 1

Panel Data

Panel Data

- Same units observed repeatedly over time
- Examples
  - Same individuals surveyed each year (NLSY, PSID, etc)
  - Same firms surveyed each year
  - Cities in Levitt (1997)
  - States in Autor, Manning, and Smith (2016)
- Seeing the same units repeatedly introduces dependence, but also allows us to deal with unobserved heterogeneity in a richer way

Panel Data

- N units, indexed by i, observed for T periods, indexed by t

$$y_{it} = \beta_0 + \beta_1 x_{it} + \underbrace{\epsilon_{it}}_{\equiv \alpha_i + \gamma_t + u_{it}}$$

- α_i = fixed (or individual) effects
- γ_t = time effects
- Consistency of OLS needs $\mathrm{E}[\epsilon x] = 0$
- Fixed effects or first differencing needs $\mathrm{E}[ux] = 0$
  - Time invariant unobserved individual characteristics, captured by α_i, can be correlated with x
  - Time varying unobserved variables that affect everyone, captured by γ_t, can be correlated with x

Section 2

Weak instruments

Weak instruments

- References: Imbens and Wooldridge (2007), Angrist and Pischke (2014) pages 145-146
- In applications, instruments are sometimes barely relevant, i.e. $\widehat{\text{Cov}}(z, x) \neq 0$, but $\widehat{\text{Cov}}(z, x)$ not far from 0 relative to its standard error
- Implications:
  - Finite sample bias of $\hat{\beta}^{2SLS}$ is large
  - Usual standard error leads to inaccurate inference (wrong standard error, incorrect p-values, incorrect confidence intervals)

Weak instruments – formalization

- Use different asymptotic approximation to better reflect barely relevant instruments:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$
$$x_i = \pi_0 + \pi_{1,n} z_i + u_i$$

  where $\pi_{1,n} = \tilde{\pi}_1/\sqrt{n}$
- Implications:
  - $\hat{\beta}^{2SLS}$ is biased in the same direction as $\hat{\beta}^{OLS}$
  - $\hat{\beta}^{2SLS}$ not asymptotically normal with mean 0 and usual variance
  - Usual p-values and confidence regions for $\hat{\beta}^{2SLS}$ invalid

Weak instruments – solutions

- Possible solution(s):
  - Use an estimator with better bias properties under weak identification – e.g. LIML
  - Conduct inference using an identification robust test – Anderson-Rubin (AR), Conditional Likelihood Ratio (CLR), or Kleibergen's (each valid whether instrument is strong, weak, or completely irrelevant)
- Practical advice:
  - Always report first stage F statistic for significance of coefficients on instruments – rule of thumb: $F \geq 10$ is okay, $F < 10$ need to worry about weak instruments
    - Justification: under weak instrument asymptotics, bias of 2SLS and is < 10% when $F \geq 10$.
  - If concerned about weak instruments:
    - Report reduced form estimate (regression of y on z)
    - Use alternative estimator (LIML, k-Class)
    - Use identification robust hypothesis tests and confidence intervals

Section 3

Heterogeneous treatment effects

Heterogeneous treatment effects

- Typical linear model assumes x has same effect on y for everyone
$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$
- What if effect of x_i on y_i varies across people, e.g.
$$y_i = \beta_0 + \beta_i x_i + \epsilon_i$$
- To simplify assume a binary treatment
$$x_i = \begin{cases} 1 & \text{if treated} \\ 0 & \text{if not treated} \end{cases}$$
e.g. whether graduated from university or not

Heterogeneous treatment effects

- Typical linear model assumes x has same effect on y for everyone

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

- What if effect of x_i on y_i varies across people, e.g.

$$y_i = \beta_0 + \beta_i x_i + \epsilon_i$$

- To simplify assume a binary treatment

$$x_i = \begin{cases} 1 & \text{if treated} \\ 0 & \text{if not treated} \end{cases}$$

  e.g. whether graduated from university or not

- If x exogenous, then $\mathrm{E}[\hat{\beta}^{\mathrm{OLS}}] = \mathrm{E}[\beta_i]$

Heterogeneous treatment effects

- If x endogenous, assume we have a binary instrument z_i, e.g. randomly assign some people to zero tuition
- What will IV estimate?

