Assignment 1

Luke Motley, Jeffrey Ohl

Due Jan 31

Exercise 1

Problem 1

Let X_i with $i = 1, \ldots, N$ be a sequence of independent type 1 extreme value random variables with location parameter μ_i and scale parameter $\sigma > 0$ (T1EV (μ_i, σ)). The c.d.f. is given by:

$$\Pr\{X_i \leq x \mid \mu_i, \sigma\} = \exp\left(-\exp\left(-\frac{x-\mu_i}{\sigma}\right)\right)$$

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

By independence,

$$\begin{aligned} P\{Y \leq y\} &= P\{\max_i\{X_i\} \leq y\} \\ &= P\{X_1 \leq y, X_2 \leq y, \ldots, X_N \leq y\} \\ &= \prod_{i=1}^{N} P\{X_i \leq y\} \end{aligned}$$

Therefore, the distribution of Y is

$$\begin{aligned} F_Y(y) = P\{Y \leq y\} &= \prod_{i=1}^{N} P\{X_i \leq y\} \\ &= \prod_{i=1}^{N} \exp\left(-\exp\left(-\frac{y-\mu_i}{\sigma}\right)\right) \\ &= \exp\left(-\sum_{i=1}^{N} \exp\left(-\frac{y-\mu_i}{\sigma}\right)\right) \end{aligned}$$

Problem 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 $X - Y$.

The difference of two independent Gumbel (T1EV) random variables with the same scale follows a logistic distribution, namely

$$Z = X - Y \quad \Longrightarrow \quad Z \sim \text{Logistic}\left(\mu_X - \mu_Y, \sigma\right)$$

or

$$F_{X-Y}(z) = \frac{1}{1 + \exp\left(-\frac{z-(\mu_X - \mu_Y)}{\sigma}\right)}$$

If needed, we could solve directly by computing

$$F_Z(z) = \Pr\{X - Y \leq z\} = \int_{-\infty}^{\infty} \Pr\{X \leq z + y\} f_Y(y) dy$$

using

$$F_X(x) = \Pr\{X \leq x\} = \exp\left[-\exp\left(-\frac{x - \mu_X}{\sigma}\right)\right]$$
$$f_Y(y) = \frac{1}{\sigma}\exp\left[-\frac{y-\mu_Y}{\sigma}\right]\exp\left[-\exp\left(-\frac{y-\mu_Y}{\sigma}\right)\right]$$

Problem 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 ϵ_j are independent and identically distributed T1EV $(0, 1)$. Derive the probability that alternative j is chosen.

The individual chooses the alternative that maximizes utility. Therefore, we want

$$\begin{aligned}
\Pr\{u_j \geq \max_k u_k\} &= \Pr\{u_j \geq u_k \text{ for all } k\} \\
&= \Pr\{\mu_j + \epsilon_j \geq \mu_k + \epsilon_k \text{ for all } k\} \\
&= \Pr\{\epsilon_k - \epsilon_j \leq \mu_j - \mu_k \text{ for all } k\}
\end{aligned}$$

The standard result here is

$$\Pr\{u_j \geq \max_k u_k\} = \frac{\exp(\mu_j)}{\sum_{k=1}^{N} \exp(\mu_k)}$$

Again, if necessary, we could solve the following integral to get this result

$$\begin{aligned}\Pr\{u_j \geq u_k \text{ for all } k\} &= \int_{-\infty}^{+\infty} \Pr\left(\epsilon_k \leq \epsilon_j + [\mu_j - \mu_k] \, \forall k \neq j \mid \epsilon_j = t\right) f_{\epsilon_j}(t) dt \\ &= \int_{-\infty}^{+\infty} \left[\prod_{k \neq j} F_{\epsilon_k}\left(t + (\mu_j - \mu_k)\right)\right] f_{\epsilon_j}(t) dt\end{aligned}$$

where the product follows by independence and we have

$$F_{\epsilon_k}(x) = \exp[-\exp(-x)]$$

Problem 4

Consider a market with J products indexed by $j = 1, \ldots, J$, 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:

$$\begin{aligned}u_{ij} &= \alpha(y_i - p_j) + \epsilon_{ij} \text{ for } j = 1, \ldots, J \\ u_{i0} &= \alpha y_i + \epsilon_{i0} \text{ for } j = 0\end{aligned}$$

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.

(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.

(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 that same characteristic. Moreover, assume that $x_j > 0, x_0 = 0$.

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

(ii) Assume that β_i are i.i.d Uniform $[0, \bar{\beta}]$ with $\bar{\beta}$ 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 $x_i \geq x_j \geq x_k$)

(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 $F(\cdot)$. Derive product j

's market share and compute own and cross-price elasticities. Explain and compare with your findings in point (b) above.

(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

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

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

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

(a) As in (3), consumers will choose the product that maximizes their utility. Therefore, we have

$$s_j(i) = \Pr\left[u_{ij} \geq u_{ik} \text{ for all } k \neq j\right]$$

We can directly use the result from (3) here, so

$$s_j(i) = \frac{\exp\left(\alpha\left(y_i - p_j\right)\right)}{\sum_{k=0}^{J} \exp\left(\alpha\left(y_i - p_k\right)\right)}$$

where if we have $p_0 = 0$, the denominator is

$$\exp\left(\alpha y_i\right)\left[1 + \sum_{k=1}^{J} \exp\left[-\alpha p_k\right]\right]$$

and therefore

$$s_j(i) = \frac{\exp\left(-\alpha p_j\right)}{1 + \sum_{k=1}^{J} \exp\left(-\alpha p_k\right)}$$

Thus, since the y's cancel out, we have

$$\frac{\partial s_j(i)}{\partial y_i} = 0 \quad \text{for all } j$$

An increase in income will not change the probability of choosing any product, as it impacts the utility of all products equally.

