HE3021 Tutorial 8 Attempt

Lui Yu Sen U1930037F

3/30/2021

```
warningsoff <- getOption("warn")
options(warn = -1)
library(haven) # read dta files
library(tidyverse) # manipulate dataframes
library(ivreg) # 2 stage linear regression and its associated statistical tests
library(aod) # for chi2 and F tests

filpaths <- list.files(path = paste0(getwd(),"//rawdata//"),
                       full.names = TRUE)
wage2 <- read_dta(file = filpaths[1])
options(warn = warningsoff)
```

1

a

$$H_0 : \pi_{sibs} = \pi_{brthord} = \pi_{meduc} = 0, H_1 : otherwise, \alpha = 0.05, LM - test$$

```
instr_exo_test_model <- lm(
    formula = educ ~ sibs + brthord + meduc + married + exper + tenure, data = wage2)
instr_exo_test <- wald.test(
    Sigma = vcov(instr_exo_test_model), b = coefficients(instr_exo_test_model), Terms = c(2,3,4))
# p-value below printed value
# instrumental relevance true
```

p-value == $0 < 0.05$, thus we have sufficient evidence to reject H_0.

b

```
twostage_model <- ivreg(
    formula = lwage ~ married + exper + tenure + educ | married + exper + tenure
    + sibs + brthord + meduc, data = wage2)
twostage_model_summary <-(summary(twostage_model, diagnostics = TRUE))
summary(twostage_model, diagnostics = FALSE)
```

```
##
## Call:
## ivreg(formula = lwage ~ married + exper + tenure + educ | married +
##     exper + tenure + sibs + brthord + meduc, data = wage2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.72191 -0.23654  0.01206  0.25728  1.33428
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 4.079660   0.351142  11.618  < 2e-16 ***
## married     0.205496   0.048986   4.195 3.04e-05 ***
## exper       0.035371   0.006378   5.546 4.01e-08 ***
## tenure      0.006522   0.003090   2.111   0.0351 *
## educ        0.153649   0.021205   7.246 1.04e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4 on 779 degrees of freedom
## Multiple R-Squared: 0.05733, Adjusted R-squared: 0.05249
## Wald test: 22.26 on 4 and 779 DF,  p-value: < 2.2e-16
```

c

Regressing the residuals on each instrumental variable, we have:

$$\hat{u}_{IV} = \gamma_0 + \gamma_1 married + \gamma_2 exper + \gamma_3 tenure + \gamma_4 sibs + \gamma_5 brthord + \gamma_6 meduc$$

We test: $H_0 : \gamma_4 = \gamma_5 = \gamma_6 = 0, H_1 : otherwise, \alpha = 0.05, F - test$. The J-statistic is: 0.600554. From the χ^2-table, we have the critical value as $5.991 > 0.600554$. The p-value is $0.7406131 > 0.05$, so we have insufficient evidence to reject the null hypothesis. We assume that at least 1 of the instrument variables is exogenous, then not rejecting H_0 means that all instrumental variables are exogenous.

d

From the Sargan test, it is likely that the 2SLS model's instrumental variables fulfill instrumental exogeneity. Then, we test $H_0 : \beta_{educ,2SLS}\ and\ \beta_{educ,OLS}\ are\ consistent, Var(\beta_{educ,2SLS}) > Var(\beta_{educ,OLS}), H_1 : only\ \beta_{educ,2SLS}\ is\ consistent, \alpha = 0.05, \chi^2 - test$. If the coefficients estimated in OLS are significantly different from those in 2-stage regression, then it is likely that the OLS coefficient is inconsistent and that $Cov(x, u) \neq 0$. The $\chi^2 - statistic$ is $16.6912726 > 5.991$ and the p-value is $4.8543614 \times 10^{-5} < 0.05$. Thus we have sufficient evidence to reject H_0. It is likely that $Cov(x, u) \neq 0$.

2

a

$$\hat{\beta}_1 = \frac{c\hat{o}v(x,y)}{V\hat{a}r(x)} \tag{1}$$

$$= \frac{C\hat{o}v(x, \beta_0 + \beta_1 x + u)}{V\hat{a}r(x)} \tag{2}$$

$$= \beta_1 \frac{V\hat{a}r(x)}{V\hat{a}r(x)} + \frac{C\hat{o}v(x,u)}{V\hat{a}r(x)} \tag{3}$$

$$= \beta_1 + \frac{C\hat{o}v(x,u)}{\sqrt{V\hat{a}r(x) \cdot V\hat{a}r(u)}} \cdot \frac{\sqrt{V\hat{a}r(u)}}{\sqrt{V\hat{a}r(x)}} \tag{4}$$

$$\lim_{n\to\infty} \left[\beta_1 + \frac{C\hat{o}v(x,u)}{\sqrt{V\hat{a}r(x) \cdot V\hat{a}r(u)}} \cdot \frac{\sqrt{V\hat{a}r(u)}}{\sqrt{V\hat{a}r(x)}} \right] = \beta_1 + Corr(x,u) \cdot \frac{\sigma_u}{\sigma_x} \tag{5}$$

b

$$\hat{\beta}_{1,IV} = \frac{C\hat{o}v(z,y)}{C\hat{o}v(z,x)} \tag{6}$$

$$= \frac{C\hat{o}v(z, \beta_0, \beta_1 x + u)}{C\hat{o}v(z,x)} \tag{7}$$

$$= \beta_1 \frac{C\hat{o}v(z,x)}{C\hat{o}v(z,x)} + \frac{C\hat{o}v(z,u)}{C\hat{o}v(z,x)} \tag{8}$$

$$= \beta_1 + \frac{C\hat{o}v(z,u)}{C\hat{o}v(z,x)} \cdot \frac{\sqrt{V\hat{a}r(z)}\sqrt{V\hat{a}r(x)}}{\sqrt{V\hat{a}r(z)}\sqrt{V\hat{a}r(u)}} \cdot \frac{\sqrt{V\hat{a}r(z)}\sqrt{V\hat{a}r(u)}}{\sqrt{V\hat{a}r(z)}\sqrt{V\hat{a}r(x)}} \tag{9}$$

$$\lim_{n\to\infty} \left[\beta_1 + \frac{C\hat{o}v(z,u)}{C\hat{o}v(z,x)} \cdot \frac{\sqrt{V\hat{a}r(z)}\sqrt{V\hat{a}r(x)}}{\sqrt{V\hat{a}r(z)}\sqrt{V\hat{a}r(u)}} \cdot \frac{\sqrt{V\hat{a}r(z)}\sqrt{V\hat{a}r(u)}}{\sqrt{V\hat{a}r(z)}\sqrt{V\hat{a}r(x)}} \right] = \beta_1 + \frac{Corr(z,u)}{Corr(z,x)} \cdot \frac{\sigma_u}{\sigma_x} \tag{10}$$

c

$Cov(z,u) \neq 0,\ Cov(z,x)\ small$, so $\frac{Corr(z,u)}{Corr(z,x)} \cdot \frac{\sigma_u}{\sigma_x} \neq 0$, so

$$\lim_{n\to\infty} \left[\beta_1 + \frac{C\hat{o}v(x,u)}{\sqrt{V\hat{a}r(x) \cdot V\hat{a}r(u)}} \cdot \frac{\sqrt{V\hat{a}r(u)}}{\sqrt{V\hat{a}r(u)}} \right] = \beta_1 + Corr(x,u) \cdot \frac{\sigma_u}{\sigma_x} \neq \beta_1$$

and the estimator is inconsistent.
Neither estimator would work, since $Cov(x,u) \neq 0,\ Cov(z,u) \neq 0$; they are both endogenous.