Author:

P. Carmel Marie Frédérique ZAGRÉ

```
julia> first(noinstrument_data, 10)
10×7 DataFrame
```

Row	marketid Int64	firmid Int64	share Float64	id Int64	price Float64	caffeine_score Float64	costattributes Float64
1	1	1	0.244752	1	1.87059	0.966159	0.0866591
2	1	2	0.0126449	2	2.58938	0.105214	0.0485192
3	1	3	0.0282693	3	2.2279	0.217468	0.732495
4	2	1	0.0641745	1	1.90561	0.597104	0.00554706
5	2	2	0.153911	2	1.94316	0.843626	0.92502
6	2	3	0.116168	3	1.65621	0.632427	0.0154237
7	3	1	0.0133731	1	2.5564	0.3281	0.964173
8	3	2	0.260673	2	1.67734	0.559408	0.491374
9	3	3	0.0625817	3	1.71425	0.132648	0.175674
10	4	1	0.238495	1	1.51839	0.734724	0.892568

Figure 1: Overview of the reshape function

1 BLP Estimation

1.1 Data Analysis

The first step in our estimation consist of transforming the different given datasets into the format that is usable for the BLP estimation used in class. I decided STATA as a software, specifically, I used the command "reshape" available in stata. reshape converts data from wide to long form and vice versa. In figure 2, I presented a table of the overview of the reshape function. In our case, we had to convert the data from wide to long. Before calling the function, we had to rename the variables so that we are in the wide format. For example, in the dataset : midterm_simulated_market_data_s.csv, we renamed the variable md, th and sb by share1, share2 and share3 respectively. Then we proceed with the reshaping.

The reshape creates a new variable call "firmid" that cotain the firm ID number. In our case, md has ID =1, th has ID =2 and sb has ID =3.

We repeat this procedure for all the provided dataset, and we finally merge all of them using the merge function in STATA. We provided our code in the file under the name : BLP-DataCleaning.

Note :

- If you want to run the code, please, modify the directory.
- We combined all the variables : share, instruments, price and characteristic into one final data set call : MIDTERM1_FinalDataSet.csv. However, as it was ask to exclude the instruments, we construct separate dataset :
  - FinalDataSet1.csv : Combined data on s, x and w. See figure 1 for the first 10 markets.
  - FinalDataSet2.csv : Combined data on zd and zs.
- For the rest of the midterm, I will work with MIDTERM1_FinalDataSet.csv.

In addition, to be able to make the merger analysis, we needed the average and minimum of the caffeine score of MD and TH. I computed then within STATA using the "egen" function. These are saved under the name "caffeine_scoreAM" and "caffeine_scoreAM2".

long

i	*j*	*stub*
1	1	4.1
1	2	4.5
2	1	3.3
2	2	3.0

$\xleftrightarrow{\text{reshape}}$

wide

i	*stub1*	*stub2*
1	4.1	4.5
2	3.3	3.0

Figure 2: Overview of the reshape function

1.2 Estimating BLP Using Instruments Provided in the Simulated Dataset

Here the utility model is :

$$U_{ij} = x_j\theta_i + \xi_j + \varepsilon_{ij},$$

where

$$\theta_i = \overline{\theta} + \Theta^o z_i + \Theta^u \nu_i.$$

Substituting the full specifications for the random coefficients, and collecting the terms that represent the mean utility of product j, we can rewrite the model as

$$U_{ij} = \delta_j + x_j\theta_2\nu_i + \varepsilon_{ij}, \tag{1}$$

where

$$\delta_j = x_j\theta_1 + \xi_j. \tag{2}$$

1.2.1 Demand

Parameter estimates

The parameters of the model are θ_1, θ_2 and ξ_j. Let $\theta = (\theta_1, \theta_2)$.