(b) The market share of product j is simply the average of the individual choice probabilities, and since we do not have dependence on i, this is the same as the individual choice probability.

$$s_j = \frac{\exp\left(-\alpha p_j\right)}{1 + \sum_{k=1}^{J} \exp\left(-\alpha p_k\right)}$$

Let's use the following notation when we take derivatives for the elasticities:

$$s_j = \frac{e^{-\alpha p_j}}{D}, \quad \text{where} \quad D = 1 + \sum_{\ell=1}^{J} e^{-\alpha p_\ell}$$

Since

$$\ln s_j = -\alpha p_j - \ln D$$

we have

$$\frac{\partial s_j}{\partial p_j} = s_j \frac{\partial}{\partial p_j} [\ln s_j] = s_j \left[-\alpha - \frac{\partial}{\partial p_j} \ln D \right]$$

Then

$$\frac{\partial}{\partial p_i} \ln D = \frac{1}{D} \frac{\partial D}{\partial p_i} = \frac{1}{D} \left[-\alpha e^{-\alpha p_j} \right] = -\alpha \frac{e^{-\alpha p_j}}{D} = -\alpha s_j$$

so

$$\frac{\partial s_j}{\partial p_j} = s_j \alpha (s_j - 1) = -\alpha s_j [1 - s_j]$$

Therefore,

$$\eta_{jj} = \left[\frac{\partial s_j}{\partial p_j} \right] \frac{p_j}{s_j} = -\alpha p_j [1 - s_j]$$

As expected, this tells us that a higher price lowers the share. Taking the same approach for the cross-price elasticities, we have

$$\frac{\partial}{\partial p_k} [\ln s_j] = -\frac{\partial}{\partial p_k} \ln D = -\frac{1}{D} \left[-\alpha e^{-\alpha p_k} \right] = \alpha \frac{e^{-\alpha p_k}}{D} = \alpha s_k$$

Thus,

$$\frac{\partial s_j}{\partial p_k} = s_j [\alpha s_k] = \alpha s_j s_k$$

and

$$\eta_{jk} = \left[\frac{\partial s_j}{\partial p_k} \right] \frac{p_k}{s_j} = \alpha s_k p_k$$

This also makes sense. An increase in the price of other products will increase the share of the product in question, as they are substitutes. It also depends only on shares and prices of the other product, so when a given product has a price change, all other products respond in the same way. This is a feature of IIA, I think.

(c) Now we transition to a model where there are taste shocks.

(i) If there is no heterogeneity ($\beta_i = \beta$), then all consumers will pick the same product.

Namely,

$$s_j = \begin{cases} 1 & \text{if } j \text{ is the max of } u_{ij} \\ 0 & \text{otherwise} \end{cases}$$

where we are picking the product that maximizes

$$-\alpha p_j + \beta x_j$$

If all are less than 0, then the outside option is chosen. All consumers have the same taste shock, so they all choose the same product.

(ii) Define indirect utility as

$$v_j(\beta) = -\alpha p_j + \beta x_j, \quad j \geq 1, \quad \text{and} \quad v_0(\beta) = 0$$

The consumer picks product j if $v_j(\beta) \geq v_k(\beta)$ for all k. For the outside option, we have

$$v_j(\beta) = -\alpha p_j + \beta x_j \geq 0 \quad \Longrightarrow \quad \beta \geq \frac{\alpha p_j}{x_j}$$

For the other products, we have

$$-\alpha p_j + \beta x_j \geq -\alpha p_k + \beta x_k, \quad \Longrightarrow \quad \beta\left(x_j - x_k\right) \geq \alpha\left(p_j - p_k\right)$$

Then for $[0, \bar{\beta}]$, with uniform distribution and our simplifying assumption, one product will be chosen in each interval, where the cutoffs are

$$\frac{\alpha\left(p_j - p_k\right)}{x_j - x_k}$$

I'm not sure if I can do this more formally, come back to this. For the elasticities, increasing p lowers the utility and thus shrinks the region where the product is chosen. Thus the own-price elasticity is negative. For the cross-price elasticities, increasing p of the other good shrinks the region where it is chosen, thus increasing the region of other products. This gives a positive sign, as expected. Now, cross-price elasticities will differ however, unlike in the previous case.

