Homework 7

Wei Ye*
ECON 7910- Econometrics I

Due on Dec 2, 2021

1 Question 1

Solution:

a. The conditions we need are IVs, and the num of IVs should be at least larger than 3, i.e., the num of IVs is larger than the num of endougenous vars. First, we make first stage regressions via IVs to estimate endougenous vars, repectively. Then by the second stage regression to estimate parameters.

b. From table (1), we assume the var GDPPCInit as z. In the regression table, corjud can affect economic growth, because its coefficient is significantly different from 0. However, we can't reject the coefficient of cored is 0 under 95% confidence level.

c. In my opinion, edext is a vaid IV.

   (a) edext is correlated with cored. Because the more percentages of HH believe corruption is an extreme problem in education, the more likely they will pay for a bribe for childen's education, because people don't want to let other kids surpass their own kids on education.

   (b) edext is an index that is subjective. It's more uncorrelated with other **omited variables**(OV), so $E(edex \cdot u_{i1}) = 0$.

d. Yes, edext is a strong IV. Because by Staiger and Stock (1994), if an IV is good IV, we need to check its F-value on the first stage equation. From our first stage equation, we get t-value is 4.517, and $F - value = t^2 = 4.517^2 = 20.40329 > 10$. Thus, edext is a strong IV.

e. This subquestion is ambiguous. If we consider our traditional OLS method, the marginal effect of cored on growth is -0.03%. If we consider 2sls estimation as in d, the marginal effect is -0.027%. And they are both not statistically different from zeron under 95% confidence level.

*1st year PhD student in Economics Department at Fordham University. Email: wye22@fordham.edu

f. We can use the result of WU-Hausman test to justify the endogenity. From WU-Hausman test[1], the p-value is 0.0309, which is less than 0.05, so there still may be some endogenity in the equation.

g. If the proposed IV is not a good IV, we may need to use Wu-Hausman test, and check the p-value. If the p-value is less than 0.05, then it's exogenous, otherwise, endogenous.

h. From Lewbel (2012), if our traditional IV method is invalid, we can hold on, and construct an artificial IV $(z - \bar{z})\hat{\epsilon_2}$, we call it generated IV. In this question, I use $(GDPPCInit - mean(GDPPCInit))u_{i_2}$ as my generated IV, and use Staiger and Stock (1994) to test first stage OLS's F-value, I found the F-value is $55.83107 > 10$, which means our generated IV is pretty strong.

The R codes associated with this question is as below:

```
    #This file is for my Homework 7
###############
## Wei Ye ####
##Nov,27,2021##
###############
source("ivregh.R")
library(tidyverse)
library(stargazer)
library(ivreg)
corrupt <- read_csv('corrupt.csv')


### For question b.
head(corrupt,n=5)
growth_1_ols <- lm(AvgGrowth~cored+corjud+GDPPCInit,data=corrupt)#z=GDPPCInit
summary(growth_1_ols)
stargazer(growth_1_ols)

## For question d.
growth_2_iv <- ivreg(AvgGrowth~cored+GDPPCInit|edext+GDPPCInit,data=corrupt)
summary(growth_2_iv)
cored_1st_stage <- lm(cored~edext+GDPPCInit,data=corrupt)
summary(cored_1st_stage)
4.517^2

## For question f.
summary(growth_2_iv,diagonistics=TRUE)
```

[1] https://www.statisticshowto.com/hausman-test/

```
## For question h.
#we use (GDPPCInit-E(GDPPCInit))*u_{i2} as our generated IV
attach(corrupt)
gdp_iv_before <- GDPPCInit-mean(GDPPCInit)
cored_ols <- lm(cored~AvgGrowth+corjud+CPI2010)
u2=residuals(cored_ols)
gdp_iv_after <- gdp_iv_before*u2
cored_ols_first <- lm(cored~corjud+gdp_iv_after+CPI2010)
summary(cored_ols_first)
gdp_new_iv_t_value <- summary(cored_ols_first)$coefficients[3,3]
F_value <- gdp_new_iv_t_value^2
F_value
```

Appendix

Table 1: Question b OLS Results

	Dependent variable:
	AvgGrowth
cored	-0.0003^{*}
	(0.0002)
corjud	0.0004^{***}
	(0.0001)
GDPPCInit	-0.001^{***}
	(0.0003)
Constant	0.016^{***}
	(0.003)
Observations	60
R^2	0.376
Adjusted R^2	0.343
Residual Std. Error	0.009 (df = 56)
F Statistic	11.256^{***} (df = 3; 56)
Note:	*p<0.1; **p<0.05; ***p<0.01

Bibliography

Lewbel, Arthur (2012) "Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models," *Journal of Business & Economic Statistics*, Vol. 30, No. 1, pp. 67–80.

Staiger, Douglas O and James H Stock (1994) "Instrumental variables regression with weak instruments."