Chris Kalra

Econ 610

Homework 10

1a) 

Here, the return on siblings is negative (and statistically significant), which shows an effect in the opposite direction than what the IV regression would show; a huge problem!

1b) *educ* and *brthord* may be negatively correlated because parents may run out of money after sending their first child or children to college, as a college education may very well cost more than the amount of disposable income parents bring in. Likewise, when it is time for a younger child to go to college, the parents may have retired or may be closer to retirement, thus having less money to spend on their children’s education

1c) 

*brthord* is indeed statistically significant with a p-value less than .01, and its effect is negative; being born one place later in birth order results in .28 fewer years of education

1d) We need π2 ≠ 0. We can do this by running the regression (regressing *educ* on *sibs* and *brthord*), in which we obtain a p-value for π2 = .0076 and a t-value of -2.675, so we reject the null hypothesis, which stated π2 = 0, and we thus conclude that π2 ≠ 0

1e) 

The standard error for  is much larger here than it was in part (c), increasing from .032 to .075. Because of this, it is also no longer statistically significant at the .05 significance level. Likewise, the standard error for  is fairly large, so it has a t-value of only 0.112 and a p-value of 0.903, so it is incredibly statistically insignificant

1f) The correlation between *edûc* and *sibs* is extremely low (r = -0.93), so we run into the problem is collinearity, which may result in  being fairly inaccurate

2a) 

Each additional year of education results in 0.0906 fewer children on average. Therefore, if 100 women receive another year of education, they are predicted to have 9 fewer children in total

2b) 

Most importantly, we need to make sure the coefficient on *frsthalf* is statistically significant. With a t-value of -7.55 and a p-value less than .01, it is statistically significant, meaning *frsthalf* is a reasonable IV candidate for *educ*

2c) 

 is still statistically significant, although now it is nearly twice the size as it was in (a), and it remains positive

3a) Texas had the most executions (34 in 1993). In total, 16 states had executed at least 1 prisoner during between 1991 and 1993

3b) 

The coefficient on *exec* is not statistically significant, meaning execution does not appear to have a deterrent effect

3c) 

Here, the negative coefficient on change in executions is statistically significant with a t-value of -2.39 and a p-value of 0.02073, implying that there is a deterrent effect. According to the coefficient estimate, 10 additional executions is estimated to reduce the murder rate by 1.04, which is just over one fewer murder per 100,00 people (murders tend to be reported at a rate of per 100,000)

3d) 

There appears to be a negative correlation between *cexec* and *cexec\_1* , as shown by *cexec\_1* having a negative coefficient. The coefficient can be interpreted as: for each year, the following year has nearly a one-to-one decrease in executions

3e) 

The standard error on the coefficient of *cexec* is much larger, making it no longer statistically significant, as its p-value increased from 0.02073 to 0.1262

4a) 

All independent variables are statistically significant at the .05 level, and all besides *exper* have p-values below .01

4b) 

(Note that in the output section, *educ* was called *pred* in order to avoid rewriting the original *educ* variable)

The coefficient estimates are identical to what was returned in (a), but the standard errors are slightly different. However, the significance results are the same here as they were in (a)

*Code (italics)* and Output (Monaco font)

1a) *wage=data*

*regr1a=lm(lwage~sibs, data=wage)*

*summary(regr1a)*

Call:

lm(formula = lwage ~ sibs, data = wage)

Residuals:

Min 1Q Median 3Q Max

-1.97662 -0.25857 0.02503 0.28572 1.22677

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 6.861076 0.022078 310.771 < 2e-16 \*\*\*

sibs -0.027904 0.005908 -4.723 2.68e-06 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4164 on 933 degrees of freedom

Multiple R-squared: 0.02335, Adjusted R-squared: 0.0223

F-statistic: 22.31 on 1 and 933 DF, p-value: 2.68e-06

1b) *regr1b=lm(educ~brthord, data=wage)*

*summary(regr1b)*

Call:

lm(formula = educ ~ brthord, data = wage)

Residuals:

Min 1Q Median 3Q Max

-4.8668 -1.5842 -0.7362 2.1332 6.1117

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 14.14945 0.12868 109.962 < 2e-16 \*\*\*

brthord -0.28264 0.04629 -6.106 1.55e-09 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.155 on 850 degrees of freedom

(83 observations deleted due to missingness)

Multiple R-squared: 0.04202, Adjusted R-squared: 0.04089

F-statistic: 37.29 on 1 and 850 DF, p-value: 1.551e-09

1c) *install.packages("AER")*

*library('AER')*

*regr1c=ivreg(lwage ~ educ | brthord, data=wage)*

*summary(regr1c)*

Call:

ivreg(formula = lwage ~ educ | brthord, data = wage)

Residuals:

Min 1Q Median 3Q Max

-1.8532 -0.2557 0.0435 0.2970 1.3033

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 5.03040 0.43295 11.619 < 2e-16 \*\*\*

educ 0.13064 0.03204 4.078 4.97e-05 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4215 on 850 degrees of freedom