(d) This is a mix of the previous two cases.

$$u_{ij} = \underbrace{\alpha\left(y_i - p_j\right)}_{\text{mean utility}} + \underbrace{\beta_i x_j}_{\text{random coefficient part}} + \underbrace{v_{ij}}_{\text{T1EV error}}$$

Conditional on β_i, the consumer solves the first T1EV problem,

$$s_j(i \mid \beta) = \frac{\exp\left[\alpha\left(y_i - p_j\right) + \beta x_j\right]}{\sum_{k=0}^{J} \exp\left[\alpha\left(y_i - p_k\right) + \beta x_k\right]}$$

so

$$s_j(\beta) = \frac{\exp\left[-\alpha p_j + \beta x_j\right]}{\sum_{k=0}^{J} \exp\left[-\alpha p_k + \beta x_k\right]}$$

(again taking $p_0 = 0$ and $x_0 = 0$). To get unconditional shares, we would have to integrate

$$s_j = \int s_j(\beta) dF(\beta) = \int \frac{\exp\left[-\alpha p_j + \beta x_j\right]}{\sum_{k=0}^{J} \exp\left[-\alpha p_k + \beta x_k\right]} f(\beta) d\beta$$

Similar, we will have conditional elasticities in the same form as above, so

$$\eta_{jj}(\beta) = \frac{\partial s_j(\beta)}{\partial p_j} \frac{p_j}{s_j(\beta)} = -\alpha p_j \left[1 - s_j(\beta)\right]$$

and

$$\eta_{jk}(\beta) = \frac{\partial s_j(\beta)}{\partial p_k} \frac{p_k}{s_j(\beta)} = \alpha p_k s_k(\beta)$$

We would need to integrate over β to get the unconditional elasticities. Specifically, we will get something that looks like the expectation of the conditional elasticity weighted by the conditional share. As we move from (b) to (d), we increase in complexity but gain in realism.

(e)

(i)

We measure welfare at the given prices $\{p_j\}$ by

$$W \equiv \mathbb{E}\left[\max_{j=0,\ldots,J} u_{ij}\right] = \mathbb{E}\left[\max_j \left(v_j + \epsilon_{ij}\right)\right].$$

A well-known property of i.i.d. T1EV $(0,1)$ errors is that

$$\mathbb{E}\left[\max_j \left(v_j + \epsilon_{ij}\right)\right] = \gamma + \ln\left(\sum_{k=0}^{J} e^{v_k}\right),$$

Hence

$$W = \gamma + \ln\left(1 + \sum_{k=1}^{J} e^{v_k}\right)$$

because $v_0 = 0$ implies $e^{v_0} = 1$.

Recall

$$s_0 = \frac{1}{1 + \sum_{k=1}^{J} e^{v_k}}$$

Hence

$$1 + \sum_{k=1}^{J} e^{v_k} = \frac{1}{s_0}.$$

Therefore,

$$\ln\left(1 + \sum_{k=1}^{J} e^{v_k}\right) = \ln\left(\frac{1}{s_0}\right) = -\ln\left[s_0\right].$$

Thus,

$$W = \gamma - \ln\left[s_0\right]$$

(ii)

Now suppose we add a new product $j = J + 1$ into the choice set, with deterministic utility v_{J+1}. In discrete choice theory with i.i.d. Gumbel errors, it is a standard result that:

$$\max_{j=0,\ldots,J,J+1}\left[v_j + \epsilon_{ij}\right] \geq \max_{j=0,\ldots,J}\left[v_j + \epsilon_{ij}\right]$$

Hence,

$$\mathbb{E}\left[\max_{j=0,\ldots,J,J+1}\left(v_j + \epsilon_{ij}\right)\right] \geq \mathbb{E}\left[\max_{j=0,\ldots,J}\left(v_j + \epsilon_{ij}\right)\right]$$

Thus, if we denote the new welfare with product $J + 1$ as

$$W^{(\text{new })} = \mathbb{E}\left[\max_{j=0,\ldots,J+1} u_{ij}\right]$$

then

$$W^{(\text{new })} \geq W$$

In fact, with i.i.d. T1EV $(0,1)$, the new expected maximum utility is

$$W^{(\text{new })}=\gamma+\ln \left(1+\sum_{k=1}^J e^{v_k}+e^{v_{J+1}}\right)$$

and obviously

$$1+\sum_{k=1}^J e^{v_k}+e^{v_{J+1}} \geq 1+\sum_{k=1}^J e^{v_k}$$

Hence $W^{(\text{new})} \geq W$. In other words, adding a new alternative cannot reduce the expected maximum utility (as expected).

Exercise 2

Problem 1

Assuming that the variables 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.

Parameter	Estimate	StdError	tValue	PValue
Alpha	0.24	0.00	-108.02	0.00
Beta	0.29	0.01	44.58	0.00

Table 1: IV Regression Results: Estimating Alpha and Beta

Taking the ratio of product j 's share to the outside good's share and taking logs yields:

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

Denote

$$y_{jt}=\ln \left(s_{jt}\right)-\ln \left(s_{0t}\right),$$

so the linear model in the data is

$$y_{jt} = -\alpha p_{jt} + \beta x_{jt} + \xi_{jt}.$$

Because p_{jt} may be endogenous (correlated with ξ_{jt}), we use the instrument z_{jt}. The standard logit IV/ GMM moment condition is that the instrument is uncorrelated with the unobserved characteristic:

$$\mathbb{E}\big[z_{jt}\,\xi_{jt}\big] \;=\; 0.$$

Equivalently, in sample,

$$\frac{1}{NT}\sum_{j,t} z_{jt}\left(y_{jt} + \alpha\, p_{jt} - \beta\, x_{jt}\right) \;=\; 0.$$

Problem 2

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?