Heterogeneous treatment effects

- If x endogenous, assume we have a binary instrument z_i, e.g. randomly assign some people to zero tuition
- What will IV estimate?
  - Show

$$\hat{\beta}^{IV} = \frac{\widehat{\text{Cov}}(y,z)}{\widehat{\text{Cov}}(x,z)} = \frac{\hat{\text{E}}[y|z=1] - \hat{\text{E}}[y|z=0]}{\hat{\text{E}}[x|z=1] - \hat{\text{E}}[x|z=0]}$$

Heterogeneous treatment effects

- If x endogenous, assume we have a binary instrument z_i, e.g. randomly assign some people to zero tuition
- What will IV estimate?
  - Show

$$\hat{\beta}^{IV} = \frac{\widehat{\text{Cov}}(y,z)}{\widehat{\text{Cov}}(x,z)} = \frac{\hat{\text{E}}[y|z=1] - \hat{\text{E}}[y|z=0]}{\hat{\text{E}}[x|z=1] - \hat{\text{E}}[x|z=0]}$$

  - Potential outcomes for x, x_i^0 = value of x_i if z_i had been 0, x_i^1 = value of x_i if z_i had been 1

Heterogeneous treatment effects

- If x endogenous, assume we have a binary instrument z_i, e.g. randomly assign some people to zero tuition
- What will IV estimate?
  - Show

$$\hat{\beta}^{IV} = \frac{\widehat{\mathrm{Cov}}(y,z)}{\widehat{\mathrm{Cov}}(x,z)} = \frac{\hat{\mathrm{E}}[y|z=1] - \hat{\mathrm{E}}[y|z=0]}{\hat{\mathrm{E}}[x|z=1] - \hat{\mathrm{E}}[x|z=0]}$$

  - Potential outcomes for x, $x_i^0 =$ value of x_i if z_i had been 0, $x_i^1 =$ value of x_i if z_i had been 1
  - Show

$$\operatorname{plim} \hat{\beta}^{IV} = \frac{\mathrm{E}[\beta_i | x_i^1 = 1, x_i^0 = 0]\mathrm{P}(x_i^1 = 1, x_i^0 = 0) - \mathrm{E}[\beta_i | x_i^1 = 0, x_i^1 = 0]\mathrm{P}(x_i^1 = 0, x_i^0 =}{\mathrm{P}(x_i^1 = 1, x_i^0 = 0) - \mathrm{P}(x_i^1 = 0, x_i^0 = 1)}$$

Heterogeneous treatment effects

- If x endogenous, assume we have a binary instrument z_i, e.g. randomly assign some people to zero tuition
- What will IV estimate?
  - Show

$$\hat{\beta}^{IV} = \frac{\widehat{\mathrm{Cov}}(y,z)}{\widehat{\mathrm{Cov}}(x,z)} = \frac{\hat{\mathrm{E}}[y|z=1] - \hat{\mathrm{E}}[y|z=0]}{\hat{\mathrm{E}}[x|z=1] - \hat{\mathrm{E}}[x|z=0]}$$

  - Potential outcomes for x, x_i^0 = value of x_i if z_i had been 0, x_i^1 = value of x_i if z_i had been 1
  - Show

$$\operatorname{plim} \hat{\beta}^{IV} = \frac{\mathrm{E}[\beta_i | x_i^1 = 1, x_i^0 = 0]\mathrm{P}(x_i^1 = 1, x_i^0 = 0) - \mathrm{E}[\beta_i | x_i^1 = 0, x_i^1 = 0]\mathrm{P}(x_i^1 = 0, x_i^0 =}{\mathrm{P}(x_i^1 = 1, x_i^0 = 0) - \mathrm{P}(x_i^1 = 0, x_i^0 = 1)}$$

  - Assume no "non-compliers" : $\mathrm{P}(x_i^1 = 0, x_i^0 = 1) = 0$
  - Conclude

$$\operatorname{plim} \hat{\beta}^{IV} = \mathrm{E}[\beta_i | x_i^1 = 1, x_i^0 = 0]$$

IV estimates average effect of x among people for whom the instrument changed x

Section 4

Simultaneous Equations

Simultaneous Equations

- Economic quantities often determined by equilibrium relationships
    - Demand and supply
    - Competitive equilibrium
    - Nash equilibrium
- Equilibrium outcome satisfies a system of equations
- References: Wooldridge (2013) chapter 15, Angrist, Graddy, and Imbens (2000)