Multiple R-Squared: -0.02862, Adjusted R-squared: -0.02983

Wald test: 16.63 on 1 and 850 DF, p-value: 4.975e-05

1d) *regr1d=lm(educ~sibs+brthord, data=wage)*

*summary(regr1d)*

Call:

lm(formula = educ ~ sibs + brthord, data = wage)

Residuals:

Min 1Q Median 3Q Max

-5.1438 -1.6854 -0.6852 2.0090 5.9950

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 14.29650 0.13329 107.260 < 2e-16 \*\*\*

sibs -0.15287 0.03987 -3.834 0.000135 \*\*\*

brthord -0.15267 0.05708 -2.675 0.007619 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.137 on 849 degrees of freedom

(83 observations deleted due to missingness)

Multiple R-squared: 0.05833, Adjusted R-squared: 0.05611

F-statistic: 26.29 on 2 and 849 DF, p-value: 8.33e-12

1e) *regr1e=ivreg(lwage~educ+sibs|brthord + sibs, data=wage)*

*summary(regr1e)*

Call:

ivreg(formula = lwage ~ educ + sibs | brthord + sibs, data = wage)

Residuals:

Min 1Q Median 3Q Max

-1.84808 -0.26227 0.03841 0.29901 1.30836

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 4.938527 1.055690 4.678 3.37e-06 \*\*\*

educ 0.136994 0.074681 1.834 0.0669 .

sibs 0.002111 0.017372 0.122 0.9033

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.427 on 849 degrees of freedom

Multiple R-Squared: -0.05428, Adjusted R-squared: -0.05676

Wald test: 10.9 on 2 and 849 DF, p-value: 2.124e-05

1f) *fit=fitted.values(regr1d)*

*cor(wage$sibs, fit, use="pairwise.complete.obs")*

[1] -0.9302391

2a) fertility=data

regr2a=lm(children~educ + age + I(age^2), data=fertility)

summary(regr2a)

Call:

lm(formula = children ~ educ + age + I(age^2), data = fertility)

Residuals:

Min 1Q Median 3Q Max

-5.8351 -0.7135 -0.0054 0.7141 7.5055

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -4.1383066 0.2405942 -17.200 <2e-16 \*\*\*

educ -0.0905755 0.0059207 -15.298 <2e-16 \*\*\*

age 0.3324486 0.0165495 20.088 <2e-16 \*\*\*

I(age^2) -0.0026308 0.0002726 -9.651 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.46 on 4357 degrees of freedom

Multiple R-squared: 0.5687, Adjusted R-squared: 0.5684

F-statistic: 1915 on 3 and 4357 DF, p-value: < 2.2e-16

2b) regr2b=lm(educ~age+I(age^2)+frsthalf, data=fertility)

summary(regr2b)

Call:

lm(formula = educ ~ age + I(age^2) + frsthalf, data = fertility)

Residuals:

Min 1Q Median 3Q Max

-7.9599 -2.4941 0.2663 2.2663 14.9934

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 9.6928643 0.5980686 16.207 < 2e-16 \*\*\*

age -0.1079504 0.0420402 -2.568 0.0103 \*

I(age^2) -0.0005056 0.0006929 -0.730 0.4657

frsthalf -0.8522854 0.1128296 -7.554 5.12e-14 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.711 on 4357 degrees of freedom

Multiple R-squared: 0.1077, Adjusted R-squared: 0.107

F-statistic: 175.2 on 3 and 4357 DF, p-value: < 2.2e-16

2c) regr2c=ivreg(children~educ+age+I(age^2) | frsthalf + age + I(age^2), data=fertility)

summary(regr2c)

Call:

ivreg(formula = children ~ educ + age + I(age^2) | frsthalf +

age + I(age^2), data = fertility)

Residuals:

Min 1Q Median 3Q Max

-6.05272 -0.71481 0.06224 0.76236 7.23693

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -3.3878054 0.5481502 -6.180 6.98e-10 \*\*\*

educ -0.1714989 0.0531796 -3.225 0.00127 \*\*

age 0.3236052 0.0178596 18.119 < 2e-16 \*\*\*

I(age^2) -0.0026723 0.0002797 -9.555 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.491 on 4357 degrees of freedom

Multiple R-Squared: 0.5502, Adjusted R-squared: 0.5499

Wald test: 1765 on 3 and 4357 DF, p-value: < 2.2e-16

3a) *which.max(murder$exec)*

[1] 132

3b) *library(plm)*

*regr3b = plm(mrdrte~d93 + exec + unem, data=murder, subset=d90+d93==1, model='pooling', index=c("id", "year"))*

*summary(regr3b)*

Pooling Model

Call:

plm(formula = mrdrte ~ d93 + exec + unem, data = murder, subset = d90 +

d93 == 1, model = "pooling", index = c("id", "year"))

Balanced Panel: n = 51, T = 2, N = 102

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-13.0666 -3.3556 -1.6472 1.6071 66.3873

Coefficients:

Estimate Std. Error t-value Pr(>|t|)

(Intercept) -5.27800 4.42781 -1.1920 0.236134

d93 -2.06742 2.14463 -0.9640 0.337421

exec 0.12773 0.26324 0.4852 0.628599

unem 2.52889 0.78172 3.2350 0.001659 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares: 11401

Residual Sum of Squares: 10243

R-Squared: 0.10161

Adj. R-Squared: 0.07411

F-statistic: 3.69475 on 3 and 98 DF, p-value: 0.0144

3c) *regr3c=plm(mrdrte~d93 + exec + unem, data=murder, subset=d90+d93==1, model='within', index=c("id", "year"))*

*summary(regr3c)*

Oneway (individual) effect Within Model

Call:

plm(formula = mrdrte ~ d93 + exec + unem, data = murder, subset = d90 +

d93 == 1, model = "within", index = c("id", "year"))

Balanced Panel: n = 51, T = 2, N = 102

Residuals:

Min. 1st Qu. Median 3rd Qu. Max.

-1.27948 -0.34897 0.00000 0.34897 1.27948

Coefficients:

Estimate Std. Error t-value Pr(>|t|)

d93 0.413266 0.209385 1.9737 0.05418 .

exec -0.103840 0.043414 -2.3918 0.02073 \*

unem -0.066591 0.158686 -0.4196 0.67662

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Total Sum of Squares: 33.47

Residual Sum of Squares: 27.936

R-Squared: 0.16533

Adj. R-Squared: -0.75627

F-statistic: 3.16936 on 3 and 48 DF, p-value: 0.032614

3d) *regr3d=lm(cexec ~ cexec\_1, data=murder)*

*summary(regr3d)*

Call:

lm(formula = cexec ~ cexec\_1, data = murder)

Residuals:

Min 1Q Median 3Q Max

-7.7618 -0.3499 -0.3499 -0.3499 10.7439

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.3499 0.3698 0.946 0.349

cexec\_1 -1.0824 0.1690 -6.403 5.59e-08 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.62 on 49 degrees of freedom

(102 observations deleted due to missingness)

Multiple R-squared: 0.4555, Adjusted R-squared: 0.4444

F-statistic: 41 on 1 and 49 DF, p-value: 5.586e-08

3e) *regr3e=ivreg(cmrdrte ~ cexec + cunem | cexec\_1 + cunem, data=murder)*

*summary(regr3e)*

Call:

ivreg(formula = cmrdrte ~ cexec + cunem | cexec\_1 + cunem, data = murder)

Residuals:

Min 1Q Median 3Q Max

-2.29762 -0.62775 -0.06348 0.83255 2.56891

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.41097 0.21142 1.944 0.0578 .

cexec -0.10010 0.06432 -1.556 0.1262

cunem -0.06673 0.15871 -0.420 0.6760

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.079 on 48 degrees of freedom

Multiple R-Squared: 0.1096, Adjusted R-squared: 0.07251

Wald test: 1.309 on 2 and 48 DF, p-value: 0.2796

4a) *regr4a=ivreg(lwage ~ educ + exper + tenure + black | sibs + exper + tenure + black, data=wage)*

*summary(regr4a)*

Call:

ivreg(formula = lwage ~ educ + exper + tenure + black | sibs +

exper + tenure + black, data = wage)

Residuals:

Min 1Q Median 3Q Max

-1.8176 -0.2403 0.0139 0.2567 1.3225

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 5.215976 0.543451 9.598 < 2e-16 \*\*\*

educ 0.093632 0.033719 2.777 0.00560 \*\*

exper 0.020922 0.008388 2.494 0.01279 \*

tenure 0.011548 0.002740 4.215 2.74e-05 \*\*\*

black -0.183329 0.050136 -3.657 0.00027 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3848 on 930 degrees of freedom

Multiple R-Squared: 0.1685, Adjusted R-squared: 0.165

Wald test: 24.92 on 4 and 930 DF, p-value: < 2.2e-16

4b) *regr4b\_1=lm(educ~sibs+exper+tenure+black, data=wage)*

*wage$pred=predict.lm(regr4b\_1)*

*regr4b\_2=lm(lwage ~ pred + exper + tenure + black, data=wage)*

*summary(regr4b\_2)*

Call:

lm(formula = lwage ~ pred + exper + tenure + black, data = wage)

Residuals:

Min 1Q Median 3Q Max

-1.97409 -0.25720 0.00997 0.26147 1.25198

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 5.215976 0.568815 9.170 < 2e-16 \*\*\*

pred 0.093632 0.035293 2.653 0.008114 \*\*

exper 0.020922 0.008779 2.383 0.017368 \*

tenure 0.011548 0.002868 4.027 6.1e-05 \*\*\*

black -0.183329 0.052476 -3.494 0.000499 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.4028 on 930 degrees of freedom

Multiple R-squared: 0.08912, Adjusted R-squared: 0.0852

F-statistic: 22.75 on 4 and 930 DF, p-value: < 2.2e-16