Problem Set 1

Richard Chen and Pryce Davies

Winter 2026

Exercise 1

1. Let $X_i,\ i = 1, \ldots, N$ be a sequence of independent type 1 extreme value random variables with location parameter μ_i and scale parameter $\sigma > 0$. The c.d.f is given by:

$$F_{\mu_i,\sigma}(x) \equiv \mathbb{P}(X_i \leq x | u_i, \sigma) = \exp(-\exp(-\frac{x - \mu_i}{\sigma}))$$

Derive the distribution of $Y = \max_i\{X_i\}$

The probability that the maximum of X_i is less than or equal to x is equal to the probability that all N RVs are less than x. Since the X_i's are independent and $X_i \sim T1EV(\mu_i, \sigma)$,

$$\begin{aligned}
\mathbb{P}(\max_i\{X_i\} \leq x) &= \prod_i^N \mathbb{P}(X_i \leq x) \\
&= \prod_i^N (\exp(-\exp(-\frac{x - \mu_i}{\sigma})) \\
&= \exp(-\sum_i^N \exp(-\frac{x}{\sigma} + \frac{\mu_i}{\sigma})) \\
&= \exp(-\exp(-\frac{x}{\sigma}) \sum_i^N \exp(\frac{\mu_i}{\sigma})) \\
&= \exp(-\exp(-\frac{x}{\sigma}) \exp(\frac{\mu}{\sigma})) \quad \text{where} \quad \mu \equiv \sigma \ln(\sum_1^N \exp(\frac{\mu_i}{\sigma})) \\
&= \exp(-\exp(-\frac{x - \mu}{\sigma})) \\
\implies Y \sim T1EV(\mu, \sigma)
\end{aligned}$$

2. Let X and Y be two independent T1EV random variables with location parameters μ_x and μ_y respectively and common scale parameter $\sigma > 0$. Derive the distribution of $X - Y$.

We want to derive $\mathbb{P}(X - Y \leq \epsilon) \iff \mathbb{P}(X \leq Y + \epsilon)$, which, since X and Y are independent, is given by:

$$\int_{-\infty}^{\infty} \left[\int_{-\infty}^{y+\epsilon} f_{\mu_x,\sigma}(x) dx \right] f_{\mu_y,\sigma} dy$$

Where $f_{\mu,\sigma}$ is the pdf of a $T1EV(\sigma, \mu)$ distribution. Then,

$$f_{\mu,\sigma}(x) = \frac{1}{\sigma} \exp(-\frac{x - \mu}{\sigma}) \exp(-\exp(-\frac{x - \mu_i}{\sigma})) = \frac{1}{\sigma} \exp(-\frac{x - \mu}{\sigma}) F_{\mu,\sigma}(x)$$

and so

$$
\begin{aligned}
\mathbb{P}(X - Y \leq \epsilon) &= \int_{-\infty}^{\infty} \left[F_{\mu_x,\sigma}(y+\epsilon)\right] f_{\mu_y,\sigma} dy \\
&= \int_{-\infty}^{\infty} F_{\mu_x,\sigma}(y+\epsilon) \frac{1}{\sigma} \exp(-\frac{y-\mu_y}{\sigma}) F_{\mu_y,\sigma}(y) dy \\
&= \frac{1}{\sigma} \int_{-\infty}^{\infty} \exp(-\exp(-\frac{y+\epsilon-\mu_x}{\sigma})) \exp(-\frac{y-\mu_y}{\sigma}) \exp(-\exp(-\frac{y-\mu_y}{\sigma})) dy \\
&= \frac{1}{\sigma} \int_{-\infty}^{\infty} \exp\left[-\exp(-\frac{y-\mu_y}{\sigma}) - \exp(-\frac{y-\mu_y+\epsilon+\mu_y-\mu_x}{\sigma})\right] \exp(-\frac{y-\mu_y}{\sigma}) dy \\
&= \frac{1}{\sigma} \int_{-\infty}^{\infty} \exp\left[-\exp(-\frac{y-\mu_y}{\sigma}) \underbrace{\left(1+\exp(\frac{-(\epsilon+\mu_y-\mu_x)}{\sigma})\right)}_{\equiv a}\right] \exp(-\frac{y-\mu_y}{\sigma}) dy \\
&= \frac{1}{a} \int_{-\infty}^{\infty} \underbrace{\frac{a}{\sigma} \exp(-\frac{y-\mu_y}{\sigma}) \exp\left[-\exp(-\frac{y-\mu_y}{\sigma}) a\right]}_{=f_{\mu_y,\frac{\sigma}{a}}(x)} dy \\
&= \frac{1}{a} = \frac{1}{1+\exp(\frac{-(\epsilon+\mu_y-\mu_x)}{\sigma})} \\
&= \frac{1}{1+\exp(\frac{-(\epsilon-(\mu_x-\mu_y))}{\sigma})}
\end{aligned}
$$

which is the CDF for the logistic distribution function with mean $\mu_x - \mu_y$ and scale parameter σ. Therefore $X - Y \sim Logistic(\mu_x - \mu_y, \sigma)$

3. Consider an individual who has to choose one product among N possible alternatives. The utility derived from alternative j is given by:

$$u_j = \mu_j + \epsilon_j$$

where μ_j is non-random and $\epsilon_j \sim i.i.d. T1EV(0,1)$. Derive the probability that alternative j is chosen.

Note that for $x \sim T1EV(0,1)$ and scalar y, $y + x \sim T1EV(y,1)$. Therefore, $u_j \sim T1EV(\mu_j, 1)$. The likelihood that alternative j is chosen is:

$$
\begin{aligned}
\mathbb{P}(\text{i choose j}) &= \mathbb{P}(u_j \geq u_k, \forall k \neq j) \\
&= \mathbb{P}(u_j \geq \underbrace{\max_k u_k}_{\equiv u_{-j}})
\end{aligned}
$$

Note that from part 1, we know that $u_{-j} \sim T1EV(\mu_{-j})$ where $\mu_{-j} \equiv \ln(\sum_{k\neq j} \exp(\mu_k))$. Then, using this fact and the distribution of the difference of two T1EV distributed RVs derived in part 2,

$$
\begin{aligned}
\mathbb{P}(\text{choose j}) &= \mathbb{P}(u_{-j} - u_j \leq 0) \\
&= \frac{1}{1+\exp(\mu_{-j} - \mu_j)} \\
&= \frac{\exp(\mu_j)}{\exp(\mu_j) + \exp(\mu_{-j})} \\
&= \frac{\exp(\mu_j)}{\sum_{k=1}^{N} \exp(\mu_k)}
\end{aligned}
$$

4. Consider a market with J products indexed by $j = 1, \dots, J$, an outside good denoted by $j = 0$ and a large number of consumers indexed by $i \in \mathcal{I}$, each of whom only buys one of the products. Consumer i's indirect utility from consuming product j is given by:

$$u_{ij} = \alpha(y_i - p_j) + \epsilon_{ij} \quad \text{for } j = i, \dots, J$$

$$u_{i0} = \alpha y_i + \epsilon_{i0}$$

where p_j is the price of product j, y_i is consumer i's income, and ϵ_{ij} is an idiosyncratic taste shock that makes products horizontally differentiated.

(a) Assume ϵ_{ij} are i.i.d T1EV(0,1). Denote consumer i's individual choice probability of selecting product j as $s_j(i)$. Derive $s_j(i)$ and compute $\frac{\partial s_j(i)}{\partial y_i}$. Interpret your results.

