Problem Set 3

Hyoungchul Kim*

October 13, 2025

Contents

*I acknowledge helpful discussions with Hyunjun, Inkoo, Taka, and David.

1

This section presents the descriptive statistics of the key variables used throughout the analysis. The summary statistics provide an overview of the main data features, such as the mean, standard deviation, and range of each variable. These descriptive measures are essential for understanding the variation in the dataset and for anticipating potential identification issues or estimation challenges that might arise later in the empirical sections.

	N	mean	sd	min	max
market_share	1431	0.06	0.07	0.00	0.44
calories	1431	1.87	0.19	1.55	2.16
organic	1431	0.35	0.48	0.00	1.00
price	1431	1.56	0.38	0.59	2.57

2

Next, I estimate a series of models, including the standard logit model estimated by OLS, the Hausman IV model, and the Berry-Levinsohn-Pakes (BLP) random coefficients model. The table below summarizes the coefficient estimates obtained from each specification.

model	term	estimate	std.error
OLS	price	-2.972	0.122
OLS	calories	-1.697	0.460
OLS	organic	1.102	0.205
hausman	fit_price	-1.244	0.085
blp	fit_price	-0.250	0.102

The results appear to be largely sensible. In particular, the price coefficient has the expected negative sign, indicating that demand decreases with higher prices. However, in the case of the BLP specification, the price coefficient is statistically insignificant. This suggests the possibility of a weak instrumental variable (IV) problem, meaning that the instruments may not be providing sufficient variation in prices to achieve consistent estimation. This weakness becomes more apparent and consequential in subsequent sections of the analysis, where we explore model fit and elasticity patterns.

3

In this section, I compare the results from my own BLP implementation to those obtained using the `PyBLP` package. For my own implementation, I constructed BLP-like instruments and estimated the model accordingly. Interestingly, the price coefficient turned out to be positive (1.92), which is contrary to economic intuition. Such a result likely arises due to the weak IV issue mentioned earlier.

model	term	estimate	std.error
blp	price	1.979	1.816
blp	β^{org}	-2.215	1.601
blp	β^{cal}	-2.517	1.176
blp	σ_α	-0.012	10.07
blp	σ_β	1.076	0.876

To further validate this observation, I also estimated the model using `PyBLP`, which provides a reliable and standardized implementation of the BLP framework. The results from `PyBLP` are presented below:

As shown, the estimated price coefficient remains positive. This further reinforces the idea that the BLP-like instruments used in this setup do not generate sufficient variation in endogenous prices. Consequently, both my implementation and the `PyBLP` estimation exhibit similar weaknesses due to the limited strength of the chosen instruments.

4

For the IV logit specification, recall that the own-price and cross-price elasticities are given by the following expressions:

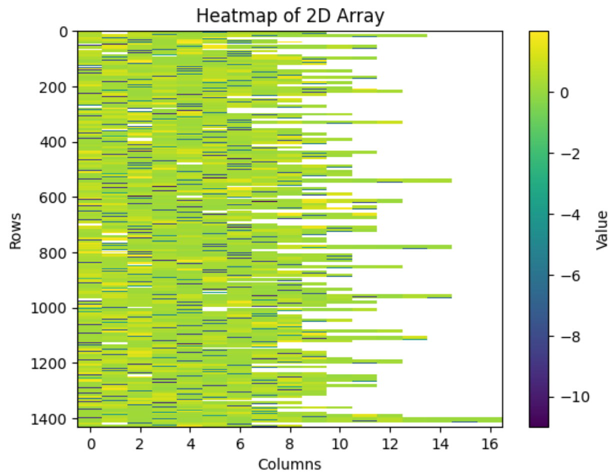
$$\begin{aligned} \eta_{jt} &= \alpha p_{jt}(1 - s_{jt}) \\ \eta_{jkt} &= -\alpha p_{kt} s_{kt} \end{aligned} \tag{1}$$

Using these formulas, I calculate and report portions of the price elasticities below. For the BLP random coefficients model, I rely on `PyBLP`'s built-in function to compute the elasticities. The resulting heatmap of price elasticities is displayed below.

You can clearly see that something is problematic in the estimated elasticity matrix. Several values appear to be missing, and the color gradient of the heatmap is inconsistent with theoretical expectations. This inconsistency likely stems from the irregular or implausible parameter estimates obtained in the preceding GMM estimation. Therefore, the derived elasticity estimates should be interpreted with caution, as they may not accurately reflect realistic substitution patterns across products.

