

Homework 5

Economics 7103

Due Monday, February 20th by 11:59 pm

You have access to imaginary vehicle sales data from 2017 (*instrumentalvehicles.csv*). You are interested in estimating the hedonic price of an additional mile per gallon as part of a larger analysis of willingness to pay for fuel efficiency. In your data, you have the following variables:

| Variable | Description |
|---------------|---|
| <i>price</i> | Sales price of the vehicle in USD |
| <i>car</i> | Class of the vehicle. =1 if the vehicle is a sedan, =0 if the vehicle is an SUV |
| <i>mpg</i> | Fuel efficiency in miles per gallon |
| <i>weight</i> | Weight of the vehicle in pounds |
| <i>height</i> | Height of the vehicle in inches |
| <i>length</i> | Length of the vehicle in inches |

Table 1: Variable descriptions for homework 6.

1 Python

1. Run the ordinary-least-squares regression of price on *mpg*, the *car* indicator variable, and a constant. Report and interpret the coefficient on miles per gallon (do not construct a table).
2. What forms of endogeneity are you concerned about when estimating the coefficient on *mpg*?
3. To correct for this endogeneity, you would like to use instrumental variables. Specifically, you are interested in the system of equations:

$$price_v = \beta_0 + \beta_1 mpg_v + \beta_2 car_v + e_v \quad (1)$$

$$mpg_v = \gamma_0 + \gamma_1 z_v + \gamma_2 car_v + u_v, \quad (2)$$

where z_v is the value of the instrument for vehicle v and e_v and u_v are error terms. Report the estimated second-stage coefficients, standard errors or confidence intervals, and the first-stage F -statistic for the excluded instrument in the same table for the following procedures (just use a regular F -statistic for this exercise rather than the robust Montiel-Olea-Pflueger F -statistic):

- (a) Perform two-stage-least-squares estimation by hand using *weight* as the excluded instrument. (First regress *mpg* on all of the instruments. Save the fitted values from the first stage \hat{mpg} and use the fitted values in place of the endogenous variable in the second stage price regression.)
 - (b) Perform two-stage-least-squares estimation by hand using *weight*² as the excluded instrument.
 - (c) Perform two-stage-least-squares estimation by hand using *height* as the excluded instrument.
 - (d) In words, what are the different exclusion restrictions required for parts (a)-(c)? Does this seem reasonable for these instruments?
 - (e) Compare and contrast the estimated coefficient on *mpg* from parts (a)-(c). What explains the discrepancies?
4. Calculate the IV estimate using GMM with *weight* as the excluded instrument. (Look for the Linear-models function `IVGMM`). Report the estimated second-stage coefficient and standard error or confidence interval for *mpg*. What factors account for the differences in the standard errors?

2 Stata

1. Use the `ivregress liml` command to compute the limited information maximum likelihood estimate using `weight` as the excluded instrument. Report your second-stage results in a nicely-formatted table using `outreg2`. Use heteroskedasticity-robust standard errors.
2. Use `weakivtest` to estimate the Montiel-Olea-Pflueger effective F-statistic. What is the 5% critical value, the F-statistic, and conclusion?