The probability of selecting product j, $s_j(i)$, is given by

$$\begin{aligned} s_j(i) &= \mathbb{P}(u_{ij} > \max_{0,k}\{u_{ik}\}) \\ &= \frac{\exp(\alpha(y_i - p_j))}{\sum_0^J \exp(\alpha(y_i - p_k))} \quad \text{by exercise 1.3} \\ &= \frac{\exp(\alpha y_i)\exp(-\alpha p_j)}{\exp(\alpha y_i)\sum_0^J \exp(-\alpha p_k)} = \frac{\exp(-\alpha p_j)}{\sum_0^J \exp(-\alpha p_k)} \\ &= \frac{\exp(-\alpha p_j)}{1 + \sum_1^J \exp(-\alpha p_k)} \end{aligned}$$

Then, $\frac{\partial s_j(i)}{\partial y_i} = 0$. Intuitively, the consumer's demand is a function of the difference between their income and the price of the good and an orthogonal T1EV taste shock. If their income changes, then the ranking of the goods by the distance between the good's price and the consumer's income is unchanged. Therefore, the probability of select a good is unchanged, given that their taste shock is orthogonal to any change in income.

(b) Assume ϵ_{ij} are i.i.d T1EV(0,1). Derive s_j (the market share of product j) and compute own and cross-price elasticities. Are the latter reasonable? Explain.

With many consumer's and ϵ_{ij} i.i.d, then the market share of good j is well approximated by an individuals probability of selecting the product, $s_j(i)$. Therefore,

$$s_j = \frac{\exp(-\alpha p_j)}{\sum_0^J \exp(-\alpha p_k)}$$

Then,

$$\begin{aligned}
\frac{\partial s_j}{\partial p_j} &= \frac{-\sum_0^J \exp(-\alpha p_k)\alpha \exp(-\alpha p_j) + \exp(-\alpha p_j)\alpha \exp(-\alpha p_j)}{(\sum_0^J \exp(-\alpha p_k))^2} \\
&= \frac{-\alpha \exp(-\alpha p_j)(\sum_0^J \exp(-\alpha p_k) - \exp(-\alpha p_j))}{(\sum_0^J \exp(-\alpha p_k))^2} \\
\frac{\partial s_j}{\partial p_j}\frac{p_j}{s_j} &= p_j \frac{-\alpha \exp(-\alpha p_j)(\sum_0^J \exp(-\alpha p_k) - \exp(-\alpha p_j))}{(\sum_0^J \exp(-\alpha p_k))^2} \frac{\sum_0^J \exp(-\alpha p_k)}{\exp(-\alpha p_j)} \\
&= -\alpha p_j (1 - s_j)
\end{aligned}$$

Likewise,

$$\begin{aligned}
\frac{\partial s_j}{\partial p_k} &= \frac{-\alpha \exp(-\alpha p_k)\exp(-\alpha p_j)}{(\sum_0^J \exp(-\alpha p_k))^2} \\
\frac{\partial s_j}{\partial p_k}\frac{p_k}{s_j} &= p_k \frac{-\alpha \exp(-\alpha p_k)\exp(-\alpha p_j)}{(\sum_0^J \exp(-\alpha p_k))^2} \frac{\sum_0^J \exp(-\alpha p_k)}{\exp(-\alpha p_j)} \\
&= \alpha p_k s_k
\end{aligned}$$

We have shown that cross-price elasticities only depend on the price and the market share of the good for which the price is being changed. This means that the percent increase in market share for any other goods r and j should be the same. This makes sense only if price and the i.i.d taste shock are the only determinants of demand, as we would expect differential changes in market share across goods if, for example, good r is a closer substitute to good k than good j based on characteristics. Notice that this expression for cross-price elasticity means that goods with a higher market share (and therefore a lower price) will have a greater absolute increase in market share than goods with a lower market share (and therefore a higher price), which is intuitive as consumers are more likely to switch over to lower cost goods in those model.

(c) Assume that $\epsilon_{ij} = \beta_i x_j$ where x_j represents a non-random product characteristic that consumers value, and β_i represents an idiosyncratic taste shock for the same characteristic. Moreover, assume that $x_j > 0$, $x_0 = 0$.

i Assume that $\beta_i = \beta$ for all i. Derive product j market share, s_j. Interpret your results.

A consumer's utility is now $u_{ij} = \alpha(y_i - p_j) + \beta x_j$ and $u_{i0} = \alpha y_i$. Then, consumer's will choose the product $j^* \in 1, .., J$ for which u_{ij} is maximized. Since αy_i is common across products, this reduces to $j^* = \arg\max_{k \in 0,...,J} \beta x_k - \alpha p_k$. Since there is no heterogeneity in preferences, all consumers will choose this product j^*. Therefore, $s_{j^*} = 1$ and $s_k = 0\ \forall k \neq j^*$. Intuitively, since all consumer's have the same preferences, they choose the same good, which happens to be the good for which $\beta x_k - \alpha p_k$ is largest (most bang per buck).

ii Assume that β_i are i.i.d Uniform$[0, \bar{B}]$ with $\bar{B}$ sufficiently large. Derive product j market share, s_j, and compute own and cross-price elasticities. Are the latter reasonable? Explain and compare with your findings in points (b) above. (For simplicity assume that $\frac{p_i - p_j}{x_i - x_j} \geq \frac{p_j - p_k}{x_j - x_k}$ whenever $x_i \geq x_j \geq x_k$).

Given that β_i is now a non-degenerate random variable, there will be variation in consumers' choice of products. Suppose that we can rank the x_j's as follows:

$$x_J \geq x_{J-1} \geq \cdots \geq x_1$$

And suppose that consumer i is indifferent between good j and $j-1$, then

$$\beta_i x_j - \alpha p_j = \beta_i x_{j-1} - \alpha p_{j-1} \implies \beta_i = \alpha \frac{(p_j - p_{j-1})}{(x_j - x_{j-1})} \equiv c_j$$

Notice that the ordered ratio assumption given implies that:

$$c_1 \leq \cdots \leq c_j \leq c_{j+1} \leq \cdots \leq c_J$$

Suppose $\beta_i \in (c_j, c_{j+1})$. This implies that

$$\beta_i < \alpha \frac{p_{j+1} - p_j}{x_{j+1} - x_j} \implies \beta_i x_{j+1} - \alpha p_{j+1} < \beta_i x_j - \alpha p_j \implies u_{ij+1} < u_{ij}$$

and

$$\beta_i > \alpha \frac{p_j - p_{j-1}}{x_j - x_{j-1}} \implies u_{ij} > u_{ij-1}$$

Now to show that product j is chosen by i, we want to show that $\beta_i \in (c_j, c_{j+1})$ also implies that $u_{ij} > u_{ik}$ for all k other than $j-1, j+1$. To do so, we will invoke an inductive argument. For example, suppose $u_{ij} \geq u_{ij+2}$, then

$$u_{ij} \geq u_{ij+2} \implies \beta_i(x_j - \alpha p_j) \geq \beta_i(x_{j+2} - \alpha p_{j+2}) \implies \beta_i \leq \alpha \frac{p_{j+2} - p_j}{x_{j+2} - x_j}$$

Then,

$$\alpha \frac{p_{j+2} - p_j}{x_{j+2} - x_j} = \underbrace{\alpha \frac{p_{j+2} - p_{j+1}}{x_{j+2} - x_{j+1}}}_{> c_{j+1}} \frac{x_{j+2} - x_{j+1}}{x_{j+2} - x_j} + \underbrace{\alpha \frac{p_{j+1} - p_j}{x_{j+1} - x_j}}_{= c_{j+1}} \frac{x_{j+1} - x_j}{x_{j+2} - x_j}$$

where $\frac{x_{j+1} - x_j}{x_{j+2} - x_j} + \frac{x_{j+2} - x_{j+1}}{x_{j+2} - x_j} = 1$. Therefore $\beta_i < \alpha \frac{p_{j+1} - p_j}{x_{j+1} - x_j} \implies \beta_i < \alpha \frac{p_{j+2} - p_j}{x_{j+2} - x_j} \implies u_{ij} > u_{ij+2}$. By induction, we can extend this result to all neighbors on either side of $j-1$ and $j+1$. We have therefore shown that $u_{ij} \geq \max_k u_{ik}$.