5

Given the instability in the BLP results, I proceed by focusing on the Hausman IV model, which tends to provide more stable and interpretable estimates in this dataset. Using the estimated parameters, I compute the implied marginal costs and markups under the assumption of single-product Nash-Bertrand competition. The resulting distributions of marginal cost and markups are displayed in the following figures.



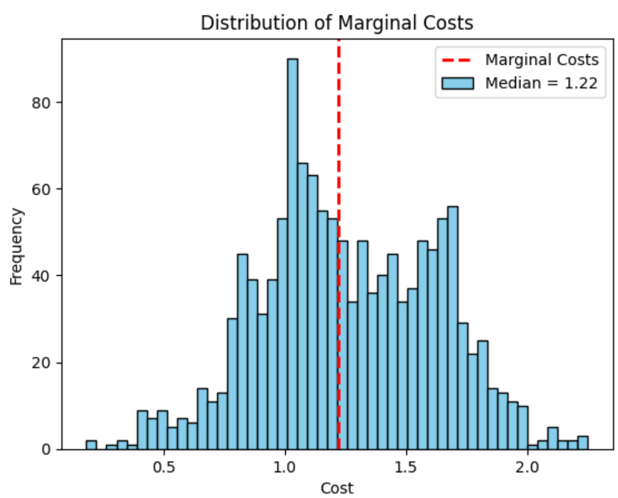


Figure 1: Marginal cost

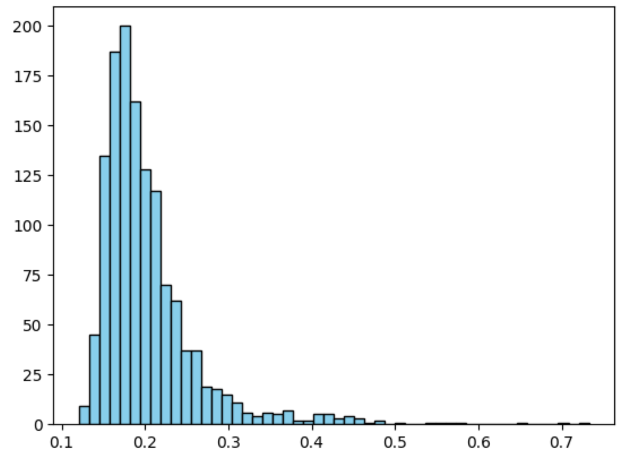


Figure 2: Markups

In the single-product case, the estimated markups appear relatively modest. Most of the markups cluster close to zero, with a median value around 0.2. This result is consistent with the notion that in markets where firms compete independently and do not internalize cross-product interactions, prices tend to be close to marginal cost.

By contrast, when we instead assume joint pricing across all products—similar to a scenario of full coordination or collusive pricing behavior—the markups change substantially. The results for this case are shown below.

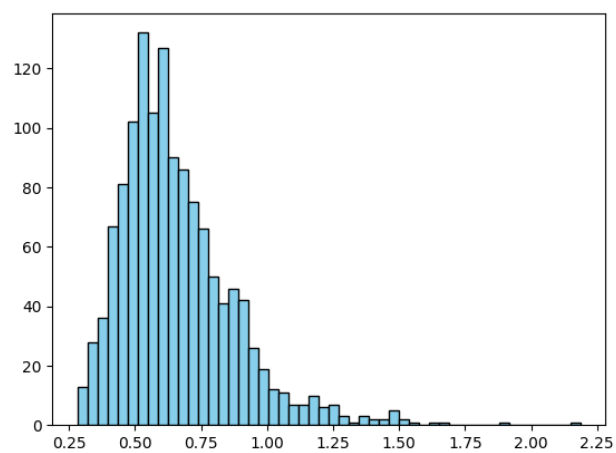


Figure 3: Markups

Under the joint-pricing assumption, the magnitude of markups increases considerably. The median markup rises to approximately 0.6, indicating a stronger degree of market power when firms set prices jointly across their product portfolios. This comparison illustrates the substantial effect that pricing assumptions can have on inferred market conduct and profitability. It also highlights how different competitive structures—ranging from independent pricing to coordinated behavior—can lead to markedly different equilibrium outcomes.