	Product1	Product2	Product3	Product4	Product5	Product6
Product1	-0.64	0.17	0.07	0.07	0.06	0.07
Product2	0.16	-0.64	0.07	0.07	0.06	0.07
Product3	0.16	0.17	-0.66	0.07	0.06	0.07
Product4	0.16	0.17	0.07	-0.66	0.06	0.07
Product5	0.16	0.17	0.07	0.07	-0.66	0.07
Product6	0.16	0.17	0.07	0.07	0.06	-0.66

Table 2: Average Own- and Cross-Price Elasticities (6x6)

In the homogeneous logit model:

$$s_j(\mathbf{p}) = \frac{e^{\delta_j}}{1 + \sum_{k=1}^{J} e^{\delta_k}}, \quad \delta_j = -\alpha p_j + \beta x_j + \xi_j$$

The derivative of s_j w.r.t. p_k is:

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

Hence the price elasticity of s_j with respect to p_k is:

$$\varepsilon_{j,k} = \frac{\partial s_j}{\partial p_k} \frac{p_k}{s_j} = \begin{cases} \alpha p_j \left(1 - s_j\right), & k = j, \\ \alpha p_k s_k, & k \neq j. \end{cases}$$

Implementation:

1. For each market t and each product j, plug in estimates $\hat{\alpha}$ and the observed (p_{jt}, s_{jt}).
2. Compute the 6×6 elasticity matrix (own- and cross-) for each market.
3. Average across markets.

We notice the IIA property, where cross-price elasticities are the same for all products (within product). This is a feature of the logit model, and is not always realistic.

Problem 3

Using your 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-products 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?

product	avg_mc	avg_price	avg_share
1	-1.89	3.36	0.20
2	-1.89	3.37	0.20
3	-1.56	3.03	0.09
4	-1.55	3.04	0.09
5	-1.55	3.03	0.09
6	-1.56	3.04	0.09

Table 3: Average Marginal Costs, Prices, and Shares by Product

We assume each product j is produced by a single-product firm. The standard first-order condition for a firm producing only product j in a logit demand model is:

$$\pi_j = \left(p_j - c_j\right) Q_j, \quad Q_j = M s_j.$$

Taking derivative of π_j w.r.t. p_j and setting to 0 yields:

$$s_j + (p_j - c_j)\frac{\partial s_j}{\partial p_j} = 0 \Longrightarrow p_j - c_j = -\frac{s_j}{\frac{\partial s_j}{\partial p_j}} = \frac{1}{\alpha\left[1 - s_j\right]}$$

Hence

$$c_j = p_j - \frac{1}{\alpha\left[1 - s_j\right]}$$

We use the estimated $\hat{\alpha}$ (and the observed p_{jt}, s_{jt}) to back out c_{jt}. I am not sure why marginal costs are negative here, but everything else seems to add up.

Problem 4

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

product	price_cf_avg	share_cf_avg
2	3.66	0.24
3	3.15	0.11
4	3.16	0.11
5	3.15	0.11
6	3.16	0.11

Table 4: Counterfactual Average Prices and Shares (products 2..6)

Now suppose product 1 is removed. Everything else (the product characteristics x_j the marginal cost c_j) stays the same. The other 5 products (and the outside good) re-solve the Nash-Bertrand equilibrium. We then obtain a new equilibrium. For single-product firms with the remaining $j = 2, \ldots, 6$, each firm's FOC is

$$p_j - c_j = \frac{1}{\alpha\left[1 - s_j(\mathbf{p})\right]}$$

But now the share $s_j(\mathbf{p})$ is computed from a 5 -product logit formula:

$$s_j(\mathbf{p}) = \frac{\exp\left(\delta_j\left(p_j\right)\right)}{1 + \sum_{k \in \{2,\ldots,6\}} \exp\left(\delta_k\left(p_k\right)\right)}, \quad \delta_j = -\alpha p_j + \beta x_j + \xi_j$$

This is a system of 5 equations in 5 unknowns $\{p_2, \ldots, p_6\}$, which we solve in the code with fixed-point iteration.

1. Initialize $\left\{p_2^{(0)}, \ldots, p_6^{(0)}\right\}$ at the old prices or at marginal cost + guess.
2. Compute new shares $s_j^{(k)}$ using the logit formula with the current guess $\left\{p_2^{(k)}, \ldots, p_6^{(k)}\right\}$.
3. Update each price using the best-response formula.
4. Iterate until convergence (i.e., $\max_j \left|p_j^{(k+1)} - p_j^{(k)}\right| <$ tolerance).

Problem 5

Finally, for each market compute the change in firms' profits and in consumer welfare following the exit of firm $j = 1$. Report the average changes across markets. Who wins and who loses?

product	avg_profit_base	avg_profit_cf	avg_delta_profit
1	1.08	0.00	-1.08
2	1.09	1.38	0.29
3	0.42	0.54	0.12
4	0.41	0.53	0.12
5	0.41	0.53	0.12
6	0.42	0.54	0.12

Table 5: Average Baseline vs. Counterfactual Profits by Product

Mean.CS.Change
-1.09

Table 6: Average Change in Consumer Surplus (per capita)

Profits are obtained with $\pi_j = (p_j - c_j)\, Q_j$. Change in consumer surplus is $\Delta\text{CS} \;=\; \frac{1}{\alpha}\Big[\ln\big(1+\sum_{j\in CF} e^{\delta_j^1}\big) \;-\; \ln\big(1+\sum_{j\in Base} e^{\delta_j^0}\big)\Big]$.

As expected, consumer surplus falls from higher prices and less variety. Product 1 's profit goes to zero. The remaining firms' prices, profits, and market shares increase (with product 2 firm gaining the most as the closest substitute).

3

3a

We consider *three* products: $\{0, 1, 2\}$, where

(i) $j = 0$ is the outside good, (ii) $j = 1$ and $j = 2$ are inside products in the same group.