If we assume that $0 \leq c_1 \leq \cdots \leq c_J \leq \bar{B}$, then since β_i are i.i.d Uniform$[0, \bar{B}]$ and there are many consumers,

$$s_j = \mathbb{P}(\beta_i \in [c_j, c_{j+1}]) = \frac{c_{j+1} - c_j}{\bar{B}} \tag{1}$$

$$= \frac{\alpha}{\bar{B}} \left(\frac{p_{j+1} - p_j}{x_{j+1} - x_j} - \frac{p_j - p_{j-1}}{x_j - x_{j-1}} \right) \tag{2}$$

where we have $\cdots \geq x_{j+1} \geq x_j \geq x_{j-1} \geq \ldots$

Then:

$$\frac{\partial s_j}{\partial p_j} = -\frac{\alpha}{\bar{B}}\left(\frac{1}{x_{j+1}-x_j} + \frac{1}{x_j - x_{j-1}}\right)$$
$$\frac{\partial s_j}{\partial p_j}\frac{p_j}{s_j} = -p_j\left(\frac{x_{j+1}-x_{j-1}}{\underbrace{(p_{j+1}-p_j)(x_j-x_{j-1})-(p_j-p_{j-1})(x_{j+1}-x_j)}_{>0 \text{ by assumption}}}\right) < 0$$

and:

$$\frac{\partial s_j}{\partial p_k} = 0 \quad \forall k \neq j-1, j+1$$
$$\frac{\partial s_j}{\partial p_{j-1}} = \frac{\alpha}{\bar{B}}\left(\frac{1}{x_j - x_{j-1}}\right)$$
$$\frac{\partial s_j}{\partial p_{j-1}}\frac{p_{j-1}}{s_j} = p_{j-1}\left(\frac{x_{j+1}-x_j}{(p_{j+1}-p_j)(x_j-x_{j-1})-(p_j-p_{j-1})(x_{j+1}-x_j)}\right) > 0$$

Finally,

$$\frac{\partial s_j}{\partial p_{j+1}} = \frac{\alpha}{\bar{B}}\left(\frac{1}{x_{j+1} - x_j}\right)$$
$$\frac{\partial s_j}{\partial p_{j+1}}\frac{p_{j+1}}{s_j} = p_{j+1}\left(\frac{x_j-x_{j-1}}{(p_{j+1}-p_j)(x_j-x_{j-1})-(p_j-p_{j-1})(x_{j+1}-x_j)}\right) > 0$$

In this demand system, due to our assumption about the ordering of $\frac{p_i-p_j}{x_i-x_j}$, ones demand for a product depends only on the relative prices of the goods with x_k one rank above and below x_j (which we call goods $j+1$,$j-1$). Therefore, the cross-price elasticities for all other goods with $k \neq j-1, j+1$ are 0. We can think of these goods as being unrelated to the relative good j. As expected, the own-price elasticity is negative, but because of the structure of the taste shock, it depends on the prices and characteristics of other goods $j+1$, $j-1$, unlike in part 4.b. Intuitively, this is because a change in the price of good j will change both end points c_j and c_{j+1}, causing some consumers to change over to good $j-1$ or good $j+1$, with the degree of substitution from good j to good $j+1$ $(j-1)$ depending on the relative similarity of the prices and characteristics of j and $j-1$ $(j+1)$. Finally, since goods $j-1$ and $j+1$ are the (only) substitutes for good j, the demand elasticity with respect to these goods is positive. Intuitively, they are substitutes as they have x_k that is close to x_j, so a small change in the thresholds that determine the product share of good j, c_j, c_{j+1}, could result in some consumers of good j switching to good $j-1$ or good $j+1$. We see that the magnitude of the cross-price elasticities with respect to good $j+1$ and $j-1$ are *both* a function of the distance between the characteristic and price of good j and $j+1$ and good j and $j-1$. Intuitively, this is because an increase in price of good $j+1$ will raise end point c_{j+1} and therefore increase s_j. The percent increase in s_j depends on the existing distance between intervals $c_{j+1} - c_j = \alpha(\frac{p_{j+1}-p_j}{x_{j+1}-x_j} - \frac{p_j-p_{j-1}}{x_j-x_{j-1}})$, causing the resulting demand elasticity to depend implicitly on p_{j-1}, x_{j-1}. Therefore, the relative degree of substitution from good j to good $j+1$ $(j-1)$ depends on the relative similarity of the prices and characteristics of j and $j-1$ $(j+1)$.

(d) Assume that $\epsilon_{ij} = \beta_i x_j + v_{ij}$ where x_j represents a non-random product characteristic, β_i represents an idiosyncratic taste shock for that same characteristic and v_{ij} are i.i.d $T1EV(0,1)$. Moreover, assume that β_i

are i.i.d with generic c.d.f $F(.)$. Derive product j's market share and compute own and cross-price elasticities. Explain and compare with your findings in point (b) above.

The utility function for consumer i and product j is

$$u_{ij} = \beta_i x_j - \alpha p_j + \nu_{ij}$$

Because ν_{ij} is T1EV iid, for a given individual with characteristic taste β_i, the probability of them choosing good j is

$$P(\text{i choose j}) = \frac{\exp(\beta_i x_j - \alpha p_j)}{1 + \sum \exp(\beta_i x_k - \alpha p_k)}$$

To get the product share, we integrate over the support of β_i.

$$E[s_{ij}] = s_j = \int_{-\infty}^{\infty} \frac{\exp(\beta_i x_j - \alpha p_j)}{1 + \sum \exp(\beta_i x_k - \alpha p_k)} f(\beta_i) d\beta_i$$

The own price elasticity can be calculated as follows using the Leibniz integral rule.

$$\begin{aligned}
\frac{\partial s_j}{\partial p_j} &= \int_{\infty}^{\infty} \frac{(1 + \sum \exp(\beta_i x_k - \alpha p_k)(-\alpha \exp(\beta_i x_j - \alpha p_j)) - (\exp(\beta_i x_j - \alpha p_j))(-\alpha \exp(\beta_i x_j - \alpha p_j))}{(1 + \sum \exp(\beta_i x_k - \alpha p_k))^2} f(\beta_i) d\beta_i \\
&= \int_{-\infty}^{\infty} \frac{(-\alpha \exp(\beta_i x_j - \alpha p_j))(1 + \sum \exp(\beta_i x_k - \alpha p_k) - \exp(\beta_i x_j - \alpha p_j))}{1 + \sum \exp(\beta_i x_k - \alpha p_k))^2} dF(\beta_i) \\
&= -\alpha \int_{-\infty}^{\infty} s_{ij}(\beta) \frac{1 + \sum \exp(\beta_i x_k - \alpha p_k) - \exp(\beta_i x_j - \alpha p_j))}{1 + \sum \exp(\beta_i x_k - \alpha p_k)} dF(\beta_i) \\
&= -\alpha \int_{-\infty}^{\infty} s_{ij}(1 - s_{ij}) dF(\beta_i)
\end{aligned}$$

and therefore

$$\frac{\partial s_j}{\partial p_j} \frac{p_j}{s_j} = -\alpha \frac{p_j}{s_j} \int_{-\infty}^{\infty} s_{ij}(1 - s_{ij}) dF(\beta_i)$$

We can then do something similar to calculate the cross-price elasticity:

$$\begin{aligned}
\frac{\partial s_j}{\partial p_k} &= \int_{-\infty}^{\infty} \alpha \frac{\exp(\beta_i x_j - \alpha p_j)}{(1 + \sum \exp(\beta_i x_k - \alpha p_k))^2} \exp(\beta_i x_k - \alpha p_k) f(\beta_i) d\beta_i \\
&= \alpha \int_{-\infty}^{\infty} s_{ij} s_{ik} dF(\beta_i) \implies \\
\frac{\partial s_j}{\partial p_k} \frac{p_k}{s_j} &= \alpha \frac{p_j}{s_j} \int_{-\infty}^{\infty} s_{ij} s_{ik} dF(\beta_i)
\end{aligned}$$

