Assignment: Estimating Static Games

Industrial Organization

1. *Retail.* Please use whatever package you like (e.g., Stata) to clean the data, but please do all analysis using MATLAB (or R or python if you prefer).

   This question is based on the code and data posted for replication here for a study of discount retail by Jia (2008). If you are not already familiar with the paper, don't become familiar (at least not until after you've submitted your work). It's better to approach this question with a fresh perspective.
   We will work with the file `XMat.out`, in the folder `\Data\1988`. This contains data on the number of retail establishments by county in the US, along with some county characteristics. The file `DataDocumentation.doc` contains variable definitions.
   Assume that each county i is a distinct market with population S_i and let $N_i \in \{0,1,2\}$ denote the number of big-box discount retailers (Kmart or Walmart) in the county. Assume that profits to a discount retailer in county i are given by

   $$\pi(S_i, N_i, F_i) = S_i v(N_i) - F_i$$

   where $v()$ is per-capita variable profit, strictly decreasing in its argument, and F_i is a fixed cost distributed i.i.d. normal with mean μ and standard deviation σ across counties. Importantly, this model treats Walmart and K-mart as identical and ignores other retailers. The model also ignores the fact that Walmart and Kmart may coordinate their entry decisions across markets. (These issues are addressed in detail in Jia's study.)

   (a) Estimate the parameters of the model.

   (b) Offer an economic interpretation of each parameter's magnitude.

   (c) Re-estimate the model, but now assume that the mean of fixed costs is μ_1 for the first entrant and $\mu_2 \geq \mu_1$ for the second entrant. Report the parameter values you estimate.

   (d) How does your estimate of $v()$ change between (a) and (c)? Why?

   (e) Return to the setting in part (a) but drop the assumption that F_i is distributed normally. Instead, assume that F_i is i.i.d. with a smooth CDF. Estimate $v()$.

   (f) How do your results in (e) compare to those in (a)? What does this suggest about the importance of the normality assumption on F_i?

2. *Type.* This question is based on Mazzeo (2002). You do not have to read the paper to solve the problem but it will give you some context about how this type of model can be applied.

There is a set of markets indexed by m. In each market there are two firms, 1 and 2. The firms play the following game. Firm 1 chooses a binary attribute $x_{m1} \in \{0,1\}$, after which firm 2 observes x_{m1} and chooses $x_{m2} \in \{0,1\}$. The firms realize payoffs given by

$$\pi_{mi} = \eta_m \left(x_{mi} - \frac{1}{2} \right) - \theta \left(1 - |x_{m1} - x_{m2}| \right)$$

where η_m is distributed i.i.d. across markets m according to distribution $\tilde{\Phi}$, which is symmetric around 0. We may think of $\theta \geq 0$ as the incentive to differentiate from one's competitor. We may think of η_m as a market-level shock to consumer preferences for firm attributes. The shock η_m is common knowledge to the firms in market m.

(a) Characterize the equilibrium of the game for a given market m as a function of θ and η_m.

(b) Let $\rho \equiv \mathrm{Cov}\left(x_{m1}, x_{m2}\right) / \sqrt{\mathrm{Var}\left(x_{m1}\right) \mathrm{Var}\left(x_{m2}\right)}$ denote the population value of the correlation between the choices of firms 1 and 2. Characterize ρ as a function of $\tilde{\Phi}$ and θ. Provide an intuition.

(c) Is θ nonparametrically identified from ρ?

(d) Suppose that η_m is distributed i.i.d. standard normal. Derive the maximum likelihood estimate of θ.