Homework 6: Industrial Organisation

Dhananjay Ghei

December 11, 2018

N.B. The code for this exercise was written in `R` *and is available on my Github account.* `www.github.com/dhananjayghei/io_estimation`.

Some basics

1. Read the data into a statistical package and look at summary statistics to convince yourself that the data was read in correctly. Try a simple OLS regression of log(QUANTITY) on a constant, log(PRICE), LAKES, and (twelve of) the seasonal dummy variables. If you were to view this as an estimate of a demand curve what would the price elasticity of demand be? Why does this number seem unreasonable?

Table 1 shows the summary statistics from Porter's data set. The number are the same as in Table II of Porter. Thus, the data has been read in correctly. Figure 1 shows the plot of price as a function of time (in weeks). This replicates the Figure I from Porter.

Table 1 Summary statistics

The table shows the summary statistics from Porter's data set.

	Mean	Sdev	Min	Max
Price	0.246	0.067	0.125	0.400
Quantity	25384.000	11633.000	4810.000	76407.000
Lakes	0.573	0.495	0.000	1.000
Collusion	0.619	0.486	0.000	1.000

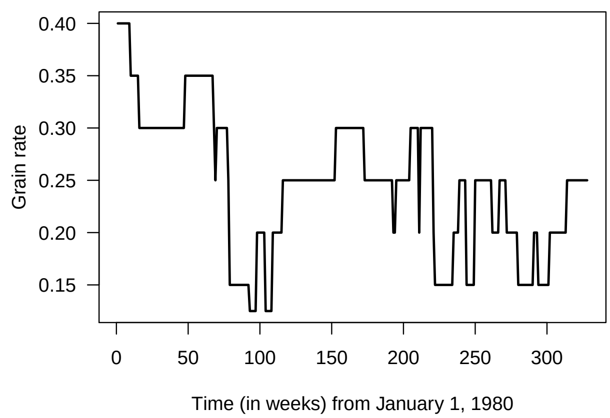
Column I of Table 2 shows the OLS estimates of the demand equation. Demand elasticity from OLS is equal to 0.639. The demand is inelastic. This seems unreasonable because if we expect JEC to act as a monopolist, prices should go to infinity. One potential reason for this is that the price is endogenous which gives the estimate of elasticity to be biased towards zero.

2. Try doing the regression instead using instrumental variable with the COLLUSION variable as the instrument for PRICE. How does the reported price elasticity change. Is the estimate closer to that in Porter's paper or that in Ellison's paper and why? How do you interpret the coefficient on the LAKES variable? On the seasonal dummies? What is the R-squared of the regression and what do you make of it?

Column II of Table 2 shows the IV estimate using collusion as the instrument for price. Instrumenting increases the estimated demand elasticity from .639 to .867. This is closer to the Ellison's paper. For the Lakes variable, going from 0 to 1 indicates the opening of great lakes for navigation. The difference in expected log quantities is given by $\log(Q_1) - \log(Q_0) = -.423$ which gives $\frac{Q_1}{Q_0} = \exp(-.423) = .66$.

Figure 1 Plot of price as a function of time

The figure shows the plot of price as a function of time. The time is in weeks starting from Jan 1, 1980.



Thus, opening of lakes lead to a reduction of 34% in the quantity shipped given price. The interpretation on the seasonal dummies is similar to the one for lakes. In fact, there are some seasonal dummies which show statistical significance. In the IV case, the R-square coefficients do not hold much significance as the IV equation is an estimating equation and not the structural equation. This is the reason we do not get an F-stat for the two regressions.

3. Try the regression with the DM1-DM4 and COLLUSION as instruments for price. Do the estimates "improve" in any way?

Column III of Table 2 shows the instrumental variable estimation using the collusion variable and the dummy variable as instrument. This does not help in changing the elasticity.