Now, the elasticity of good j with respect to price k depends on the characteristics of good j, whereas in part (b) it only depended on the characteristics of good k. This property is closely connected to the IIA assumption, as it shows that the substitution away from good k to good j does not depend on the characteristics of good j. We can see that with randomly distributed β_i, we have relaxed the IIA assumption as the substitution from good k to good j depends both on the characteristics of good k and good j. Specifically, since $s_{ik} \in [0, 1]$, the cross-price elasticity of good j with respect to good k will be maximized when $s_{ik} = s_{ij}$. Intuively, this means that goods that are more similar (have more similar values of s_i) are stronger substitutes, a desirable property that IIA rules out. The reason IIA no longer holds is because by allowing β_i to vary over the population, the common shock ϵ_{ij} is no longer the only determinant of which product each consumer chooses. Rather, we allow for different types of individuals who systematically prefer different types of goods, resulting in heterogeneous substitution patterns.

(e) Assume, as in point (a) above, that ϵ_{ij} are i.i.d $T1EV(0,1)$. Moreover, suppose we want to measure welfare at given prices $(p_1, \ldots, p_J)$ as

$$W \equiv \mathbb{E}[\max_{j=0,\ldots,J} u_{ij}]$$

i Rewrite W as a function of the market share of the outside option s_0

Welfare is measured by the expected utility of the consumer from choosing their preferred product.Using that $u_{ij} = \alpha(y_i - p_j) + \epsilon_{ij}$ for the first equality and following note 27 on page 69 of the textbook to get the second equality, we have that

$$\begin{aligned} W &\equiv \mathbb{E}[\max_{j=0,\ldots,J} -\alpha p_j + \epsilon_{ij}] \qquad \epsilon_{ij} \sim T1EV(0,1) \\ &= \ln\left[\sum_{j=1}^{J} \exp(-\alpha p_j))\right] + \underbrace{C}_{\approx..5772} \\ &= \ln\left[\frac{1}{s_0} - 1\right] + C = \ln(1 - s_0) - \ln(s_0) + C \end{aligned}$$

Where we use that $s_0 = \frac{1}{1+\sum_{j=1}^{J} \exp(-\alpha p_j)}$ to get the third equality.

ii Suppose that a new product $J+1$ is introduced in the market. What happens to W? Interpret your results.

If a new product $J+1$ is introduced, the market share of the outside option s_0 will decrease (weakly). Therefore, $\frac{1}{s_0}$ will increase and thus W will increase. Intuitively, some consumers will choose this new product $J+1$, which means that they now achieve (weakly) higher utility then they did under the old choice set. Therefore, welfare will (weakly) increase.

Exercise 2

The file *ps1ex2.csv* contains aggregate data on a large number $T = 1000$ of markets in which $J = 6$ products compete between each other with an outside good $j = 0$. The utility of consumer i is give by:

$$u_{ijt} = -\alpha p_{jt} + \beta x_{jt} + \xi_{jt} + \epsilon_{ijt} \quad j = 1, \ldots, 6$$

$$u_{i0t} = \epsilon_{i0t}$$

where p_{jt} is the price of porduct j in market t, x_{jt} is an observed product characteristic, ξ_{jt} is an unobserved product characteristic and ϵ_{ijt} is i.i.d. $T1EV(0,1)$. Our goal is to estimate demand parameters (α, β) and perform some counterfactual exercise.

1. Assuming that the variable z in the dataset is a valid instrument for prices, write down the moment condition that allows you to consistently estimate (α, β) and obtain an estimate for both parameters.

The relevant moment condition is:

$$\mathbb{E}[\xi_{jt}|z_jt] = 0$$

Using our expression for s_j derived in Exercise 1.4.b, we have that

$$\frac{s_{jt}}{s_{0t}} = \exp(-\alpha p_{jt} + \beta x_{jt} + \xi_{jt}) \implies \ln(\frac{s_{jt}}{s_{0t}}) = -\alpha p_{jt} + \beta x_{jt} + \xi_{jt}$$

and so the relevant moment condition in terms of (α, β) is

$$\mathbb{E}\left[\left(\ln(\frac{s_{jt}}{s_{0t}}) - (-\alpha p_{jt} + \beta x_{jt})\right)|z_{jt}\right] = 0$$

Which results in the following GMM problem:

$$\min_{\alpha,\beta} \sum_{j,t} \left((\ln(\frac{s_{jt}}{s_{0t}}) - (-\alpha p_{jt} + \beta x_{jt}))z_{jt}\right)' \Omega \left((\ln(\frac{s_{jt}}{s_{0t}}) - (-\alpha p_{jt} + \beta x_{jt}))z_{jt}\right)$$

Note that this problem is equivalent to estimating

$$\ln(\frac{s_{jt}}{s_{0t}}) = -\alpha p_{jt} + \beta x_{jt}$$

using z_{jt} as an instrument for p_{jt} .

Our estimates are as follows:

$$\hat{\alpha} = -0.4675$$

$$\hat{\beta} = 0.3047$$

2. For each market, compute own and cross-product elasticities. Average your results across markets and present them in a $J \times J$ table whose (i, j) element contains the (average) elasticity of product i with respect to an increase in the price of product j. What do you notice?

As derived in Exercise 1, the price elasticities are given by the following expression:

$$\frac{\partial s_j}{\partial p_k}\frac{p_k}{s_j} = \begin{cases} -\alpha p_k(1-s_k) & \text{if } j = k \\ \alpha p_k s_k & \text{if } j \neq k \end{cases}$$

Using our parameter estimates, the data yields the following average price elasticities:

Table 1: Matrix of Price Elasticities

	1	2	3	4	5	6
1	**-1.25**	0.323	0.129	0.127	0.126	0.129
2	0.321	**-1.25**	0.129	0.127	0.126	0.129
3	0.321	0.323	**-1.29**	0.127	0.126	0.129
4	0.321	0.323	0.129	**-1.29**	0.126	0.129
5	0.321	0.323	0.129	0.127	**-1.29**	0.129
6	0.321	0.323	0.129	0.127	0.126	**-1.29**

The cross-price elasticities are entirely dependent on market share, so the cross price elasticity of good j with respect to good k does not depend at all on any characteristics of good j and only on the market share of good k. This gives inflexible substitution patterns (IIA).

3. Using you demand estimates, for each product in each market recover the marginal cost c_{jt} implied by Nash-Bertrand competition. For simplicity, you can assume that in each market each product is produced by a different firm (i.e. there is no multi-product firms). Report the average (across markets) marginal cost for each product. Could differences in marginal costs explain the differences in the average (across markets) market shares and prices that you observe in the data?

In each market t, assuming each firm only produces one product j, firms competing Nash-Bertrand will solve the following problem:

$$\max_{p_{jt}}(p_{jt} - c_{jt})Ms_{jt}(p) - C_f$$

FOC:

$$(p_{jt} - c_{jt})\frac{\partial s_{jt}}{\partial p_{jt}} + s_{jt}(p) = 0$$

$$\implies c_{jt} = p_{jt}(1 + \frac{1}{e_{jjt}})$$

where

$$e_{jjt} \equiv -\alpha p_{jt}(1 - s_{jt}) \quad \text{(own-price elasticity)}$$

Plugging in our estimate of α, we get the following average marginal costs:

Table 2: Average Marginal Cost

Firm/Product	Marginal Cost
1	0.667
2	0.672
3	0.679
4	0.689
5	0.683
6	0.683

The marginal costs of each product are quite similar, whereas the average prices and shares are as follows:

Table 3: Average Price and Market Shares

Firm/Product	Price	Share
1	3.36	.203
2	3.37	.203
3	3.03	.090
4	3.04	.089
5	3.03	.088
6	3.04	.091

We can see that even though the marginal costs of goods 3-6 are slightly larger than the marginal costs of good 1 and 2, goods 1-2 are around 10% more expensive and have more than double the average market share, meaning that the difference in prices or shares cannot be explained by the differences in marginal cost. Rather, products 1 and 2 have smaller (in absolute terms) own-price elasticities, resulting in them having higher mark-ups and higher prices. This lower elasticity may be because these goods have greater observed and or unobserved quality, generating greater consumer demand.