- **Outside good ($j = 0$):** The utility for consumer i at time t is

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

where $\varepsilon_{i0t} \sim$ Gumbel with cumulative distribution

$$F_{\varepsilon_0}(\varepsilon_0) \;=\; \exp\bigl[-\exp(-\,\varepsilon_0)\bigr].$$

Moreover, ε_{i0t} is *independent* of all inside-product errors.

- **Inside group:** Products 1 and 2 have

$$u_{i1t} \;=\; \beta x_{1t} + \xi_{1t} + \varepsilon_{i1t}, \qquad u_{i2t} \;=\; \beta x_{2t} + \xi_{2t} + \varepsilon_{i2t}.$$

The pair $(\varepsilon_{i1t}, \varepsilon_{i2t})$ follows a nested-GEV distribution with parameter $0 < \rho \le 1$. Concretely:

$$F(\varepsilon_1, \varepsilon_2) \;=\; \exp\Bigl(-\bigl[\exp(-\,\varepsilon_1/\rho) + \exp(-\,\varepsilon_2/\rho)\bigr]^{\rho}\Bigr).$$

Goal: We want the probability that the consumer chooses *either* product 1 *or* product 2 (the inside group) rather than the outside good 0. That is,

$$\Pr\bigl(\max\{u_{i1t}, u_{i2t}\} \;>\; u_{i0t}\bigr).$$

Given $\varepsilon_{i0t} = \varepsilon_0$, the outside good's utility $u_{i0t} \equiv \varepsilon_0$ is a constant. Hence the event "inside group chosen" becomes

$$\bigl\{\max\{u_{i1t}, u_{i2t}\} \;>\; \varepsilon_0\bigr\}.$$

We need to evaluate this now. We have three alternatives (0,1,2).

- **Outside good** $j = 0$. Its utility is

$$u_0 \;=\; \varepsilon_0,$$

where ε_0 is Gumbel(location $= 0, \alpha = 1$) with cdf

$$F_{\varepsilon_0}(x) \;=\; \exp\bigl[-e^{-\,x}\bigr].$$

- **Inside products** $j = 1, 2$. Their utilities are

$$u_1 \;=\; \delta_1 + \varepsilon_1, \quad u_2 \;=\; \delta_2 + \varepsilon_2,$$

where $\delta_j = \beta x_j + \xi_j$ is the "observed plus unobserved quality," and $\varepsilon_1, \varepsilon_2$ follow a nested-GEV law with parameter $0 < \rho \le 1$. Concretely:

$$F(\varepsilon_1, \varepsilon_2) \;=\; \exp\Bigl(-\bigl[\exp(-\varepsilon_1/\rho) \;+\; \exp(-\varepsilon_2/\rho)\bigr]^{\rho}\Bigr).$$

- ε_0 is *independent* of $\varepsilon_1, \varepsilon_2$.

We want to show that

$$\Pr\bigl(\max\{u_1, u_2\} \;>\; u_0\bigr) \;=\; \frac{\bigl(e^{u_1/\rho} + e^{u_2/\rho}\bigr)^{\rho}}{\bigl(e^{u_1/\rho} + e^{u_2/\rho}\bigr)^{\rho} \;+\; e^{\,u_0}}.$$

$$\Pr\bigl(\max\{u_1,u_2\} > u_0\bigr) \;=\; 1 \;-\; \Pr\bigl(\max\{u_1,u_2\} \le u_0\bigr).$$

But

$$\{\max\{u_1,u_2\} \le u_0\} \;=\; \{u_1 \le u_0\} \;\cap\; \{u_2 \le u_0\}.$$

Thus,

$$\Pr\bigl(\max\{u_1,u_2\} \le u_0\bigr) \;=\; \Pr(u_1 \le u_0,\; u_2 \le u_0).$$

Because ε_0 is Gumbel and independent, write

$$\Pr(u_1 \le u_0,\; u_2 \le u_0) \;=\; \int_{-\infty}^{\infty} \Pr\bigl(u_1 \le x,\; u_2 \le x\bigr)\, dF_{\varepsilon_0}(x),$$

noting that $u_0 = \varepsilon_0$ has cdf $F_{\varepsilon_0}(\cdot)$. So

$$\Pr(u_1 \le x,\; u_2 \le x) \;=\; \Pr(\,\varepsilon_1 \le x - \delta_1,\; \varepsilon_2 \le x - \delta_2\,).$$

By the given 2-product GEV cdf,

$$F(\varepsilon_1,\varepsilon_2) \;=\; \exp\Bigl(-\bigl[\exp(-\varepsilon_1/\rho) + \exp(-\varepsilon_2/\rho)\bigr]^{\rho}\Bigr),$$

we set $\varepsilon_1 = x - \delta_1$ and $\varepsilon_2 = x - \delta_2$. Hence

$$\Pr(u_1 \le x,\; u_2 \le x) \;=\; \exp\Bigl(\,-\bigl[\exp\bigl(-(x-\delta_1)/\rho\bigr) \;+\; \exp\bigl(-(x-\delta_2)/\rho\bigr)\bigr]^{\rho}\Bigr).$$

Factor out $e^{-x/\rho}$:

$$= \exp\Bigl(\,-\, e^{-\,x}\bigl(e^{\delta_1/\rho} + e^{\delta_2/\rho}\bigr)^{\rho}\Bigr).$$

