Problem Set 1

ECON 40201 - Advanced Industrial Organization II

DUE: January 31st, 2025

Please typeset your assignment in LaTeX. For the coding part, feel free to use your preferred programming language. Please comment your code clearly, as the clarity of your code will also be evaluated. Submit both your write-ups and codes through Canvas in a single zip file. You can work in a group of two (please clearly indicate both authors in your submitted PDF) and submit a single file per group.

Exercise 1. Derivation of Basic Expressions

1. Let X_i with $i = 1, ..., N$ 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:

$$\Pr\{X_i \leq x | \mu_i, \sigma\} = \exp\Big(- \exp\Big(- \frac{x - \mu_i}{\sigma} \Big) \Big)$$

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

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 $X - Y$.

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.

4. Consider a market with J products indexed by $j = 1, ..., J$, 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:

$$u_{ij} = \alpha(y_i - p_j) + \epsilon_{ij} \quad \text{for} \quad j = 1, ..., J$$

$$u_{i0} = \alpha y_i + \epsilon_{i0} \quad \text{for} \quad j = 0$$

   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 $[0, \overline{\beta}]$ with $\overline{\beta}$ 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 $(p_1, ..., p_J)$ as

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

   (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.

Exercise 2. Homogeneous Logit, Costs, and Counterfactuals

The file ps1_ex2.csv contains aggregate data on a large number $T = 1000$ of markets in which $J = 6$ products compete between each other together with an outside good $j = 0$. The utility of consumer i is given by:

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

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

where p_{jt} is the price of product 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 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.

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?

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?

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.

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?

Exercise 3. Nested Logit and Instruments

One of the possible ways to relax the restrictive substitution pattern from 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 $g \in \{1, 2, ..., G\}$. 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 $\{\epsilon_{ijt}\}$ is of the form

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

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

$$F(\epsilon_0) = \exp\Big(- \exp(-\epsilon_0) \Big)$$

for all i and t.

1. Derive the expression for $\Pr(i \text{ Chooses } B_g)$.

   (*Hint*: You may start by considering a simpler case when there is only one group with two product 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.)

2. Derive the expression for $\Pr(i \text{ Chooses } j \mid i \text{ Chooses } B_g)$ when $j \in B_g$.

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

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

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

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

5. Given the relationship derived in the previous part, some researchers have suggested that we can estimate the model parameters by regressing $\log(s_{jt}) - \log(s_{0t})$ 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})?

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 from 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 you assumed in the data generation process.

Exercise 4. BLP Implementation

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

$$u_{ijt} = \tilde{x}'_{jt}\beta + \xi_{jt} + \tilde{x}'_{jt}\Gamma v_i + \epsilon_{ijt}, \quad j = 1, ..., 6$$

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

where 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 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 (α, Γ). Finally, you can assume that Γ is lower triangular e.g.,

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

such that $\Gamma\Gamma' = \Omega$ is a positive definite matrix and that v_i is a 2 dimensional vector of i.i.d random taste shocks.

1. Assume that $v_i \sim N(0, I)$ so that $\Gamma v_i \sim N(0, \Omega)$. 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.

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's the main difference when compared with the table of elasticities you found in 2.2?

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

4. Compare your results with PyBLP (this does not need to be done in Python - you can call PyBLP from R, Matlab, and Julia).