

# 710a Metrics Notes

Travis Cao

## Stats

### Geometric series sum

$$\sum_{j=0}^n r^j = \frac{1-r^{n+1}}{1-r}$$

### Conditional expectation

$$E[Y] = \sum_l E[Y|Z=l]Pr(Z=l)$$

### Block inversion formula

$M$  is invertible iff  $A - BD^{-1}C$  is invertible, and  $M^{-1} =$

$$\begin{bmatrix} (A - BD^{-1}C)^{-1} & -(A - BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1} \end{bmatrix}$$

### Sherman-Morrison formula

$A + uv'$  is invertible iff  $1 + v'A^{-1}u \neq 0$ , and

$$(A + uv')^{-1} = A^{-1} - \frac{A^{-1}uv'A^{-1}}{1 + v'A^{-1}u}$$

### Partial and average partial effect

- Partial effect (at  $X = x$ ):  $\frac{\partial}{\partial x} E[Y|X = x]$
- Average partial effect:  $E[\frac{\partial}{\partial X} E[Y|X]]$

### Inequalities

- Chebyshev's:  $Pr(|X - E[X]| \geq \epsilon) \leq \frac{Var(X)}{\epsilon^2}$  for all  $\epsilon > 0$
- Markov:  $Pr(X \geq \epsilon) \leq \frac{E[X]}{\epsilon}$  for all  $\epsilon > 0$
- Jensen's: For convex function  $g(\cdot)$ ,  $g(E[X]) \leq E[g(X)]$
- Cauchy Schwarz:  $E[XY]^2 \leq E[X^2]E[Y^2]$

### Martingale CLT

If  $\{Z_t\}_{t=1}^T$  satisfy

- $\{Z_t\}_{t=1}^T$  is strictly stationary
- $E[Z_1^2] < \infty$
- $E[Z_t|Z_{t-1}, Z_{t-2}, \dots, Z_1] = 0$
- $\frac{1}{T} \sum_{t=1}^T Z_t^2 \xrightarrow{p} E[Z_1^2]$

then  $\frac{1}{\sqrt{T}} \sum_{t=1}^T Z_t \xrightarrow{d} N(0, E[Z_1^2])$  as  $T \rightarrow \infty$

## IV

Model:  $Y_i = X_i' \beta + U_i$

Assumptions:

- Validity of instrument
  - Exogeneity:  $E[U|Z] = 0$
  - Relevance:  $E[ZX']$  and  $\sum_{i=1}^n Z_i X_i'$  are invertible
- $E[|Y|^2 + ||X||^2 + ||Z||^2] < \infty$

- $\{(Y_i, X_i', Z_i')\}$  are i.i.d.

IV estimator:  $\hat{\beta}_{IV} = (\sum_{i=1}^n Z_i X_i')^{-1} (\sum_{i=1}^n Z_i Y_i')$

### Small sample property

Model:  $Y_i = \beta_0 + X_i \beta_1 + U_i$

$$X_i = \pi_0 + Z_i \pi_1 + V_i$$

$$\begin{pmatrix} U_i \\ V_i \end{pmatrix} \Bigg| Z_i \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{pmatrix} \right)$$

$$\Rightarrow E[\hat{\beta}_{IV} | \mathbb{X}, \mathbb{Z}] = \beta + \frac{\sigma_{uv}}{\sigma_v^2} \frac{\pi_1}{se(\hat{\pi}_1)} \frac{v}{+v}, \text{ where } v \sim N(0, 1)$$

### Large sample property

Large sample approximation is reasonable depends on both

- sample size  $n$  (the bigger, the better), and
- the covariance between the instrument and endogenous regressor  $\pi_1$  (the bigger, the better)

Which means there's a need of testing whether the instrument is relevant (whether  $\pi_1 \neq 0$ )

Model:  $X_1 = Z_1 \pi_1 + X_2' \pi_2 + V$ . Let  $\hat{\pi}_1$  be OLS estimate of  $\pi_1$ .

Then the F-statistic is  $F = \frac{(\hat{\pi}_1 - 0)^2}{se(\hat{\pi}_1)^2} = \frac{\hat{\pi}_1^2}{se(\hat{\pi}_1)^2}$

Empirically, if  $F \geq 10$ , then some rough insurance with nominal 95% confidence intervals have actual coverage of at least 80%.

### Weak instrument

Model:  $Y = X_1 \beta_1 + U$ . Weak instrument is  $Z_1$ .