Thus

$$\Pr\bigl(\max\{u_1,u_2\} \le u_0\bigr) \;=\; \int_{-\infty}^{\infty} \exp\Bigl[-\, e^{-\,x}\bigl(e^{\delta_1/\rho} + e^{\delta_2/\rho}\bigr)^{\rho}\Bigr]\, dF_{\varepsilon_0}(x).$$

Recall $F_{\varepsilon_0}(x) = \exp[-e^{-x}]$ and ε_0 has pdf $f_{\varepsilon_0}(x) = e^{-x}\, e^{-e^{-x}}$. We want

$$\int_{-\infty}^{\infty} \exp\Bigl[-\,C\,e^{-x}\Bigr]\; d\Bigl[\exp\bigl(-\,e^{-x}\bigr)\Bigr],$$

where $C = (e^{\delta_1/\rho} + e^{\delta_2/\rho})^{\rho} > 0$. A standard result (or do a substitution $u = e^{-x}$) shows

$$\int_{-\infty}^{\infty} \exp\bigl[-C\,e^{-x}\bigr]\; d\Bigl[\exp\bigl(-e^{-x}\bigr)\Bigr] \;=\; \frac{1}{1+C}.$$

Hence

$$\Pr(\max\{u_1,u_2\} \le u_0) \;=\; \frac{1}{1+C} \quad\text{where}\quad C = \bigl(e^{\delta_1/\rho} + e^{\delta_2/\rho}\bigr)^{\rho}.$$

Therefore,

$$\Pr\bigl(\max\{u_1,u_2\} > u_0\bigr) \;=\; 1 - \frac{1}{1+C} \;=\; \frac{C}{1+C}.$$

Rewriting C and exponentiating in terms of $u_j = \delta_j + \varepsilon_j$, we can equivalently write

$$\boxed{\Pr\big(\max\{u_1, u_2\} > u_0\big) \;=\; \frac{\big(e^{\delta_1/\rho} + e^{\delta_2/\rho}\big)^{\rho}}{\big(e^{\delta_1/\rho} + e^{\delta_2/\rho}\big)^{\rho} + 1}}.$$

Extension to General Case Adopt the $\delta_j = x_{jt} + \xi_{jt}$ notation. Our goal will be to show this equals

$$P(i\ chooses\ \mathcal{B}_g) = \frac{[\sum_{j \in \mathcal{B}_g} \exp\big(\rho^{-1}\delta_j\big)]\big)^{\rho}}{\sum_{h=1}^{G}[\sum_{k \in \mathcal{B}_h} \exp(\rho^{-1}\delta_k]^{\rho} + 1)} \tag{1}$$

The way the generalization works is:

1. Repeat the 1 nest 2-good case with 1-nest N-good, our argument above in no way depended on having exactly 2 goods, so the N good case proceeds identically.
2. Now, define a "nest-specific outside option", which has probability of being chosen: $\frac{K_{-g}}{K_{-g}+[\sum_{j \in \mathcal{B}_g} \exp \rho^{-1}\delta_j])^{\rho}}$, where K_{-g} may depend on features of the other nests, but not on g, i.e. we relabel the "outside option" as "any other nest besides g, or the outside option"
3. Now, note that the probability of a (non-outside good) from nest $\mathcal{B}_g$ being chosen must be proportional to $[\sum_{j \in \mathcal{B}_g} \exp \rho^{-1}\delta_j])^{\rho}$ $\forall g$ But if both (i) the probability of each nest is proportional to $[\sum_{j \in \mathcal{B}_g} \exp \rho^{-1}\delta_j])^{\rho}$ (relative to the outside good), and (ii) probabilities of all possible choices must sum to 1, then it must be that, for any given nest, $K_{-g} + K_g = \sum_{h=1}^{G}[\sum_{k \in \mathcal{B}_h} \exp\big(\rho^{-1}\delta_k]^{\rho} + 1\big)$, which yields the denominator for the final expression
4. Hence, combining the steps, the final probability is as in Equation (1).

3b

Here, the goal is to show that, given the max is inside nest B_g, we are back to the multinomial logit case. Particularly, we want to show that the final answer is

$$P(i\ chooses\ j|i\ Chooses\ B_g, j \in B_g) = \frac{\exp(1/\rho(\delta_j))}{\sum_{k \in B_g} \exp(\delta_k/\rho)}$$

Conditioning on the maxmax $beinginside B_g B_g, butnotknowingwhatthe$ max max $is, isakintointegratingout$ B_g as

$$F(\epsilon_j | j \in B_g) = \exp(-\exp(\epsilon/\rho))$$

i.e. a type 1 extreme value distribution with shape parameter, $1/\rho$, which, by properties of the extreme value type 1 distribution, yields what we wanted to show:

$$P(i\ chooses\ j|i\ Chooses\ B_g, j \in B_g) = \frac{\exp(1/\rho(\delta_j))}{\sum_{k\in B_g}\exp(\delta_k/\rho)}$$

3c

Mechanical from 3a and 3b,

By law of total probability

Hence the unconditional probability of choosing product j (which belongs to nest B_g) is

$$\frac{\exp(1/\rho(\delta_j))}{\sum_{k\in B_g}\exp(\delta_k/\rho)}\frac{(\sum_{j\in B_g}\exp\big(\rho^{-1}(\delta_j)\big))^{\rho}}{\sum_{h=1}^{G}(\sum_{k\in B_h}\exp(\rho^{-1}(\delta_k))^{\rho}+1}$$

$$\exp(\delta_j/\rho)\frac{(\sum_{k\in B_g}\exp\big(\rho^{-1}(\delta_k)\big))^{\rho-1}}{\sum_{h=1}^{G}(\sum_{l\in B_h}\exp(\rho^{-1}(\delta_l))^{\rho}+1}$$