Example: demand & supply

- Supply: $q^s = \beta_0^s + \beta_1^s p^s + \epsilon^s$
- Demand: $q^d = \beta_0^d + \beta_1^d p^d + \epsilon^d$
- Equilibrium: $q^s = q^d = q$ and $p^s = p^d = p$
- Inverse demand: $p^d = \frac{-\beta_0^d}{\beta_1^d} + \frac{1}{\beta_1^d} q^d + \frac{-\epsilon^d}{\beta_1^d}$
- Observed p and q satisfy:

$$
\begin{aligned}
q =& \beta_0^s + \beta_1^s p + \epsilon^s \\
p =& \frac{-\beta_0^d}{\beta_1^d} + \frac{1}{\beta_1^d} q + \frac{-\epsilon^d}{\beta_1^d}
\end{aligned}
$$

Generic setup

- Two equation system:

$$y_{i1} = \beta_{01} + \alpha_{21} y_{i2} + \beta_{11} z_{i1} + \beta_{21} z_{i2} + \epsilon_{i1}$$
$$y_{i2} = \beta_{02} + \alpha_{12} y_{i1} + \beta_{12} z_{i1} + \beta_{22} z_{i2} + \epsilon_{i2}$$

- Assume z_{i1} and z_{i2} are exogenous – $\mathrm{E}[z_{i1}\epsilon_{i1}] = 0$, $\mathrm{E}[z_{i1}\epsilon_{i2} = 0]$, $\mathrm{E}[z_{i2}\epsilon_{i1}] = 0$, $\mathrm{E}[z_{i2}\epsilon_{i2} = 0]$
- y_{i1} and y_{i2} are endogenous

Identification

- Identification: given the distribution of observable variables (y and z), are there unique values of the parameters (β and α) that satisfy the model?
  - In OLS and 2SLS the rank condition ensures identification
- Reduced form: solve for y

$$y_{i1} = \underbrace{\frac{\beta_{01} + \alpha_{21}\beta_{02}}{1 - \alpha_{21}\alpha_{12}}}_{\pi_{01}} + \underbrace{\frac{\beta_{11} + \alpha_{21}\beta_{12}}{1 - \alpha_{21}\alpha_{12}}}_{\pi_{11}} z_{i1} + \underbrace{\frac{\beta_{21} + \alpha_{21}\beta_{22}}{1 - \alpha_{21}\alpha_{12}}}_{\pi_{21}} z_{i2} + \underbrace{\frac{\epsilon_{i1} + \alpha_{21}\epsilon_{i2}}{1 - \alpha_{21}\alpha_{12}}}_{u_{i1}}$$

$$y_{i2} = \underbrace{\frac{\beta_{02} + \alpha_{12}\beta_{01}}{1 - \alpha_{12}\alpha_{21}}}_{\pi_{02}} + \underbrace{\frac{\beta_{12} + \alpha_{12}\beta_{21}}{1 - \alpha_{12}\alpha_{21}}}_{\pi_{12}} z_{i1} + \underbrace{\frac{\beta_{22} + \alpha_{12}\beta_{21}}{1 - \alpha_{12}\alpha_{21}}}_{\pi_{22}} z_{i2} + \underbrace{\frac{\epsilon_{i2} + \alpha_{12}\epsilon_{i1}}{1 - \alpha_{12}\alpha_{21}}}_{u_{i2}}$$

- No identification without some restriction

Identification

- Equation 1 identified if there is an exogenous variable excluded from equation 1 and included in equation 2
  - E.g. $\beta_{21} = 0$ and $\beta_{22} \neq 0$.
  - Intuition: can use z_{i2} as instrument for y_{i2} and estimate equation 1 by 2SLS
- Equation 2 identified if there is an exogenous variable excluded from equation 2 and included in equation 1
  - E.g. $\beta_{12} = 0$ and $\beta_{11} \neq 0$.
  - Intuition: can use z_{i1} as instrument for y_{i1} and estimate equation 2 by 2SLS

Section 5

Limited Dependent Variables

Limited dependent variables

- Limited dependent variables = dependent variable has limited range
- Examples:
  - Binary $y_i = 0$ or 1, such as employment, program participation, etc
  - Multinomial y_i one of a finite number of categories, such as which type of car someone buys
  - Censoring and truncation $y_i \geq 0$, such as hours worked

Binary dependent variables

- Observed: $y_i = 0$ or 1
- Model:

$$y_i^* = \beta_0 + \beta_1 x_i + \epsilon_i$$
$$y_i = \begin{cases} 1 & \text{if } y_i^* \geq 0 \\ 0 & \text{if } y_i^* < 0 \end{cases}$$

- Assume $\epsilon \perp\!\!\!\perp x$ and ϵ has CDF F_ϵ

$$\begin{aligned} \mathrm{P}(y = 1|x) =& \mathrm{P}(\beta_0 + \beta_1 x + \epsilon \geq 0|x) \\ =& \mathrm{P}(\epsilon \geq -(\beta_0 + \beta_1 x)|x) \\ =& 1 - F_\epsilon(-\beta_0 - \beta_1 x) \end{aligned}$$

