Econ 8400, Fall 2022
Problem Set 3 (Due October 13)

On the course web site you will find a simulated data set in the file PS3_data.csv. The variables are labeled in the first row and include: a market id (ranging from 1-5), period (1-30), product id (1-20), market share (note that the shares do not sum up to 1 because of the outside good), calories (a continuous characteristic), organic (a discrete characteristic), and price. Overall, there are 1,431 observations. Note that not all products are available in all markets and time periods. The products are produced by single product firms.

1. Provide descriptive statistics that describe the key variables.

2. Estimate the following (Logit) model:

$$u_{ijt} = \alpha p_{jt} + x_{jt}\beta + \xi_{jt} + \epsilon_{ijt} \quad i = 1, ..., I \quad j = 1, ..., J \quad t = 1, ..., T$$

   The notation follows that used in class (note that t indexes both markets and periods, and therefore $T = 150$). For characteristics use "calories" and "organic" (and well as a constant). Assume that current price is correlated with the error term. Estimate the model using OLS and IV. For IVs use (1) Hausman IVs (i.e., the avg price of the product in other markets in the same period); and (2) BLP-like instruments (number of competing products, average calories of competing products, number of organic products, number of organic products interacted with the organic dummy and closest competitor in "calories space", i.e. $min_{k \neq j} |calories_j - calories_k|$). Remember to always include the own characteristics in the IV set. For each specification report the estimated coefficients and their standard errors.

3. Estimate the following (RC Logit) model:

$$u_{ijt} = \alpha_i p_{jt} + x_{jt}\beta_i + \xi_{jt} + \epsilon_{ijt} \quad i = 1, ..., I \quad j = 1, ..., J \quad t = 1, ..., T$$

   where α_i is normally distributed with mean α and standard deviation σ_α, and β_i is constant for organic and normally distributed with mean β^{cal} and standard deviation σ_β. Estimate the model using GMM based on the same IVs you used in the previous question For the starting guess for δ_{jt} use $log(s_{jt}) - log(s_{0t})$. For the starting value for the nonlinear parameters use $\sigma_\alpha = \sigma_\beta = 1$. For the random draws use the data in PS3_rnd.csv. These are draws from standard normal distributions (using these instead of drawing your own will making it easier to check the results). Note: you can write your own code or use codes that are available on the web (most notably the code written by Conlon and Gortmaker). You will probably learn more by writing your own code, but learning how to work with the Conlon-Gortmaker code is also valuable.

4. Using the results for the IV Logit specification and the GMM results compute the own- and cross-price elasticities. Compute these for each t (market-period) and report the median across t. To save space you can report these for a subset of the products. Are there any particular patterns in the elasticities both within model (i.e., Logit and RC) and across the models? Do these make sense? How would you need to generalize the model for the elasticities to make more sense?

5. Using these results compute the markups predicted by a (i) single-product Nash-Bertrand and (ii) joint pricing of all the products. For each (pricing and demand) model report the median of the distribution of the markups and margins. Are there any particular patterns in these numbers?