Assumptions on  $Z_1$ :

- Exogeneity:  $E[U|Z_1] = 0$
- Allow for relevance
- $E[Z_1^2]$  and  $\frac{1}{n} \sum_i Z_{1i}^2$  are both non-zero (i.e. invertible)

Under weak instrument, want to test

$$H_0 : \beta_1 = c \quad H_1 : \beta_1 \neq c$$

Identification under  $H_0$ :  $E[Z_1(Y - X_1 c)] = 0$

$\Rightarrow T = \frac{1}{n} \sum_i Z_{1i} (Y_i - X_{1i} c)$ . Reject  $H_0$  when  $|T|$  is large

Anderson-Rubin Test:

$$AR := \frac{\sqrt{n}T}{S} \xrightarrow{d} N(0, 1)$$

where  $S^2 = \frac{1}{n} \sum_i Z_{1i}^2 (Y_i - X_{1i} c)^2 \xrightarrow{p} E[Z_1^2 U^2]$

With a size of 5%, reject  $H_0$  when  $|AR| > 1.96$

Confidence set:  $\{c \in \mathbb{R} : |AR(c)| \leq 1.96\}$

### Optimal instrument

Say optimal instrument is  $h^*(Z)$ . Identification yields from

$E[h(Z)U] = 0$ . So  $\hat{\beta}_{IV}^h = (\frac{1}{n} \sum_i h(Z_i) X_i')^{-1} \frac{1}{n} \sum_i h(Z_i) Y_i$

Under homoskedasticity, asymptotic variance (AVAR) is

$$\begin{aligned} \Omega^h &= \frac{E[h(Z)^2]}{E[h(Z)X]^2} \sigma_u^2 = \frac{E[h(Z)^2]}{E[h(Z)E[X|Z]]^2} \sigma_u^2 \\ &\geq \frac{E[h(Z)^2]}{E[h(Z)]^2 E[E[X|Z]^2]} \sigma_u^2 = \frac{\sigma_u^2}{E[E[X|Z]^2]} \end{aligned}$$

To achieve the lower bound of AVAR,  $h^*(Z) = E[X|Z]$

## Random coefficient model

Model:  $Y = X' \beta_0 \cdot U = X' \beta_0 + X' \beta_0 \cdot (U - 1)$

Assumptions:

- $E[U|Z] = 1$ , and  $E[ZX']$  invertible
- $Z$  independent of  $Y$
- Let  $X = (1, X)'$ ,  $Z = (1, Z)'$ ,  $Z \in \{0, 1\}$ ,  $X \in \{0, 1\}$

$$\hat{\beta}_{IV} \xrightarrow{p} \frac{Cov(Z, Y)}{Cov(Z, X)} = \frac{E[Y|Z=1] - E[Y|Z=0]}{E[X|Z=1] - E[X|Z=0]}$$

For binary  $X$ , its response to  $Z$  given unobservable  $U$  is  $X_U(Z)$

	$X_U(0) = 0$	$X_U(0) = 1$
$X_U(1) = 0$	Never taking	Defying
$X_U(1) = 1$	Complying	Always taking

Assume no defyers, and  $Pr(\text{Complying}) > 0$ , then

$$\begin{aligned} E[X|Z=1] - E[X|Z=0] &= E[X_U(1) - X_U(0)] \\ &= Pr(X_U(1) - X_U(0) = 1) - Pr(X_U(1) - X_U(0) = -1) \\ &= Pr(\text{Complying}) \\ E[Y|Z=1] - E[Y|Z=0] &= E[Y_U(1)X_U(1) + Y_U(0)(1 - X_U(1))] \\ &\quad - E[Y_U(1)X_U(0) + Y_U(0)(1 - X_U(0))] \\ &\quad \text{(since } Y = Y_U(1)X + Y_U(0)(1 - X)) \\ &= E[(Y_U(1) - Y_U(0))(X_U(1) - X_U(0))] \\ &= E[(Y_U(1) - Y_U(0)) | X_U(1) - X_U(0) = 1] \times Pr(\text{Complying}) \end{aligned}$$

$$\text{Thus, } \hat{\beta}_{IV} \xrightarrow{p} E[ \underbrace{(Y_U(1) - Y_U(0))}_{\text{mean response of } Y \text{ to } X} \mid \underbrace{X_U(1) - X_U(0) = 1}_{\text{for compliers}} ]$$

## Time Series

### TS Models

- Static:  $Y_t = \alpha_0 + X_t' \delta_0 + U_t$
- FDL(s):  $Y_t = \alpha_0 + X_t' \delta_0 + \dots + X_{t-s}' \delta_s + U_t$
- AR(p):  $Y_t = \alpha_0 + Y_{t-1} \rho_1 + \dots + Y_{t-p} \rho_p + U_t$
- MA(q):  $Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$
- Trend:
  - Linear time trend:  $Y_t = \beta_0 + \beta_1 t + U_t$
  - Exponential trend:  $\log(Y_t) = \beta_0 + \beta_1 t + U_t$
- Seasonality:  $Y_t = \alpha_0 + \alpha_1 1_{\{t/12 \text{ is an integer}\}} + U_t$

### Stationarity

- Strict:  $(Y_{t_1}, Y_{t_2}, \dots, Y_{t_k}) \sim (Y_{t_1+l}, Y_{t_2+l}, \dots, Y_{t_k+l})$  (Joint distribution is  $t$  independent)
- Weak (Covariance): For all  $t$ ,  $E[Y_t]$  and  $\gamma(k) = Cov(Y_t, Y_{t+k})$  ( $\leftarrow$  autocovariance function) are both independent of  $t$