4. Suppose that product $j = 1$ exits the market. Assuming that marginal costs and product characteristics for the other products remain unchanged, use you estimated marginal costs and demand parameters to simulate counterfactual prices and market shares in each market. Report the average prices and shares.

The new market share of product $j \neq 1$ in market t is given by

$$s_{jt}^n = \frac{\exp(-\alpha p_{jt}^n + \beta x_{jt} + \xi_{jt})}{1 + \sum_{k \neq 1} \exp(-\alpha p_{kt}^n + \beta x_{kt} + \xi_{kt})}$$

where p_{jt}^n are the new prices following from the exit of firm 1. Notice that we can also express the new prices p_{jt}^n in terms of the new shares s_{jt}^n using the FOC derived in Exercise 2.3 :

$$\begin{aligned} c_{jt} &= p_{jt}^n\left(1 + \frac{1}{e_{jjt}^n}\right) \qquad e_{jjt}^n \equiv -\alpha p_{jt}^n (1 - s_{jt}^n) \\ &= p_{jt}^n - \frac{1}{\alpha(1 - s_{jt}^n)} \implies \\ p_{jt}^n &= c_{jt} + \frac{1}{\alpha(1 - s_{jt}^n)} \end{aligned}$$

Therefore, we have defined a system of 5 non-linear equations for each market t:

$$s_{jt}^n = \frac{\exp((-\alpha c_{jt} - \frac{1}{(1 - s_{jt}^n)}) + \beta x_{jt} + \xi_{jt})}{1 + \sum_{k \neq 1} \exp(-\alpha c_{kt} - \frac{1}{(1 - s_{kt}^n)}) + \beta x_{kt} + \xi_{kt})}$$

where $k = 2, \ldots, 6$. This defines a fixed point problem $s_{jt}^n = f(s_{jt}^n, \{s_{kt}^n\})$, which can be solved by iterating over values of s_{jt} until the left-hand side equals the right hand side for each product and market.

We compute the following prices and shares:

Table 4: Average Prices and Market Shares by Product

Product	Avg. New Price	Avg. Old Price	Avg. New Share	Avg. Old Share
2	3.518	3.368	0.244	0.203
3	3.095	3.033	0.113	0.090
4	3.100	3.040	0.111	0.089
5	3.092	3.031	0.111	0.088
6	3.100	3.038	0.114	0.091

5. Finally, for each market compute the change in firms profits and in consumer welfare following the exit of firm $j = i$. Report the average changes across markets. Who wins and who loses?

The change in profit for firm j selling product j in market t is given by:

$$\Delta\Pi_{jt} = M((p_{jt}^n - c_{jt})s_{jt}^n - (p_{jt}^o - c_{jt})s_{jt}^o)$$

where M is the number of individuals in the market and can be ignored for our purposes. Likewise, the change in welfare from consumers in market t is given by:

$$\begin{aligned}\Delta W_t &= \ln\left(\frac{1 - s_{0t}^n}{s_{0t}^n}\right) - \ln\left(\frac{1 - s_{0t}^o}{s_{0t}^o}\right) \\ &\approx -0.358 \qquad \text{(our estimate)}\end{aligned}$$

We compute the following changes in profits:

Table 5: Average Change in Profit per Firm

Firm	Δ profit
1	NaN
2	0.150
3	0.062
4	0.061
5	0.061
6	0.062

As we can see, the firms win at the expense of the consumers. Intuitively, with fewer choices in the market, each firm faces less competition and can increase their markups. As a result, consumer's retain less of the surplus.

Exercise 3

One of the possible ways to relax the restrictive substitution pattern from the homogeneous logit model is to use the nested logit model. Suppose that the utility of each product $j \in \mathcal{J}$ to consumer i is given by:

$$u_{ijt} = \beta x_{jt} + \xi_{jt} + \epsilon_{ijt}$$

$$u_{i0t} = \epsilon_{i0t}$$

where x_{jt} is an observed product characteristic and ξ_{jt} is an unobserved product characteristic. However, the products are now "grouped" into G disjoint subsets. Each group is denoted by B_g, where $g \in \{1, 2, \ldots, G\}$. The outside good ($j = 0$) belongs to none of the groups (or you can alternatively treat it as the only product in the "outside group", $g = 0$). Given the setting, suppose the joint distribution of $\{\epsilon_{ij}\}$ is of the form

$$F(\{\epsilon_j\}_{j\in\mathcal{J}}) = \exp\left(-\sum_{g=1}^{G}\left[\sum_{j\in B_g}\exp\left(-\rho^{-1}\epsilon_j\right)\right]^{\rho}\right), \qquad 0 < \rho \leq 1,$$

and the marginal distribution of ϵ_{i0t} (distributed independently from the other ϵ) is

$$F(\epsilon_0) = \exp\left(-\exp(-\epsilon_0)\right).$$

1. Derive the expression for $\mathbb{P}(i \text{ Chooses } B_g)$.
(*Hint*: You may start by considering a simpler case when there is only one group with two products and an outside group. In other words, $j \in \{0, 1, 2\}$) where products 1 and 2 are in group $g = 1$ and 0 is an outside good. Then, use the intuition from the simple case above to generalize the expression.)

For the time being we will drop the t subscript. Intuitively, in this simpler case, a consumer will choose group $g = 1$ if either the utility from product 1 or product 2 is greater than the utility of the outside good. Formally:

$$\begin{aligned}
\mathbb{P}(i \text{ Chooses } B_g) &= \mathbb{P}(\max_{j=1,2}\{u_{ij}\} \geq u_{i0}) \\
&= 1 - \mathbb{P}(\epsilon_{i1} \leq \epsilon_{i0} - \beta x_1 - \xi_1, \epsilon_{i2} \leq \epsilon_{i0} - \beta x_2 - \xi_2) \\
&= 1 - \int_{-\infty}^{\infty} \exp\left(-\left[\exp(-\rho^{-1}(\epsilon_{i0} - \beta x_1 - \xi_1)) + \exp(-\rho^{-1}(\epsilon_{i0} - \beta x_2 - \xi_2))\right]^{\rho}\right) f(\epsilon_{i0}) d\epsilon_{io} \\
&= 1 - \int_{-\infty}^{\infty} \exp\left(-\left[\exp(-\rho^{-1}\epsilon_{i0})\exp(\rho^{-1}\underbrace{\delta_1}_{\equiv\beta x_1+\xi_1})) + \exp(-\rho^{-1}\epsilon_{i0})\exp(\rho^{-1}\underbrace{\delta_2}_{\equiv\beta x_2+\xi_2}))\right]^{\rho}\right) f(\epsilon_{i0}) d\epsilon_{io} \\
&= 1 - \int_{-\infty}^{\infty} \exp\left(-\exp(-\epsilon_{i0})\underbrace{\left(\left(\exp(\rho^{-1}\delta_1)\right) + \exp(\rho^{-1}\delta_2)\right)^{\rho}}_{\equiv P}\right) \exp(-\epsilon_{i0})\exp(-\exp(-\epsilon_{i0}))) d\epsilon_{io} \\
&= 1 - \int_{-\infty}^{\infty} e^{-\epsilon_{i0}} e^{-(1+P)e^{-\epsilon_{i0}}} d\epsilon_{io}
\end{aligned}$$

