Economics 1660: Big Data Prof. Daniel Björkegren

PS 7: Price Setting

Continued:
Your friend runs a firm. After hearing about your experience using big data, she has asked to help her optimize the price of her product. She has provided historical data for you to analyze.

Your friend found the results from last week confusing. Let's unpack this.

1. **Revisiting Elasticities**

Consider the regression you ran in the previous problem set:

$$\log(D) = b + \varepsilon \log(p) + u$$

a. **What assumption on *u* is required for an OLS estimate of ε to represent the price elasticity of demand?**

b. **Describe a scenario that could result in a positive OLS estimate of ε, when the true underlying price elasticity is negative.** (It may be useful to think about a standard supply/demand diagram here)

Revisit the file demand_monopoly.csv. It includes price and quantity observations for her firm.

It also includes data on *mcShifter*, which is a variable affecting the marginal cost of producing a product.

c. **What assumptions are required in order to use *mcShifter* as an instrument to estimate the price elasticity of demand using instrumental variables?** (In terms of the standard supply/demand diagram: How does a change in the production costs likely affect the diagram?)

d. **Construct an instrumental variables estimate of ε. You may use any statistical software for part (d) including STATA, R or Python.**
   - **First, regress *p* on *mcShifter*. Describe and interpret your results.**
   - **Second, compute the 2SLS estimate of ε using *mcShifter* as an instrument.**[1] **Interpret your results.**
   - **Does your IV estimate suggest that your friend should raise or lower prices?**

[1] In R, you can use the ivreg command after installing the package AER.

e. Theory / Projection

We actually have two mechanisms at work simultaneously:

- A demand equation, where a given consumer takes price as given and decides how much to buy at any given price
- A supply equation, where the firm sets the price given a demand function.

This gives the following system:

$$\log(\mathrm{D}) = \alpha_d + \beta_d \log(\mathrm{p}) + u_d$$
$$\log(\mathrm{p}) = \alpha_s + \beta_s \log(\mathrm{D}) + u_s$$

- **In terms of the variables of the true regression function, what do you obtain when you regress log(D) on log(p) using OLS? How does this compare to the parameter we're interested in?** (Hint: Use $\hat{\beta}_{OLS} = \frac{\mathrm{Cov(X,Y)}}{\mathrm{Var(X)}}$ and plug in the expression for log(p) you obtain by solving the system.)

- **What about if you regress D on p using mcShifter as an instrument for p?** (Hint: Use $\hat{\beta}_{IV} = \frac{\mathrm{Cov(Z,Y)}}{\mathrm{Cov(Z,X)}}$.)

f. Simulation

For the remainder of this problem set we will use Python.
The true demand equation is actually given by the following:

$$\mathrm{D(p)} = \frac{e^{v_i - ap_i}}{e^{v_i - ap_i} + e^{u_{out}}}$$

where:
v_i is is the "base utility" a consumer gets from product i if they didn't have to pay for it
u_{out} is the utility consumers get from buying nothing (the outside option)
a is the price sensitivity of consumers
p_i is the product you charge for product i

- Write a function in Python that will compute the market share and profit that results from a particular price.
  What profits do you obtain, as a function of your price? Graph profits vs. price. Comment.
  For this part, fix the marginal cost (mc) to 1, the base utility of the product v_i to 2, the price sensitivity a to 1, and the outside utility u_{out} to 1.

- Now write a function in Python that will find the profit maximizing price and corresponding quantity for any given parameters a, mc, u_{out} and v_i.
  With the true model, what are the optimal prices, as a function of the supply shifter (MC)? Draw a graph that depicts this optimal price as a function of marginal costs.

What are the profits that result? Also draw the corresponding profits as a function of marginal costs. (Keep the other parameters fixed as before.)

- **What are the profits generated by the true model if instead you set prices according to:**
    - **The OLS estimates (from last week)**
    - **The IV estimates (from part d)**

  In other words, you pretend the relationship between price and quantity is as estimated in your regression. You use this relationship in your profit maximization. Once you've decided on a price, you want to see how well this would do in the real world. Usually we can't test this (we'd have to just change our price and see what happens in the market), but because we have the true model, you can calculate the profits you would have obtained with your price.
  For both parts, start by deriving an expression for the optimal price as a function of the marginal costs. Depict the corresponding profits as a function of marginal costs graphically.

- **Fix the marginal costs again at 1. If you had to set prices based on evidence:**
    - **If you had the OLS estimates, how much would it be worth to pay for the IV estimates?**
    - **If you had the IV estimates, how much would it be worth to pay for the true model?**

g. We will finish with a brief market simulation. You actually already wrote almost all of the code that you need for this.
The data in demand_monopoly.csv was generated in a market where there were shocks to both supply and demand. Now, we'd like you to generate data from a market where there are only shocks to supply. Once you've generated the data, we'll see how different estimation techniques perform.

For each period, draw a random value for marginal cost (this will represent your supply shock. We recommend drawing a value from a uniform distribution ranging between 0.2 and 4.0). Find the price that an optimizing firm would charge under that marginal cost (if they had access to the true model), and compute the resulting market share. Compute outcomes for 1000 periods.

Save those 1000 observations (of prices, market shares, and marginal costs) and estimate the price elasticity from this dataset using both OLS and IV. What do you observe?

h. A classmate asks, generally, when it is ok to use OLS estimates and when you need to worry about identification. What can you tell them?

i. BONUS: Explore adding demand shocks to your simulation from part (g) (random shocks to the utility from the outside option). Try different magnitudes of shocks. How large do these shocks have to be to affect the usefulness of the OLS estimates?

2. Final Project Proposal

Also turn in a short proposal for your final project (2 pages is a good length to shoot for). This is to help you formalize ideas for your project and to give us a chance to provide feedback before you get too far.

Each group can turn in one proposal.

Your proposal should describe:

- Who is in your group
- Description of any data sets you will be using (make sure you have or can obtain access)
- The questions you want to answer
- How you plan to answer those questions

Your proposal won't be assigned a grade; we want you to use it to help you develop a better final project. So, we encourage you to use this as an opportunity to request specific feedback on your ideas: for example, if you are weighing two possible ideas you could describe both and ask for advice.