Homework5

Min-kyeong (Min) Cha

Spring semester 2023

1 Python

1.

The OLS estimate of the regression coefficient of mpg is -22.21, with a standard error of 16.57. In other words, one unit increase of mpg leads to an average decrease in price by -22 USD.

2.

There could be various forms of endogeneity, but the most concerning is omitted variable bias. For instance, a car maker might be affecting mpg and price. Mercedes Benz or Rolls-Royce cars might have low mpg but high prices because they are luxury vehicles.

3.

(a-c)

The two-stage-least-squares estimates with different instrument variable (a.$weight$, b.$weight^2$, c.$height$) are presented in Table 1.

	(a)	(b)	(c)
constant	17627.64	17441.23	-264024.2
	(1667.27)	(1616.45)	(717510.86)
mpg	150.43	157.06	10165.74
	(59.21)	(57.47)	(25513.86)
car(sedan)	-4676.09	-4732.67	-90156.39
	(548.12)	(537.09)	(217753.1)
F-statistic	75.46	79.81	0.0
Observations	1000	1000	1000

Table 1: Results of 2SLS using different instrumental variables. Robust standard errors are presented in parentheses.

(d)

The exclusion restriction states that the covariance of the instrument and the error term in the second stage should be zero. In other words, instrument variables should influence price only through mpg. In this exercise, all three instruments could directly affect car price, not only through mpg. In general, bigger and taller cars are more expensive than small ones.

(e)

The estimated coefficients of mpg, car (indicator variable with sedan=1), and constant change with respect to a different instrument. The estimates of coefficients and their standard errors in columns (a) and (b) are similar. Using weight as an instrument, we can infer one unit mpg increase leads to a 150USD increase in car price in (a) and (b). Interestingly, even though different forms of instrument are used (linear and quadratic), estimates are similar. On the other hand, the result in column (c) using height as an instrument is hugely divergent from (a) and (b), and we can also see absurdly large standard errors.

The first-stage F-statistics of mpg, computed by squaring the t-value of mpg, are larger than 10 in columns (a) and (b) while smaller than 1 in column (c), indicating height as a weak instrument. A further test of a weak instrument might be necessary.

4.

constant	17627.64
	(14153.05, 21102.23)
car	-4676.09
	(-5831.89, -3520.29)
mpg	150.43
	(26.86, 274.01)

Table 2: Results of IV estimate using GMM with weight as an instrumental variable. 95% Confidence intervals are presented in parentheses. Heteroskedasticity-robust standard error is used

IV estimate of mpg using GMM is the same as in Table 1 (a),150.43, though standard error is now different - 95% confidence interval is (26.86, 274.01), indicating bigger standard error. The difference in standard errors might be caused by the number of steps in each estimate - since GMM estimates both stages at the same time, standard errors might be adjusted, while in 2SLS, the first stage standard error is not adjusted for the second stage...?

2 Stata

1.

The result of IV regression with limited information maximum likelihood estimate is presented in Table 3. The estimates are the same as the ones we get in Python, though standard errors are different. The coefficient of mpg means that one unit increase of mpg leads to a 150 USD higher price of car.

2.

The Montiel-Olea-Pflueger effective F-statistic is 78.36, and 5% critical value of LIML is 37.42. Since F-statistic is larger than critical value, we can reject the null hypothesis of weak instrument variable.

	(1)
VARIABLES	price
mpg	150.4**
	(63.05)
car	-4,676***
	(589.7)
Constant	17,628***
	(1,773)
Observations	1,000
R-squared	0.104

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Results of IV estimate with the limited information maximum likelihood method. 'weight' variable is used as an instrument. Robust standard errors are presented in parentheses.