After applying integration by substitution with $y = e^{-\epsilon_{i0}}$, we get that:

$$\begin{aligned}
\mathbb{P}(i \text{ Chooses } B_g) &= 1 - \frac{1}{1+P} \\
&= \frac{\left[\exp(\rho^{-1}\delta_1) + \exp(\rho^{-1}\delta_2)\right]^{\rho}}{1 + \left[\exp(\rho^{-1}\delta_1) + \exp(\rho^{-1}\delta_2)\right]^{\rho}} \qquad \delta_j = \beta x_j + \xi_j
\end{aligned}$$

We can see that expression is quite similar to the expression for the simple multinomial logit choice probability, where we have an exponential term related to the chosen product in the numerator and the sum of exponential terms for all products in the denominator. In this case with nests, these exponential terms are the sum of the exponential terms with product characteristics for all products in a given nest g. Therefore, we can generalize the expression for when we have multiple nests $g \in \{1, \ldots, \}$ with products $j \in B_g$ as:

$$\mathbb{P}(i \text{ Chooses } B_g) = \frac{\left[\sum_{j \in B_g} \exp(\rho^{-1}\delta_j)\right]^{\rho}}{1 + \sum_{g=1}^{G} \left[\sum_{j \in B_g} \exp(\rho^{-1}\delta_j)\right]^{\rho}} \qquad \delta_j = \beta x_j + \xi_j$$

2. Derive the expression for $\mathbb{P}(i \text{ Chooses } j | i \text{ Chooses } B_g)$ when $j \in B_g$.

We can write this probability as follows:

$$\begin{aligned}\mathbb{P}(i \text{ Choose } j | i \text{ Choose} B_g) &= \frac{\mathbb{P}(i \text{ Choose } j)}{\mathbb{P}(i \text{ Choose } B_g)} \qquad \text{Bayes rule, } P(\text{choose } B_g|\text{choose } j) = 1 \\ &= \frac{\mathbb{P}(i \text{ Choose } j)}{\sum_{k \in B_g} \mathbb{P}(i \text{ Choose } k)}\end{aligned}$$

While the derivation of the following theorem is much too difficult for me, McFadden (1978) and some other papers from the 70's show that we can express $\mathbb{P}(i \text{ Choose } j)$ as:

$$\mathbb{P}(i \text{ Choose } j) = \frac{1}{1+G}\frac{\partial G}{\partial \delta_j} \qquad \delta_j \equiv \beta x_j + \xi_j$$

where G comes from the CDF for the maximum of the joint probability:

$$\begin{aligned}\mathbb{P}(\max_j u_{ij} \leq x) &= \mathbb{P}(\epsilon_{ij} \leq x - \delta_j \ \forall j) \\ &= \exp\left(-\sum_{g=1}^{G}\left[\exp(-x\rho^{-1})\sum_{j \in B_g} \exp\left(\rho^{-1}\delta_j\right)\right]^{\rho}\right) \\ &= \exp\left(-\exp(-x)\underbrace{\sum_{g=1}^{G}\left[\sum_{j \in B_g}\exp\left(\rho^{-1}\delta_j\right)\right]^{\rho}}_{\equiv G}\right)\end{aligned}$$

Then, knowing G, we have that:

$$\begin{aligned}\frac{\partial G}{\partial \delta_j} &= \rho\left[\sum_{j \in B_g}\exp\left(\rho^{-1}\delta_j\right)\right]^{\rho-1}\frac{1}{\rho}\exp(\rho^{-1}\delta_j) \\ &= \left[\sum_{j \in B_g}\exp\left(\rho^{-1}\delta_j\right)\right]^{\rho-1}\exp(\rho^{-1}\delta_j)\end{aligned}$$

and so

$$\mathbb{P}(i \text{ Choose } j) = \frac{1}{1+G}\left(\sum_{j \in B_g}\exp\left(\rho^{-1}\delta_j\right)\right)^{\rho-1}\exp(\rho^{-1}\delta_j)$$

and

$$\sum_{k\in B_g}\mathbb{P}(i \text{ Choose } k) = \frac{1}{1+G}\left(\sum_{j\in B_g}\exp\left(\rho^{-1}\delta_j\right)\right)^{\rho-1}\sum_{j\in B_g}\exp(\rho^{-1}\delta_j)$$
$$= \frac{1}{1+G}\left(\sum_{j\in B_g}\exp\left(\rho^{-1}\delta_j\right)\right)^{\rho}$$

Then finally,

$$\mathbb{P}(i \text{ Choose } j|i \text{ Choose} B_g) = \frac{\mathbb{P}(i \text{ Choose } j)}{\sum_{k\in B_g}\mathbb{P}(i \text{ Choose } k)}$$
$$= \frac{\exp(\rho^{-1}\delta_j)}{\sum_{k\in B_g}\exp(\rho^{-1}\delta_k)}$$

3. Derive the expression for $\mathbb{P}(i \text{ Chooses } j)$.

$$\mathbb{P}(i \text{ Choose } j) = \mathbb{P}(i \text{ Choose } j|i \text{ Choose} B_g)\mathbb{P}(i \text{ Choose } B_g) \qquad \text{Bayes rule}$$
$$= \frac{\exp(\rho^{-1}\delta_j)}{\sum_{k\in B_g}\exp(\rho^{-1}\delta_k)}\frac{\left[\sum_{j\in B_g}\exp(\rho^{-1}\delta_j)\right]^{\rho}}{1+\sum_{g=1}^{G}\left[\sum_{j\in B_g}\exp(\rho^{-1}\delta_j)\right]^{\rho}}$$
$$= \frac{\exp(\rho^{-1}\delta_j)\left[\sum_{k\in B_g}\exp(\rho^{-1}\delta_k)\right]^{\rho-1}}{1+\sum_{g=1}^{G}\left[\sum_{k\in B_g}\exp(\rho^{-1}\delta_k)\right]^{\rho}} \qquad \delta_k = \beta x_k + \xi_k$$

4. Suppose we approximate each of the probabilities above using the observed market shares as

- $s_{jt} \approx \Pr(i \text{ chooses } j)$,
- $s_{jt|g_t} \approx \Pr(i \text{ chooses } j \mid i \text{ chooses } B_g)$.

Use the expressions for the probabilities derived above to find the relationship between $\log(s_{jt}) - \log(s_{0t})$, $\log(s_{jt|g_t})$, x_{jt}, and ξ_{jt}.

We can model the probability of choosing the outside option as the probability of not choosing any of the $j = 1, \ldots, \mathcal{J}$ products (where we assume the outside good belongs to some outside group $g = 0$):

$$s_{0t} = 1 - \sum_{g=1}^{G}\sum_{k\in B_g}\frac{\exp(\rho^{-1}\delta_k)\left[\sum_{k\in B_g}\exp(\rho^{-1}\delta_k)\right]^{\rho-1}}{1+\sum_{g=1}^{G}\left[\sum_{k\in B_g}\exp(\rho^{-1}\delta_k)\right]^{\rho}}$$
$$= 1 - \frac{\sum_{g=1}^{G}\left[\sum_{k\in B_g}\exp(\rho^{-1}\delta_k)\right]^{\rho}}{1+\sum_{g=1}^{G}\left[\sum_{k\in B_g}\exp(\rho^{-1}\delta_k)\right]^{\rho}}$$
$$= \frac{1}{1+\sum_{g=1}^{G}\left[\sum_{k\in B_g}\exp(\rho^{-1}\delta_k)\right]^{\rho}}$$

and then

$$\begin{aligned}
\log(s_{jt}) - \log(s_{0t}) &= \log(\frac{s_{jt}}{s_{0t}}) \\
&= \log(\exp(\rho^{-1}\delta_j)\left[\sum_{k\in B_g}\exp(\rho^{-1}\delta_k)\right]^{\rho-1}) \\
&= \frac{\delta_{jt}}{\rho} + (\rho-1)\log(\sum_{k\in B_g}\exp(\rho^{-1}\delta_k))
\end{aligned}$$

