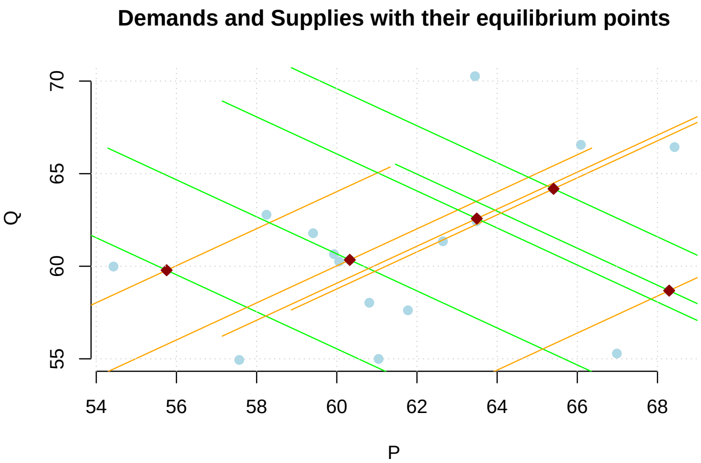
assignment8

William Hope

2023-03-27

Part 1

For this part, we need to generate new demand and supply data, using different functions (hidden above). It looks similar to the previous assignment, but the demand now depends on two exogenous regressors.



$$
\begin{aligned}
Q^d &= 100 - 1P + 2X_1 + 0.007600531X_2 + \eta \\
Q^s &= 20 + 1P - 2Z + e
\end{aligned}
$$

Estimating the Supply function

The model we want to analyze is the following supply and demand system:

$$
\begin{aligned}
Q^d &= \delta_1 + \delta_2 P + \delta_3 X_1 + \delta_4 X_2 + \eta \\
Q^s &= \alpha_1 + \alpha_2 P + \alpha_3 Z + e
\end{aligned}
$$

where X_1, X_2 and Z are exogenous variables (uncorrelated with η and e).

First, we generate a dataset using my student ID as seed by running the following code.

We only want to estimate the supply function. I want you to try the following sets of excluded exogenous variables: (i) $\{X_1\}$, (ii) $\{X_2\}$ and (iii) $\{X_1, X_2\}$. Don't forget that the set of instruments must include the included exogenous variable (Z in this case). For each set of instruments, answer the following questions.

Question 1

Estimated the reduced form and testing if the instruments are strong.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	12.8722757	0.4333668	29.70296	0
X1	-0.5008333	0.0404667	-12.37642	0

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	8.4686379	0.7509583	11.2771084	0.0000000
X2	-0.0661183	0.0752104	-0.8791111	0.3814915

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	15.2619728	0.7323478	20.8397866	0.0000000
X1	-0.5318743	0.0433664	-12.2646573	0.0000000
X2	0.0116102	0.0455464	0.2549103	0.7993377
Z	-0.2149096	0.0394187	-5.4519699	0.0000004

When testing X1 by itself, it is significant. X2 is not significant. Morever, when testing all the instrumental variables, X1 and Z are significant.

Question 2

Using `ivreg` to estimate the model by TSLS. Interpretation below.

```
##
## ==========================================================
##                              Dependent variable:
##                     --------------------------------------
##                                      Q
##                      instrumental           OLS
##                        variable
##                          (1)                (2)
## ----------------------------------------------------------
## P                      3.366***            0.224
##                        (0.345)            (0.303)
##
## X1                                        -1.574***
##                                           (0.189)
##
## Constant              32.685***          73.139***
##                        (2.780)            (4.006)
##
## ----------------------------------------------------------
## Observations             100                100
## R2                      0.192              0.647
## Adjusted R2             0.184              0.639
```

```
## Residual Std. Error 5.071 (df = 98)    3.371 (df = 97)
## F Statistic                        88.779*** (df = 2; 97)
## ==========================================================
## Note:                          *p<0.1; **p<0.05; ***p<0.01

##
## ==========================================================
##                                Dependent variable:
##                      -------------------------------------
##                                         Q
##                       instrumental           OLS
##                         variable
##                            (1)               (2)
## ----------------------------------------------------------
## P                         2.494            1.955***
##                          (2.704)           (0.258)
##
## X2                                          -0.036
##                                            (0.168)
##
## Constant                 39.517*          44.077***
##                         (21.191)           (2.608)
##
## ----------------------------------------------------------
## Observations               100               100
## R2                        0.366             0.396
## Adjusted R2               0.360             0.383
## Residual Std. Error 4.492 (df = 98)    4.408 (df = 97)
## F Statistic                        31.759*** (df = 2; 97)
## ==========================================================
## Note:                          *p<0.1; **p<0.05; ***p<0.01

##
## ==========================================================
##                                Dependent variable:
##                      -------------------------------------
##                                         Q
##                       instrumental           OLS
##                         variable
##                            (1)               (2)
## ----------------------------------------------------------
## P                         2.418            1.060***
##                          (2.704)           (0.258)
##
## X1                                          -1.055
##
##
## X2                                         -0.0001
##                                            (0.168)
##
## Z                                            0.727
##
##
## Constant                 40.112*          53.958***
```

```
##                          (21.191)           (2.608)
##
## ------------------------------------------------------------
## Observations                 100                100
## R2                          0.374              0.745
## Adjusted R2                 0.368              0.734
## Residual Std. Error 4.464 (df = 98)    2.893 (df = 95)
## F Statistic                            69.456*** (df = 4; 95)
## ============================================================
## Note:                            *p<0.1; **p<0.05; ***p<0.01
```