Our identifying assumption is:

$$E[\xi_j Z_j] = 0 \quad \forall j \quad \text{for some instruments } Z_j \quad \text{and the true unobservables } \xi_j.$$

Since we don't know the true unobservables, we will use some estimates $\hat{\xi}_t(\theta)$. Our estimator for the parameters will minimize the weighted sum of squares with the "optimal" weights :

$$\hat{\theta} = \arg\min_\theta \xi' Z(Z'Z)^{-1}Z'\xi \tag{3}$$

Solve for ξ

From equation 2 we have $\xi_j = \delta_j - x_j\theta_1$. We need to find δ.

Berry (1994) proves that there is a unique vector δ that maps the observed share from the data s_j to the theoretical share $s(\delta_j)$ $\forall j$. BLP gives us the contraction mapping that allows us to solve for this δ.

The market share implied by our model is the share of consumers who choose good j and it is an integral with no analytic solution. Thus, we need to simulate ns "individuals" (for each market t), each characterized byν_i drawn from the a normal distribution and use the simulator :

$$s_j(\delta, \theta; P^{ns}) = \sum_{r=1}^{ns} \frac{\exp\left(\delta_j + \sum_k x_{jk}\theta_{2k}\nu_{rk}\right)}{1 + \sum_s \exp\left(\delta_s + \sum_k x_{jk}\theta_{2k}\nu_{rk}\right)}. \tag{4}$$

Instruments Z

We need some instruments Z_j that are not with the ξ_j. Here we provided instruments.

Solving for θ_1, θ_2 and ξ_j

With Z and ξ in hand, we use our instruments forξ to identify θ_2, and then we use equation 3 to solve for θ_1.

Our x_j contains price and caffeine score. The resulting estimates provide us with the average price elasticity and the average caffeine score elasticity for demand.

Elasticities

We do the estimation for different for two starting value for θ_2, with three different number of draws (50, 500, and 5000) using the LBFGS, BFGS, GradientDescent, ConjugateGradient and NelderMead algorithm. We observe a convergence of our estimates regardless of the algorithm used or the initial value of θ_2 to the nearest decimal place.

		All Algorithms
θ_2 = 1 or 0	50 draws	θ_2 = -0.133 ; Price elasticity = -1.62, Caffeine score elasticity = 2.12
	500 draws	θ_2 = 0.085 ; Price elasticity =-1.61, Caffeine score elasticity = 2.11
	5000 draws	θ_2 = 0.0025 ; Price elasticity = -1.61, Caffeine score elasticity = 2.11

In the readme.file, we have a price elasticity price of -1.5. Our estimate is around -1.6, which reasonably close to the true value. Similarly, our estimate of caffeine score elasticity is around 2.1, close to the true value of 2.

1.2.2 Supply

Using the FOC of profit maximization($\frac{\partial \pi_f}{\partial p_f}$), we have :

$$s_j(p, x, \xi, \theta) + \sum_{r \in \mathcal{F}_f} (p_r - mc_r) \frac{\partial s_r(p, x, \xi; \theta)}{\partial p_j} = 0.$$

which can be rewritten in vector form :

$$0 = s + \Delta(p - mc)$$

where $\Delta$is a $J \times J$ matrix with

$$\Delta_{jr} = \begin{cases} -\frac{\partial s_r}{\partial p_j} & \exists f : r, j \in \mathcal{F}_f \\ 0 & \text{otherwise} \end{cases}.$$

Thus the vector of marginal costs for all products is

$$mc = \Delta^{-1} s + p$$

giving us the estimates :

		Cost elasticity of caffeine score
θ_2 = -0.133 ; θ_1 = [-1.62,2.12]	50 draws	$\theta_3 = 0.024$
	500 draws	θ_3 =0.024
	5000 draws	θ_3 =0.024
θ_2 = 0.085 ; θ_1 = [-1.61,2.11]	50 draws	$\theta_3 = 0.034$
	500 draws	θ_3 =0.034
	5000 draws	θ_3 =0.034
θ_2 = 0.0025; θ_1 = [-1.61,2.11]	50 draws	$\theta_3 = 0.039$
	500 draws	θ_3 =0.039
	5000 draws	θ_3 =0.039

- We had to drop negative marginal cost. We only had 6 of them in all cases.
- We used 2SLS.