Likewise,

$$\log(s_{jt|g_t}) = \frac{\delta_{jt}}{\rho} - \log(\sum_{k\in B_g}\exp(\rho^{-1}\delta_k))$$

Therefore

$$\log(s_{jt}) - \log(s_{0t}) = \beta x_{jt} + \xi_{jt} - (\rho-1)(\log(s_{jt} \mid g_t))$$

5. Given the relationship derived in the previous part, some researchers have suggested that we can estimate the model parameters by regression $\log(s_{jt}) - \log(s_{0t})$ on $\log(s_{jt|gt})$ and x_{jt}. Do you think such regression may suffer from potential endogeneity issues? Unlike the homogeneous logit case, can the regression still have endogeneity issues even when the unobserved characteristics (ξ_{jt}) are uncorrelated with any of the observed characteristics (x_{jt})?

The previous relationship suggests the following regression:

$$\log(s_{jt}) - \log(s_{0t}) = \beta x_{jt} + \underbrace{\alpha}_{\equiv 1-\rho}(\log(s_{jt} \mid g_t)) + \epsilon_{jt}$$

Of course, we have potential endogeneity issues as the unobserved (to the econometrician) variable ξ_{jt} in the error term is likely correlated with x_{jt}. Furthermore, even if $Cov(\xi_{jt}, x_{jt}) = 0$, notice that $\log(s_{jt|g_t}) = \frac{\delta_{jt}}{\rho} - \log(\sum_{k\in B_g}\exp(\rho^{-1}\delta_{kt}))$, where $\delta_{jt} = \beta x_{jt} + \xi_{jt}$. Therefore, $\log(s_{jt|g_t})$ is mechanically correlated with the unobserved variable ξ_{jt}, resulting in exogeneity.

6. To overcome the endogeneity problem discussed in the part above, one of the instruments commonly used in the literature is the number of products within each group. Assuming that the number of products within each group is determined independently of any of the product characteristics, do you think it is a good instrument? Support your claim by simulating data and checking whether you can recover the true parameters assumed in the data-generating process.

The logic behind using this instrument is that an increase in the number of products in a group will decrease the market share of each product. To see why, note than an increase in $\sum_{k\in B_g}\exp(\rho^{-1}\delta_k)$ will cause $\mathbb{P}(i \text{ Choose } j|i \text{ Choose} B_g) = \frac{\exp(\rho^{-1}\delta_j)}{\sum_{k\in B_g}\exp(\rho^{-1}\delta_k)}$ to decrease. Therefore, if exogeneity is satisfied, then the number of products in a group is a valid instrument for $\log(s_{jt} \mid g_t)$ in the following regression

$$\log(s_{jt}) - \log(s_{0t}) = \beta x_{jt} + \underbrace{\alpha}_{\equiv 1-\rho}(\log(s_{jt} \mid g_t)) + \underbrace{\epsilon_{jt}}_{\text{contains } \xi_{jt}}$$

As we can see, exogeneity in this case means that the number of products in group g_t is independent of unobserved product characteristics ξ_{jt}, conditional on x_{jt}. Therefore, if the number of products are determined independently of any of the product characteristics, than it might be a good instrument. One concern is that there may be groups where a few goods have the vast majority of the market share as they are overwhelmingly preferred, meaning that a change in the number of products could result in only a small change in $\mathbb{P}(i \text{ Choose } j | i \text{ Choose} B_g)$ for these goods. As a result, the number of products would be a weak instrument in this case, biasing the linear IV estimates towards the regular OLS estimates. Furthermore, the assumption that the number of products in a group is unrelated to unobserved product characteristics is unrealistically strong. For one, it rules out the possibility that firms may leave or enter a product group depending on the unobserved characteristics of the goods in that group, which would not make sense if firms act rationally on information about other firms. For example, if a certain cereal firm producing sweetened children's cereal has considerable brand loyalty, a typical unobserved product characteristic, then a cereal firm launching a new product would be more likely to enter a product group with less brand loyalty, such as adult fiber cereal, compared to the sweetened children's cereal product group.

To show how using an instrument as a product may work, I simulate product data x_j, ξ_j for products in one of 10 nests, with the number of products within a nest given by a random number ranging from 3 to 10. Using the expressions derived previously and parameter values $\beta = 1, \rho = .5$, I compute s_j, s_0 and $s_{j|g}$ and run the following regression using both regular OLS and IV where $s_{j|g}$ is instrumented by the number of products within that group:

$$\log(s_j) - \log(s_0) = \beta x_j + \underbrace{\alpha}_{\equiv 1-\rho} \log(s_j \mid g) + \epsilon_{jt}$$

These regressions yield the following results:

Table 6: OLS and IV Estimates

	OLS	IV
$\widehat{\beta}$ $(\beta = 1)$	0.587	1.191
	[0.422, 0.752]	[0.865, 1.516]
$\widehat{\alpha}$ $(\alpha = .5)$	0.704	0.344
	[0.631, 0.777]	[0.163, 0.524]
First-stage F-statistic	–	7.527
Number of observations	73	73

Notes: The dependent variable is $\log(s_j) - \log(s_0)$. The endogenous regressor in the IV specification is $\log(s_j \mid g)$, instrumented using the number of products in the group. OLS standard errors are non-robust. IV standard errors are heteroskedasticity-robust. 95% confidence intervals are reported in brackets.

We can see that the regular OLS regression without an instrument yields biased estimates of β and α where the true values are not in the 95% confidence intervals. This is the case as $\log(s_j|g)$ is correlated with the unobserved to the econometrician ξ_j term. However, the regression that uses the number of products as an instrument for $\log(s_j|g)$ yields unbiased estimates of β and α that do include the true values in the 95% confidence intervals. While IV estimation seems to have worked well in this simulation, it is important to note two things. First, as expected, the standard errors from the IV regression are larger, resulting in less precise estimates. Secondly, the independence between the number of products in a group and ξ_j is by design, and, as argued above, will not generally be satisfied in real world data.

Exercise 4

The file `ps1_ex4.csv` contains aggregate data on $T = 100$ markets in which $J = 6$ products compete with each other together with an outside good $j = 0$. The utility of consumer i is given by:

$$u_{ijt} = \tilde{x}'_{jt}\beta + \xi_{jt} + \tilde{x}'_{jt}\Gamma v_i + \varepsilon_{ijt}, \qquad j = 1, \ldots, 6,$$

$$u_{i0t} = \varepsilon_{i0t},$$

where $\tilde{x}_{jt}$ is a vector of observed product characteristics including the price, ξ_{jt} is an unobserved product characteristic, v_i is a vector of unobserved taste shocks for the product characteristics, and ε_{ijt} is i.i.d. Type I Extreme Value, $\varepsilon_{ijt} \sim$ T1EV$(0,1)$. Our goal is to estimate demand parameters (β, Γ) using the BLP algorithm. As you can see from the data, there are only two characteristics

$$\tilde{x}_{jt} = (p_{jt}, x_{jt}),$$

namely prices and an observed measure of product quality. Moreover, there are several valid instruments z_{jt} that you will use to construct moments to estimate (β, Γ). Finally, you can assume that Γ is lower triangular, e.g.,

$$\Gamma = \begin{pmatrix} \gamma_{11} & 0 \\ \gamma_{21} & \gamma_{22} \end{pmatrix},$$

such that $\Gamma\Gamma' = \Omega$ is a positive definite matrix, and that v_i is a two-dimensional vector of i.i.d. random taste shocks.