When reviewing the results, it is clear IV does a better job with X1, while OLS does a better job with the other models.

Question 3

Testing if the instruments are valid. Interpretation below.

```
##                    df1 df2 statistic      p-value
## Weak instruments     1  98  121.0323 8.236646e-19
## Wu-Hausman           1  97   68.9877 5.982294e-13
## Sargan               0  NA        NA           NA
```

```
##                    df1 df2  statistic   p-value
## Weak instruments     1  98 0.90148761 0.3447188
## Wu-Hausman           1  97 0.04326128 0.8356702
## Sargan               0  NA         NA        NA
```

```
##                    df1 df2 statistic      p-value
## Weak instruments     3  96 62.974017 1.354025e-22
## Wu-Hausman           1  97  7.315184 8.076470e-03
## Sargan               2  NA 49.090187 2.188777e-11
```

We have strong instruments when we use X1, X2, and Z.

Question 4

Testing if P is exogenous. Interpretation below.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	32.685237	2.0427096	16.000922	0
P	3.366471	0.2463531	13.665228	0
uhat	-3.142730	0.3835436	-8.193933	0

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	39.5167691	20.357060	1.9411825	0.0551385
P	2.4939413	2.598396	0.9598002	0.3395422

	Estimate	Std. Error	t value	Pr(>|t|)
uhat	-0.5385073	2.561477	-0.2102331	0.8339267

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	40.112122	2.6553813	15.105974	0.0000000
P	2.417902	0.3216660	7.516811	0.0000000
uhat	-1.358009	0.4872333	-2.787185	0.0063974

It appears P is exogenous when we use X2, but not exogenous in the other cases.

Ideally, it appears that the 1st method is not better among the 3. The reason why is because the 2nd method, using AER performs multiple tests and allows you to decide with fewer lines of code. It is also much easier to read and understand whether the IV are significant. The 1st does provide the same/similar results, but the 2nd and 3rd methods are much more concise.

Part 2

For this part, we want to work with the fish data used by Kathryn Graddy in the paper *Fulton Fish Market* published in the Journal of Economic Perspectives in 2006. The dataset contains daily data from December 1991 to March 1992. We use the `read.table` function to extract the data.

Note: I use my specific directory, which must be changed for external use to your directory.

Question 1:

We reproduce the Table 2 from the paper. For the IV regressions (there are 2 in the table), we test the exogeneity of price and the relevance of the instrument (Stormy). Interpretation below.

```
##
## ==========================================================================================
##                                               Dependent variable:
##                      ---------------------------------------------------------------------
##                                                   LogQuantity
##                                      OLS                                 instrumental
##                                                                            variable
##                               (1)                   (2)               (3)              (4)
## ------------------------------------------------------------------------------------------
## LogPrice                   -0.541***             -0.545***         -0.541***       -1.223**
##                            (0.170)               (0.194)           (0.170)         (0.594)
##
## Monday                                            0.032                             -0.033
##                                                  (0.178)                           (0.167)
##
## Tuesday                                          -0.493***                        -0.533***
##                                                  (0.180)                           (0.181)
##
## Wednesday                                        -0.539***                        -0.576***
##                                                  (0.190)                           (0.177)
##
```

```
## Thursday                                      0.095                                  0.118
##                                              (0.177)                                (0.182)
##
## Cold                                          -0.062                                 0.068
##                                              (0.165)                                (0.176)
##
## Rainy                                          0.067                                  0.072
##                                              (0.163)                                (0.173)
##
## Constant               8.419***              8.617***            8.419***          8.442***
##                        (0.082)               (0.147)             (0.082)           (0.241)
##
## -----------------------------------------------------------------------------------------------
## Observations             111                   111                 111               111
## R2                      0.078                 0.223               0.078             0.110
## Adjusted R2             0.069                 0.170               0.069             0.049
## Residual Std. Error    0.716 (df = 109)      0.676 (df = 103)    0.716 (df = 109) 0.723 (df = 103)
## F Statistic        9.167*** (df = 1; 109) 4.220*** (df = 7; 103)
## ===============================================================================================
## Note:                                                             *p<0.1; **p<0.05; ***p<0.01
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	8.3137875	0.1127974	73.705506	0.0000000
LogPrice	-1.0824089	0.4556215	-2.375675	0.0192792
uhat	0.6443279	0.5125841	1.257019	0.2114585

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.2717054	0.0974460	-2.7882665	0.0062922
Stormy	0.3464055	0.0749654	4.6208702	0.0000109
Monday	-0.1129225	0.1187867	-0.9506321	0.3439742
Tuesday	-0.0411493	0.1195600	-0.3441728	0.7314047
Wednesday	-0.0118250	0.1164771	-0.1015225	0.9193293
Thursday	0.0496456	0.1183507	0.4194789	0.6757232

```
##                   df1 df2 statistic      p-value
## Weak instruments    1 103 14.611914 0.0002263443
## Wu-Hausman          1 102  2.149674 0.1456763781
## Sargan              0  NA        NA           NA
```

When looking at the new table, reproduced from the paper, we can see that the coefficients are similar and the standard errors are very similar. While some are different, it is clear that our tables are the same when calculating each variable's significance from the paper and comparing the results.

When testing the exogeneity of log price, we see that the coefficient of uhat is highly insignificant, so we strongly accept the hypothesis that ed76 is exogenous.

Question 2

We cannot test the validity of the instrument when we only use *Stormy*. We estimate the demand using the instruments `Cold` and `Rainy` and perform all tests (validity, relevance and exogeneity).

```
## [1] 0.3413332
```

```
## [1] 3.841459
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	8.3063790	0.1624331	51.1372264	0.0000000
LogPrice	-1.1206597	0.7531268	-1.4880094	0.1396625
uhat	0.6159459	0.7962698	0.7735392	0.4408932

Interpretation of the results. Which instrument(s) seems to be the best choice?

We do not reject the hypothesis that the instruments are exogenous.

We can see that the coefficient of uhat is highly insignificant, we strongly reject the hypothesis that ed76 is NOT exogenous.

Question 3:

We want to see in this question if the demand function can be estimated. Assume the following demand and supply:

$$
\begin{aligned}
\log(Q^s) &= \alpha_1 + \alpha_2 \log(P) + \alpha_3 Mon + \alpha_4 Tue + \alpha_5 Wed + \alpha_6 Thu + \alpha_7 Stormy + e \\
\log(Q^d) &= \delta_1 + \delta_2 \log(P) + \delta_3 Cold + \delta_4 Rainy + u
\end{aligned}
$$

Can we estimate the supply function? We use TSLS and by performing the appropriate tests.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	8.5059113	0.1567027	54.2805632	0.0000000
LogPrice	-1.1194169	0.4587852	-2.4399589	0.0163637
Monday	-0.0254022	0.2293157	-0.1107738	0.9120072
Tuesday	-0.5307694	0.2068325	-2.5661799	0.0116930
Wednesday	-0.5663511	0.2129136	-2.6600041	0.0090392
Thursday	0.1092673	0.1822478	0.5995535	0.5500955

```
##                    df1 df2 statistic      p-value
## Weak instruments     1 105 21.517361 1.015269e-05
## Wu-Hausman           1 104  2.273104 1.346678e-01
## Sargan               0  NA        NA           NA
```

We see that the test results in strong instruments. It appears we can estimate the supply function.