4. Estimate a supply equation as in Porter and Ellison using the LAKES variable as an instrument for quantity. What does the magnitude of the coefficient on COLLUSION tell us about the effect of collusion on prices? What might the coefficient on QUANTITY in this regression indicate about the nature of costs in the JEC?

Table 3 shows the estimation of supply equation using OLS and IV where the lakes variables is used as an instrument for quantity. The coefficient on collusion in the case of IV is 0.43. Thus, the ratio of prices is equal to exp(.430) = 1.537 and therefore, the prices are 54% higher in collusion. If you compare the results from Column II of Table 3 with Column II of Table III from Porter, they are pretty similar even though we do not have the estimated dummy variable as in Porter. The coefficient on quantity is .485 which is positive and statistically significant. This could suggest that there is either increasing marginal costs or misspecification in the system.

Model derivation and interpretation

1. Suppose that rather than the log-log specification of demand you've been using so far, you tried others and found that a linear specification of demand like

$$Q_t = \alpha_0 + \alpha_1 P_t + \alpha_2 Lakes_t + u_t$$

Table 2 Demand estimation

The table shows the demand estimation using OLS and instrumental variable estimation. Column I shows the results of OLS, Column II shows the results of IV using collusion as an instrument for price and Column III shows the results of IV using collusion and the dummies (DM1-DM4) as instruments for price.

	Dependent variable:		
	OLS	IV	IV
	(1)	(2)	(3)
Constant	9.176***	8.865***	9.045***
	(0.135)	(0.196)	(0.180)
Log(Price)	−0.639***	−0.867***	−0.735***
	(0.082)	(0.132)	(0.120)
Lakes	−0.448***	−0.423***	−0.437***
	(0.120)	(0.122)	(0.120)
SEAS2	0.200*	0.222**	0.209*
	(0.106)	(0.108)	(0.107)
SEAS3	0.244**	0.267**	0.254**
	(0.106)	(0.108)	(0.107)
SEAS4	0.288***	0.283***	0.286***
	(0.108)	(0.109)	(0.108)
SEAS5	0.242*	0.205	0.226*
	(0.131)	(0.134)	(0.132)
SEAS6	0.180	0.125	0.157
	(0.164)	(0.168)	(0.165)
SEAS7	0.255	0.191	0.228
	(0.164)	(0.169)	(0.166)
SEAS8	−0.102	−0.163	−0.128
	(0.164)	(0.168)	(0.166)
SEAS9	0.136	0.073	0.109
	(0.164)	(0.169)	(0.166)
SEAS10	0.302*	0.217	0.266
	(0.165)	(0.172)	(0.169)
SEAS11	0.348**	0.283*	0.320*
	(0.164)	(0.169)	(0.167)
SEAS12	0.352**	0.310*	0.334**
	(0.163)	(0.166)	(0.165)
SEAS13	0.133	0.131	0.132
	(0.111)	(0.112)	(0.111)
Observations	328	328	328
R^2	0.313	0.296	0.310
Adjusted R^2	0.282	0.264	0.279
Residual Std. Error (df = 313)	0.397	0.402	0.398
F Statistic	10.169*** (df = 14; 313)		
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3 Supply estimation

The table shows the supply side estimation using OLS and instrumental variable estimation. Column I shows the results of OLS, Column II shows the results of IV using the lakes variable as an instrument for quantity.

	Dependent variable:	
	OLS	IV
	(1)	(2)
Constant	-0.934^{***}	-6.328^{***}
	(0.284)	(1.103)
D1	-0.241^{***}	-0.192^{***}
	(0.046)	(0.068)
D2	-0.199^{***}	-0.199^{*}
	(0.069)	(0.102)
D3	-0.308^{***}	-0.346^{***}
	(0.048)	(0.071)
D4	-0.332^{***}	-0.093
	(0.106)	(0.162)
Collusion	0.296^{***}	0.430^{***}
	(0.027)	(0.048)
Log(Quantity)	-0.044	0.485^{***}
	(0.027)	(0.108)
Observations	328	328
R^2	0.455	−0.178
Adjusted R^2	0.445	−0.200
Residual Std. Error (df = 321)	0.215	0.317
F Statistic	44.638^{***} (df = 6; 321)	
Note:		*p<0.1; **p<0.05; ***p<0.01

seemed most appropriate. Show that for this demand curve the optimal price for a monopolist with a constant marginal cost of c to set is

$$P_t = c - \frac{1}{\alpha_1} Q_t$$

where, $\alpha_1 \leq 0$. Given this result, what functional form would you choose for the supply curve in this model?

