

## Exercise 1

### Problem 1

Let  $X_i$  with  $i = 1, \dots, N$  be a sequence of independent type 1 extreme value random variables with location parameter  $\mu_i$  and scale parameter  $\sigma > 0$  (T1EV( $\mu_i, \sigma$ )). 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,

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

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  $\mu_x$  and  $\mu_y$  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

$$Z = X - Y \implies Z \sim \text{Logistic}(\mu_X - \mu_Y, \sigma)$$

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  $\mu_j$  is non-random and  $\epsilon_j$  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

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

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  $j = 1, \dots, J$ , 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:

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

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

- (a) Assume  $\epsilon_{ij}$  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  $\epsilon_{ij}$  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  $\beta_i$  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  $\beta_i$  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,  $\beta_i$  represents an idiosyncratic taste shock for that same characteristic and  $v_{ij}$  are i.i.d T1EV(0,1). Moreover, assume that  $\beta_i$  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  $\epsilon_{ij}$  are i.i.d T1EV(0,1). Moreover, suppose we want to measure welfare at given prices  $(p_1, \dots, p_J)$  as

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

- (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

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

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

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

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

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

and therefore

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

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.

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

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

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

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

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

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

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

For the other products, we have

$$-\alpha p_j + \beta x_j \geq -\alpha p_k + \beta x_k, \implies \beta (x_j - x_k) \geq \alpha (p_j - p_k)$$

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

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

I'm not sure if I can write 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.

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

Conditional on  $\beta_i$ , the consumer solves the first T1EV problem,

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

so

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

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

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

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

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

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  $\beta$  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

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

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

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

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,

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

Thus,

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

(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:

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

Hence,

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

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

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

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  $(\alpha, \beta)$  and obtain an estimate for both parameters.

Parameter	Estimate	StdError
Alpha	0.24	0.00
Beta	0.29	0.01

**Table 0.0.1:** 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

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

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  $\xi_{jt}$ ), we use the instrument  $z_{jt}$ . The standard logit IV/GMM moment condition is that the instrument is uncorrelated with the unobserved characteristic:

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

Equivalently, in sample,

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

## 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 0.0.2:** 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:

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

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

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

### 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  $6 \times 6$  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 0.0.3:** 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:

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

Taking derivative of  $\pi_j$  w.r.t.  $p_j$  and setting to 0 yields:

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

Hence

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

We use the estimated  $\hat{\alpha}$  (and the observed  $p_{jt}, s_{jt}$ ) to back out  $c_{jt}$ .

### 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.

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.

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 0.0.4:** Counterfactual Average Prices and Shares (products 2-6)

New Equilibrium Conditions

For single-product firms with the remaining  $j = 2, \dots, 6$ , each firm's FOC is

$$p_j - c_j = \frac{1}{\alpha [1 - s_j(\mathbf{p})]}, \quad j = 2, \dots, 6.$$

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

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

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

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

#### 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 0.0.5:** Average Baseline vs. Counterfactual Profits by Product

Profits:  $\pi_j = (p_j - c_j) Q_j$ . 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).

Mean.CS.Change
-1.09

**Table 0.0.6:** Average Change in Consumer Surplus (per capita)

$$\text{Consumer Surplus: } \Delta\text{CS} = \frac{1}{\alpha} \left[ \ln(1 + \sum_{j \in CF} e^{\delta_j^1}) - \ln(1 + \sum_{j \in Base} e^{\delta_j^0}) \right].$$

Consumer surplus falls from higher prices and less variety.