

## Lista de Exercícios 2:

### Modelos de Escolha Discreta com Dados Agregados

1. Nos arquivos `berrydata1.xls` e `berrydata2.xls` os senhores podem encontrar dados para 500 mercados simulados sobre quantidade (Q), preços (P), características dos produtos (X) e preços de fatores de produção (W). Os dois conjuntos de dados diferem na importância atribuída às características não observadas dos produtos – que denominaremos  $\xi$ . Os dados foram gerados tendo por base um mercado potencial de 1.000.000 de habitantes, que as empresas se engajam em competição à la Nash-Bertrand, e que a demanda segue um padrão de escolha discreta multinomial. A utilidade média do consumidor é  $\delta_{ij} = 5 - p_{ij} + 2x_{ij} + \sigma\xi_{ij}$ , sendo que  $\xi \sim N(0, 1)$  e  $i$  denota o mercado e  $j$  o produto. **Note que, como  $\xi$  é uma característica não observável, não temos dados sobre isso.** A diferença entre os dois conjuntos de dados é que, no primeiro caso, 80% da variação é explicada pelas características observadas  $\sigma = 1$ . No segundo conjunto de dados, apenas 30% da variação é explicada pelas características observadas  $\sigma = 3$ . Pergunta-se o seguinte (use o STATA para este exercício), para cada um dos conjuntos de dados:
  - (a) Converta as quantidades em probabilidades de compra (isto é, *shares*), e construa médias e desvio-padrão para todas as variáveis. Mostre a matriz de correlação entre estas variáveis. Ela parece razoável?
  - (b) Estime a demanda usando OLS. Em que direção a endogeneidade dos preços deve estar viesando as suas estimativas? Isso faz sentido de acordo com os seus resultados?
  - (c) Estime a demanda usando os custos específicos da empresa como IV para os preços, inicialmente, e depois as características dos produtos da empresa competidora como IV (use médias para gerar as variáveis). Como as suas estimativas se relacionam com as em OLS? E entre elas? (Tanto estimativas pontuais como erros-padrão).
  - (d) Existem outras IV potenciais? Se existem, qual é a sua intuição acerca do desempenho delas?

- (e) Estime a demanda usando os dois instrumentos potenciais discutidos no item (c). Como eles se comparam em relação aos outros? Calcule um teste de *overidentifying restrictions*. As duas IV's são válidas?
  - (f) Como os resultados entre as duas bases de dados se comparam? O que você pode concluir destas diferenças?
2. Um dos grandes motivos para o sucesso dos métodos de simulação é a existência de códigos para o MATLAB e Python facilmente disponíveis na internet. Neste exercício, quero que os senhores respondam o seguinte, usando o PyBLP:
- (a) Repliquem a tabela I da página 539 de Nevo (2000) "A Research Assistant's Guide for the Random Coefficients Models"

**Observação sobre lista de exercícios:**

Esta lista de exercícios pode ser entregue em até três pessoas. Gostaria que os resultados fossem apresentados em um texto word, em que fossem apresentadas as tabelas com os resultados relevantes APENAS, e os programas utilizados para a elaboração das respostas. Qualquer software econométrico é adequado desde que vcs anexem os respectivos códigos.