Homework 5

Environmental Economics II
Maghfira Ramadhani

February 18, 2024

Problem 1 Python

1. The OLS estimates from regression of price on *mpg* with *car* and a constant is -131.04. It means that more efficient car (higher *mpg*) is associated with lower price which does not make sense.

2. When estimateng the effect of fuel efficiency on price of vehicle, we should be careful with the endogeneity problem. The fuel efficiency of a vehicle is likely to be correlated with unobserved characteristics of the vehicle that are also correlated with the price of the vehicle. For example, the fuel efficiency of a vehicle may be correlated with the size of the vehicle, which is also correlated with the price of the vehicle. This is a problem because the OLS estimates will be biased and inconsistent.

(a)-(c) Table 1 show the by hand 2SLS estimates

Table 1: Two-stage least squares estimates of the effect of treatment on bycatch yield.

	(a)	(b)	(c)
Miles per gallon	150.43	157.06	10165.74
	(59.30)	(57.56)	(25552.22)
=1 if the vehicle is sedan	-4676.09	-4732.67	-90156.39
	(548.94)	(537.90)	(218080.47)
Instrumental variable	Weight	Weight^2	Height
First Stage F-statistic	256.80	257.02	203.66

(d) The instrumental variable must satisfy the relevancy conditions and exclusion restrictions. The relevancy condition requires that the instrument is correlated with the endogenous variable. The exclusion restriction requires that the instrument is not correlated with the error term, or in other words, the instrumental variable affect the outcome variable only through the endogenous variable.

Results from using the height of vehicle as the instrumental variable do not seems reasonable. I suspect that the exclusion restriction is not satisfied. The height of vehicle is likely to be correlated with whether or not the vehicle is sedan or SUV. Figure 1 indicates that the height of vehicle is also affecting the price of vehicle through the type of vehicle, which violates the exclusion restriction.
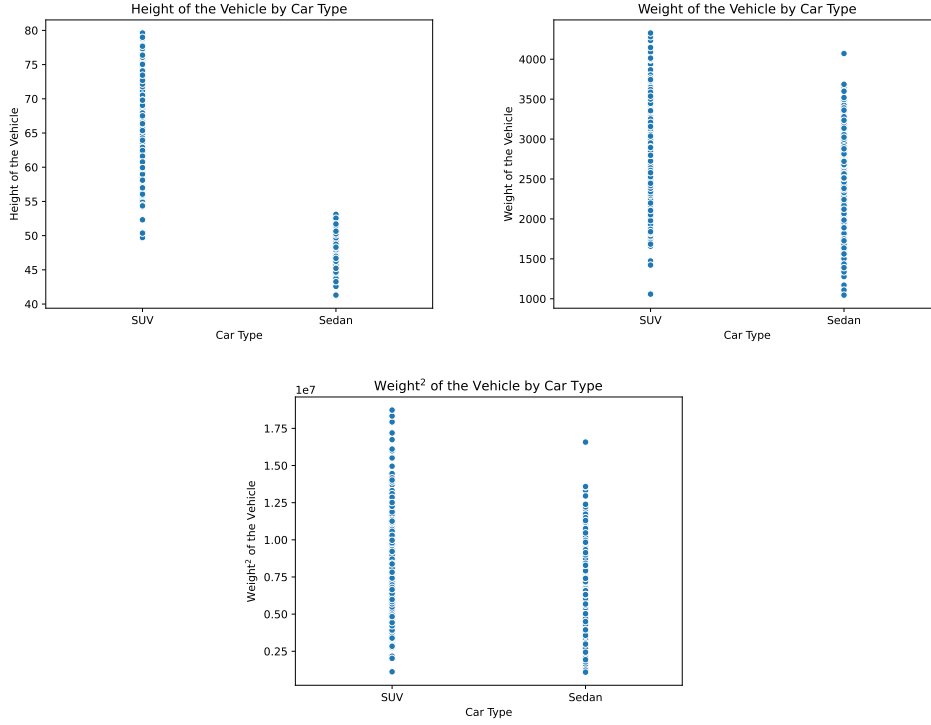
Figure 1: Scatter plot of variables used as IV and car type

(e) From the result in Table 1, the estimates from column (a) and (b) are statistically significant, and both are not statistically different. However, the estiamtes from column (c) is practically zero, with estimates that is very large.

3. Table 2 show the IVGMM estimates

Table 2: Comparing 2SLS and IVGMM estimates

	By Hand 2SLS	IVGMM
Miles per gallon	150.43	150.43
	(59.30)	(63.05)
=1 if the vehicle is sedan	-4676.09	-4676.09
	(548.94)	(589.70)

The discrepancy in the standard errors is because the standard error estimates in the by hand 2SLS is basically wrong if we use directly from the second stage estimates. To compute the correct standard errors, we need to compute the residuals correctly. The correct residuals is using the coefficient estimates from the second stage but using the original variables, i.e. $\hat{e}_i = X_i \hat{\beta}_{2^{nd}Stage}$. Instead what we use is the wrong residuals $\tilde{e}_i = \hat{X}_i \hat{\beta}_{2^{nd}Stage}$.

Problem 2 Stata

1. Table 3 shows the IV LIML estimates. The estimates are identical to the result from using IVGMM in the previous problem.

	IV LIML
	Table 3: IV LIML estimates
Miles per gallon	150.43
	(63.05)
=1 if the vehicle is sedan	-4676.09
	(589.70)
Montiel-Pflueger F-statistics	78.36
LIML critical value for $\tau = 5\%$	37.42
Observations	1000

2. The effective F-stat and the critical value is 78.36 and 37.42, respectively. Based on Pflueger and Wang (2015) if the effective F-stat is greater than the critical value, then we can conclude that we rejects the null hypothesis that the instrument is weak.

References

Pflueger, C. E. and S. Wang (2015). A robust test for weak instruments in stata. *The Stata Journal 15*(1), 216–225.