3d

As is usual when we see a $\log s_{jt} - \log s_{0t}$ we must be cancelling something to get a nice expression

$$P(i\ chooses\ j) = P(i\ choosesj|i\ chooses\ B_g)P(i\ chooses\ B_g)$$

Plugging in the approximations and the equation for $P(i\ chooses\ B_g)$

$$s_{jt} \approx s_{jt|gt}\frac{[\sum_{j\in\mathcal{B}_g}\exp\big(\rho^{-1}\delta_j\big)]^{\rho}}{\sum_{h=1}^{G}[\sum_{k\in\mathcal{B}_h}\exp(\rho^{-1}\delta_k)]^{\rho}+1}$$

Now for outside option, $P(i\ chooses\ j = 0)$, we know that $s_{0t|0t} = 1$, so we have that $P(i\ chooses\ j = 0) = P(i\ chooses\ g = 0)$

$$s_{0t} \approx (1)\frac{1}{\sum_{h=1}^{G}[\sum_{k\in\mathcal{B}_h}\exp(\rho^{-1}\delta_k)]^{\rho}+1}$$

Then we will do s_{jt}/s_{0t} (later taking logs to get the desired thing)

$$\frac{s_{jt}}{s_{0t}} = s_{jt|gt}[\sum_{k\in\mathcal{B}_g}\exp\rho^{-1}\delta_k])^{\rho}$$

Let's see what happens when we take logs.

$$\log s_{jt} - \log s_{0t} = \log s_{jt|gt} + \rho \log[\sum_{k \in \mathcal{B}_g} \exp \rho^{-1} \delta_k]$$

we should have $s_{jt}|gt$ in our answer. this suggests we might want to add and subtract it cancel out that final term

let's add and subtract $\rho \log s_{jt|gt}$

$$\log s_{jt} - \log s_{0t} = \log s_{jt|gt} - \rho \log s_{jt|gt} + \rho \log[\sum_{k \in \mathcal{B}_g} \exp \rho^{-1} \delta_k] + \rho \log s_{jt|gt}$$

By property of logs and the approximation for $s_{jt|gt}$

$$\log s_{jt} - \log s_{0t} = \log s_{jt|gt} - \rho \log s_{jt|gt} + \log[(\sum_{k \in \mathcal{B}_g} \exp \rho^{-1} \delta_k)^\rho] + \log[\frac{\exp(\delta_j/\rho)^\rho}{(\sum \exp(\delta_k/\rho)^\rho}]$$

properties of logs

$$\log s_{jt} - \log s_{0t} = \log s_{jt|gt} - \rho \log s_{jt|gt} + \log[(\sum_{k \in \mathcal{B}_g} \exp \rho^{-1} \delta_k)^\rho][\frac{\exp(\delta_j/\rho)^\rho}{(\sum \exp(\delta_k/\rho)^\rho}]$$

$$\log s_{jt} - \log s_{0t} = \log s_{jt|gt} - \rho \log s_{jt|gt} + \log[\exp(\delta_j/\rho)^\rho]$$

Properties of logs (bringing exponent out front and canceling log with exp)

$$\log s_{jt} - \log s_{0t} = \log s_{jt|gt} - \rho \log s_{jt|gt} + \rho(\delta_j/\rho)$$

by defn of δ_j, and grouping like terms

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

□

3e

Regression form would look like:

$$\log s_{jt} - \log s_{0t} = (1-\rho) \log s_{jt|gt} + x_{jt} + \xi_{jt} + \nu_{jt}$$

This endogeneity problem is an issue because, even if ξ_{jt} is uncorrelated with x_{jt}, it will be correlated with $s_{jt|gt}$ (and s_{jt}), which we do not observe.

Table 7: OLS regression

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
Intercept	0.9056	0.024	37.520	0.000	0.858	0.953
log_s_j_g	0.6494	0.003	186.612	0.000	0.643	0.656
x_ji	1.4932	0.017	88.847	0.000	1.460	1.526

Table 8: IV Estimates

	Parameter	**Std. Err.**	**T-stat**	**P-value**	**Lower CI**	**Upper CI**
Intercept	-0.0369	0.0628	-0.5883	0.5563	-0.1600	0.0861
x_ji	2.0256	0.0376	53.865	0.0000	1.9519	2.0993
log_s_j_g	0.4937	0.0101	48.834	0.0000	0.4739	0.5135

3f

The exercise here will be to randomly generate data, where each row is indexed by a market t, product j, and nest $\mathcal{B}_g$ where $j \in \mathcal{B}_g$.

Pseudocode:

1. In total, we assume 100 markets, t, 10 nests, with (1 ... 10) products each, so 55 products total, $\beta = 2, \rho = 0.5, \mu_x = 0, \sigma_x = 1, \mu_\xi = 0, \sigma_\xi = 1$
2. x and ξ are simulated at the product-level with above parameters, normal distribution, separately for each market, t
3. With β and ρ, we compute the approximate shares needed for the regression.
4. We run the regression shown in 3e with OLS, and find we cannot recover the parameters $(\beta, \rho) = (2, 0.5)$
5. We run the regression shown in 3e with the suggested instrument and find we can get quite close

4

4.1

See code.