In the readme.file, we have a cost elasticity of caffeine score of 0.1. Our estimate is around 0.02-0.03. Unfortunately, this value is not close to the true value.

1.3 Estimating BLP Using Instruments Generated by the function BLP_-instruments

BLP derive a slightly more explicit set of instrumental variables. They use the sums of the values of the same characteristics of products offered by other firms. The function BLP_-instruments creates those instruments. We only present the result for 50 draws with the NelderMead algorithm, but with different starting value of θ_2.

1.3.1 Demand

	NelderMead Algorithms
θ_2 =1	θ_2 = 1.0 ; Price elasticity = -1.85, Caffeine score elasticity = 2.12
θ_2 =0	θ_2 = 0.025 ; Price elasticity = -1.61, Caffeine score elasticity = 2.12

1.3.2 Supply

- We had 63 negative marginal cost values.

	Cost elasticity of caffeine score
θ_2=1.0, θ_1 = [-1.85,2.12]	θ_3 =-0.546
θ_2=0.025, θ_1 = [-1.61,2.12]	θ_3 = 0.04

The value of the cost elasticity of caffeine score is very sensitive to the starting value of θ_2. For θ_2 =0, we find similar value as our baseline. However, for θ_2 =1, we have a change in the sign of the elasticity. This might be explain by the fact that we are not using the optimal instrument.

2 Merger Analysis

Now that we have marginal costs and a model of demand, we can predict the impact of changes1 such as mergers.

We first solve for baseline price [Baseline : Model estimate with optimal instrument in part 1.2]. We use the results from the 50 draws. Once we got them, we computed the baseline share. Then, we tested if these values are different from the observed one. For both test, we cannot reject the null hypothesis that the simulated data are different from the observed one at the 5% level.

Efficiency scenario

We assume these changes affect marginal costs or demand parameters : The new caffeine score cs of MD + TH coffee will be $cs_{MDTH} = min(cs_{MD}, cs_{TH})$ and the new unobserved cost shifter will be also the min of the original firms/products. So marginal cost of MD +TH is lower than the originals.

Average scenario

We assume these changes affect marginal costs or demand parameters : The new caffeine score cs of MD + TH coffee will be $cs_{MDTH} = \frac{cs_{MD}+cs_{TH}}{2}$ and the new unobserved cost shifter will be also the average of the original firms/products. So marginal cost of MD + TH is in the middle of the original ones.

Analysis for the firms

Once we obtain the new price (under both scenario), we computed an hypothesis test to compare the new price to baseline for MD product and TH product produce by our new firm. We find p-value of 0.203 for MD and 0.999 for TH in the efficient case, and 0.149 for MD and 0.999 for TH in the average scenario. Thus, we cannot conclude that new prices for MD and TH are statistically higher than their price in the baseline(Similar result for the shares). The merger seems to not affect the market price of these two product even though it affect their marginal cost. This might be because the substitution effects (between MD, TH and SB) are so small between the products that there that the "gains from collusion" are not very high.

We compare the two merger in terms of price. For MD product, we cannot conclude that the two new price resulting from the two merger scenario are statistically different. However, for TH, we find that the two merger scenario result in different prices. For the share, we find no difference between post and after merge for each firm when we compare the two cases.

Considering these findings, it appears that both merger scenarios have minimal impact on prices and market shares. However, there is a difference in the prices resulting from the two merger scenarios for the TH product. Therefore, if one were to prefer a merger based solely on price considerations, the Average Scenario might be slightly more favorable as it results in different prices for the TH product, potentially offering some flexibility in pricing strategies.

Analysis for the consumers

The Average Scenario might be slightly more favorable for producing equality for consumers due to the potential for different prices for the TH product, providing more options and potentially better deals.