- β_0 and β_1 estimated by maximum likelihood:

$$(\hat\beta_0, \hat\beta_1) = \arg\max_{b_0, b_1} \sum_{i=1}^N \left[\log F_\epsilon(-b_0 - b_1 x_i)(1 - y_i) + \log\left(1 - F_\epsilon(-b_0 - b_1 x_i)\right) y_i\right]$$

- When $\epsilon \sim N(0, 1)$, called a probit
- When $F_\epsilon(z) = \frac{e^z}{1+e^z}$, called a logit

Selection

- Continuous outcome y_i observed only if a binary outcome, $d_i = 1$
  - E.g. $y_i = \log$ wage and $d_i = 1$ if employed, 0 if unemployed
- Model:

$$d_i = \begin{cases} 1 & \text{if } \gamma_0 + \gamma_1 x_i + \gamma_2 z_i + u_i \geq 0 \\ 0 & \text{if } \gamma_0 + \gamma_1 x_i + \gamma_2 z_i + u_i < 0 \end{cases}$$

$$y_i = \begin{cases} \beta_0 + \beta_1 x_i + \epsilon_i & \text{if } d_i = 1 \\ \text{not observed} & \text{if } d_i = 0 \end{cases}$$

  - Assume $\epsilon, u \perp\!\!\!\perp x, z$
  - ϵ and u might be correlated

Selection

- OLS using observed y is inconsistent

$$\begin{aligned}\operatorname{plim} \hat{\beta}_1^{\text{OLS}} =& \frac{\text{Cov}(x, y|d=1)}{\text{Var}(x|d=1)} \\ =& \beta_1 + \frac{\text{Cov}(x, \epsilon|d=1)}{\text{Var}(x|d=1)} \\ =& \beta_1 + \frac{\text{E}[x\epsilon|u_i \geq -(\gamma_0 + \gamma_1 x_i + \gamma_2 z_i)]}{\text{Var}(x|d=1)} \neq \beta_1\end{aligned}$$

- Assume ϵ, u jointly normally distributed with correlation ρ, then

$$\text{E}[\epsilon|u_i \geq -(\gamma_0 + \gamma_1 x_i + \gamma_2 z_i)] = \rho\sigma_\epsilon \frac{\phi(-(\gamma_0 + \gamma_1 x_i + \gamma_2 z_i))}{\Phi(-(\gamma_0 + \gamma_1 x_i + \gamma_2 z_i))}$$

- Heckman (1979) two-stage estimator:
  1. Estimate $\hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2$ using probit

Selection

2. Estimate $\hat{\beta}_0, \hat{\beta}_1, \widehat{\rho\sigma_\epsilon}$ by OLS regression of y on x and $\frac{\phi(-(\hat{\gamma}_0+\hat{\gamma}_1 x_i+\hat{\gamma}_2 z_i))}{\Phi(-(\hat{\gamma}_0+\hat{\gamma}_1 x_i+\hat{\gamma}_2 z_i))}$

- For reliable results important that a variable, like z_i, is included in selection equation and excluded from outcome equation

Section 6

Nonparametric models

Section 7

Structural models

References

Angrist, Joshua D., Kathryn Graddy, and Guido W. Imbens. 2000. "The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish." *The Review of Economic Studies* 67 (3):pp. 499–527. URL http://www.jstor.org/stable/2566964.

Angrist, Joshua D and Jörn-Steffen Pischke. 2014. *Mastering 'Metrics: The Path from Cause to Effect*. Princeton University Press.

Autor, David H., Alan Manning, and Christopher L. Smith. 2016. "The Contribution of the Minimum Wage to US Wage Inequality over Three Decades: A Reassessment." *American Economic Journal: Applied Economics* 8 (1):58–99. URL http://www.aeaweb.org/articles.php?doi=10.1257/app.20140073.

References

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (1):153–161. URL http://www.jstor.org/stable/1912352.

Imbens, Guido and Jeffrey Wooldridge. 2007. ""Weak instruments and many instruments"." *What's New in Econometrics?* URL http://www.nber.org/minicourse3.html.

Levitt, Steven D. 1997. "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime." *The American Economic Review* 87 (3):pp. 270–290. URL http://www.jstor.org/stable/2951346.

Wooldridge, J.M. 2013. *Introductory econometrics: A modern approach*. South-Western.