1. Assume that $v_i \sim \mathcal{N}(0, I)$ so that $\Gamma v_i \sim \mathcal{N}(0, \Omega)$. Implement the BLP routine to estimate (β, Γ). You may want to write multiple functions, including but not limited to: a share prediction function, a share inversion function, an implicit function $\xi(\beta, \Gamma)$, and an objective function to be minimized.

Our coding of the BLP routine using the Nelder-Mead optimization algorithm gives the following results:

$$\begin{bmatrix} \hat{\beta}_p \\ \hat{\beta}_x \end{bmatrix} = \begin{bmatrix} -1.550 \\ 1.342 \end{bmatrix} \qquad \hat{\Gamma} = \begin{bmatrix} 4.19 & 0 \\ 4.91 & -1.82 \end{bmatrix} \qquad \hat{\Omega} = \hat{\Gamma}\hat{\Gamma}' = \begin{bmatrix} 17.57 & 20.59 \\ 20.59 & 27.43 \end{bmatrix}$$

2. For each market, compute cross and own product elasticities. Average your results across markets and present them in a $J \times J$ table whose (i, j) element contains the (average) elasticity of product i with respect to an increase in the price of product j. What is the main difference compared with the table of elasticities you found in Exercise 2.2?

Table 7: Matrix of BLP Price Elasticities

	1	2	3	4	5	6
1	**-0.0045**	0.0039	0.1795	0.2530	0.1902	0.1317
2	0.0060	**-0.0047**	0.2110	0.1529	0.1161	0.2029
3	0.0034	0.0016	**-3.0382**	0.1100	0.0456	-0.0297
4	0.0034	0.0017	0.1304	**-3.0614**	0.0417	-0.0956
5	0.0005	0.0010	-0.2452	-0.0230	**-1.9646**	-0.0016
6	0.0006	0.0003	-0.1354	-0.0615	-0.0346	**-1.4990**

As we can see, random coefficients allow for much more flexible substitution patterns. Specifically, the demand elasticity of good j with respect to good k varies by good j, unlike in exercise 2.2. This means that we have relaxed the IIA assumption and allow substitution patterns to vary based on product and consumer characteristics .

3. Look at the average (across markets) prices, shares, and observed quality of the products in the data. Based on your estimated Γ, what do you think could be driving differences in prices and market shares?

Table 8: Summary Statistics

Product/Firm	Price	Observed Quality (x)	Share
1	0.0024	-0.019	0.0988
2	0.0023	-0.026	0.0891
3	2.0191	-0.081	0.0430
4	1.7516	-0.180	0.0393
5	3.5770	1.693	0.1517
6	4.4429	2.002	0.1932

Note that the estimate of Ω implies a correlation between β_p and β_x of:

$$\rho_{px} = \frac{20.59}{\sqrt{17.57 * 27.43}} = .938$$

This means that individuals preferences over price and quality (x) are very highly correlated. We can therefore split people into two general groups: low types, who are highly price sensitive and have low preference for x and high types, who are not price sensitive (or even enjoy high prices) and have high taste for x. Given that products 1 and 2 are orders of magnitude cheaper than the other products and have low observed quality, the low types will overwhelmingly purchase these products. Likewise, since products 5 and 6 have much higher quality, but relatively high prices, the high types will overwhelmingly purchase these products. Finally, products 3 and 4 are both much more expensive than products 1 and 2 and much lower quality than product 5 and 6, and will therefore be purchased primarily by individuals who are insensitive to high prices and have low preference for x or who receive an idiosyncratic taste shock for these products. This intuition for who buys what aligns well with the observed market shares and price elasticities. We can see that products 5 and 6 have similarly large market shares, products 1 and 2 have similarly moderate market shares , and products 3 and 4 have similarly low shares. Likewise, products 1 and 2 have cross-price elasticities that are higher than they are with respect to the other goods and have extremely low own-price elasticities, as they are the only goods with low quality and price. Products 3 and 4 have high cross price elasticities with goods 1-4 and low own-price elasticities, which reflects the fact that many consumers of goods 3 and 4 would be willing to substitute towards goods 1 and 2 due to their similar qualities and much lower prices. Finally, goods 5 and 6 have moderately low own price elasticities, as the high types are relatively indifferent between 5 and 6, but also have positive cross price elasticities, which is not very intuitive.

4. Compare your results with `PyBLP`.

The PyBLP package using the BFGS optimization algorithm yields the following parameter estimates:

$$\begin{bmatrix}\hat{\beta}_p \\ \hat{\beta}_x\end{bmatrix} = \begin{bmatrix}-1.949 \\ .846\end{bmatrix} \qquad \hat{\Gamma} = \begin{bmatrix}4.44 & 0 \\ 4.94 & -1.77\end{bmatrix} \qquad \hat{\Omega} = \hat{\Gamma}\hat{\Gamma}' = \begin{bmatrix}19.77 & 21.96 \\ 21.96 & 27.53\end{bmatrix}$$

As we can see, PyBLP returns estimates of Γ that are very similar to our own. It's estimate of β_p is around 25% larger and it's estimate of β_x is around 35% smaller. In light of these relatively small differences, the resulting distribution's look quite similar:

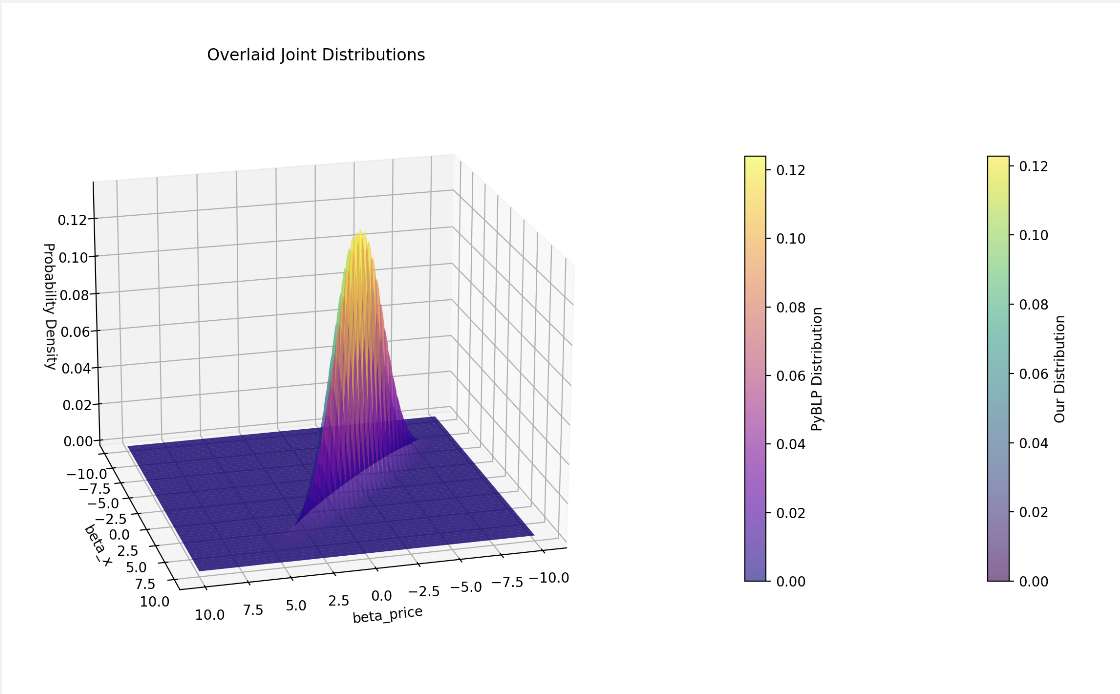

Figure 1: Own vs PyBLP $\hat{\beta}$ Distribution