The monopolist will solve:

$$\max_p (p - c)Q$$

The first order condition is given by: $(p-c)\frac{\partial Q}{\partial p} + Q = 0$. From the demand function, we know that $\frac{\partial Q}{\partial p} = \alpha_1$, so we have $p = c - \frac{1}{\alpha}Q$. The appropriate functional form for the supply curve will be:

$$p = \beta_0 + \beta_1 Q$$

where, $\beta_1 = -\frac{1}{\alpha}$.

2. What pricing rule would result with this demand curve if the industry instead consisted of perfectly competitive firms with total costs of the form $c(Q_t) = c_0 Q_t + c_1 Q_t^2$ setting price equal to marginal cost? Could one use an approach like Porter's to distinguish between these two models of behavior? Talk about why this is an important question.

If the industry was perfectly competitive, then the firms would price at marginal cost. Therefore, the supply equation will be: $p_t = c_0 + 2c_1 Q_t$ which is derived from the firm's first order condition. Note that both the monopoly pricing solution and the supply equation from the perfect competitive are increasing and linear in Q. Thus, we can not tell them apart if we were doing standard estimation.

Porter's approach works here if we assume that there was enough variation in the conduct of firms and we have an indicator for it. That is, we have the collusion variable in our model.
Recall, that Porter's approach was to add the indicator variable (I_t, collusion - which is observable in his case.) which distinguishes between the monopoly pricing and the perfectly competitive supply:

$$p_t = \beta_0 + \beta_1 Q_t = \beta_3 I_t + u_t$$

Alternatively, one could also estimate it using a hidden Markov structure using Ellison given that we assume that there were only two (or, three) states in the sample.
This is an important question, because we have an indicator variable that gives evidence of change in conduct. Without it, the identification between the monopoly and perfect competition will be difficult unless we make assumptions on the cost structure (which is not common to do in the literature) or have an additional variable which would impact the monopolist's decision but not of the perfectly competitive to identify the differences in the two regimes.

Causes of price wars

1. Using the collusion variable generate an indicator variable for the start of a price war. Perform a probit regression with this indicator as a dependent variable and with QUANTITY, LAKES, and DM1-DM4 (or a subset thereof) as explanatory variables. What inferences might you want to draw about whether price wars are more likely to occur in booms from the coefficients on the first two variables? Why are these variables not really the right ones to be using in the equation?

I construct the indicator variable for price war directly from the collusion variable. Price wars is equal to one when Collusion shifts from 1 to 0 and 0 otherwise. This gives us 11 events of price wars. As a result, we have very little variation in our dependent variable.
Table 4 shows the estimation from probit results. Heteroskedasticity consistent standard errors are used given the relatively smaller sample. The dummies turn out to be significant, whereas the lakes and quantity variable do not turn out to be significant.

Table 4 Cause of price wars

The table shows the results from probit estimation. Heteroskeadsticity consistent standard errors are used in the analysis.

	Dependent variable:
	Price war
Constant	-5.351^{***}
	(0.202)
Quantity	-0.00001
	(0.00001)
Lakes	-0.379
	(0.290)
D1	3.818^{***}
	(0.279)
D2	4.423^{***}
	(0.524)
D3	4.268^{***}
	(0.241)
D4	-0.220^{*}
	(0.122)
Observations	328
Log Likelihood	-45.106
Akaike Inf. Crit.	104.211
Note:	*p<0.1; **p<0.05; ***p<0.01