Our objective function seems to bottom out without hitting zero, which is consistent with us being overidentified (we may not be able to hit all 6 moment conditions with the 5 parameters in Γ and β).

Estimates for

$$\Gamma = \begin{bmatrix} 4.656 & 0 \\ -0.1222 & 0.00143 \end{bmatrix}$$

And thus,

$$\Gamma\Gamma' = \Omega = \begin{bmatrix} 21.686 & -0.5692 \\ -0.5692 & 0.0149 \end{bmatrix}$$

Estimates for β are:

$$[-1.73, -0.6149]$$

where β_1 is the coefficient on price, and β_2 is the coefficient on x. With more compute-time we could calculate standard errors on all of these with bootstrapping.

The negative coefficient on price is a good sign. It is a little puzzling that x doesn't have a positive sign if it's "quality" (as described in 4.3)

	Product1	Product2	Product3	Product4	Product5	Product6
Prod1	-0.001954	0.000338	0.017130	0.013033	0.063578	0.081298
Prod2	0.000486	-0.001947	0.016972	0.012885	0.065346	0.080284
Prod3	0.000276	0.000168	-1.251420	0.073299	0.477741	0.601405
Prod4	0.000315	0.000169	0.090877	-1.231844	0.499488	0.559553
Prod5	0.000176	0.000153	0.133663	0.057004	-1.108132	0.747235
Prod6	0.000158	0.000087	0.067712	0.074689	0.525427	-0.900268

Table 9: Average own and Price. In $(row = i, column = j)$ is the share elasticity of product i to a change in price of j

Table 10: Average Price, x, and Share by Product

product	avg_price	avg_x	avg_share
0	0.0000	0.0000	0.3848
1	0.0024	-0.0193	0.0988
2	0.0023	-0.0260	0.0891
3	2.0191	-0.0813	0.0430
4	1.7516	-0.1801	0.0393
5	3.5770	1.6926	0.1517
6	4.4429	2.0024	0.1932

4.2 elasticities

We must do numerical integration for this step. Particularly, since we don't have demographics, but we do have specific values for each v_i, the equation given in class reduces to the following, where in our case, $|I| = 50$, and we have expressed the dependence of shares on v_i with an i subscript.

$$\epsilon_{jk,t} = \frac{\partial s_{j,t}}{\partial p_{k,t}} \frac{p_{k,t}}{s_{j,t}} = \begin{cases} -\frac{p_{j,t}}{s_{j,t}} \sum_i s_{i,j,t}(1 - s_{i,j,t}) \frac{1}{|I|} & \text{if } j = k \\ \frac{p_{k,t}}{s_{j,t}} \sum_i s_{i,j,t} s_{i,k,t} \frac{1}{|I|} & \text{otherwise.} \end{cases}$$

As expected, they're all still negative. But now all columns need not have the same number as in 2.2. That's part of what we get from breaking IIA. Own-price elasticities are always bigger in magnitude than cross-price elasticities, which is also good to see. There also generally seems to be more heterogeneity across products (e.g. wider dispersion along the diagonal).

4.3

Γ tells us how someone's idiosyncratic shocks v_i interact with the price and observed quality x to influence utility. To think about this step carefully, it is useful to show how the random coefficients augment the homogeneous ones:

$$(\beta_1 p_{jt} + p_{jt}\,\gamma_{11}\,v_{i1}) \;+\; (\beta_2 x_{jt} + x_{jt}\,\gamma_{21}\,v_{i1} \;+\; x_{jt}\,\gamma_{22}\,v_{i2})$$

Table (10) shows that share and price are perfectly positively correlated with x once we exclude the outside good, which seems inconsistent with our negative coefficient on x. But from this simple table, x could be driving market share.

There is also potentially a big importance for heterogeneity:

- $\gamma_{1,1}$ being so large tells us that heterogeneous coefficients on price (represented by the upper corner), could swamp the common portion. If someone's v_{i1} is 1 SD above the mean (i.e. equals 1) their coefficient on price could actually swing positive.
- $\gamma_{2,1}$ being negative tells us that when v_{i1} is large, utility from x falls. So the same shock that makes them less price sensitive, also makes them care less about x.
- $\gamma_{2,2}$ is small relative to the mean values of x, meaning that the important part of heterogeneity is v_{i1}.

So, if we trust this estimate for Γ, it might be that shares and prices are primarily determined by x and their v_{i1} shock.

4.4 pyBLP

I am assuming we should use the following formulations since we permit p and x to have heterogeneous coefficients, but there is no constant in the utility (and no heterogeneity in preference for the outside good).

- X1 formulation: pyblp.Formulation('0+ prices + x') (no constant)
- X2_formulation = pyblp.Formulation('0 + prices + x') (no heterogeneous preference for the outside option)
- We simulate 50 values of v_i for the random shocks.

$$\Gamma_{pyBLP} = \begin{bmatrix} 0.0264 & 0 \\ -0.0300 & 0.0040 \end{bmatrix}.$$

$$\beta_{pyBLP} = [0.28537, -0.697]$$

The value on β_p is not the same sign as ours. Our value for β_x is pretty similar, and the signs of all the entries in the Γ matrix are the same. The scale of terms seems to be smaller in pyBLP.