

### Prova Intermediária:

#### Modelos de Escolha Discreta com Dados Agregados e Função de Produção

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 facilmente disponíveis na internet. Neste exercício, quero que os senhores entrem na página de Brownyn Hall<sup>1</sup>, baixem os códigos ali disponíveis, bem como os dados e respondam o seguinte:
- (a) Repliquem a tabela I da página 539 de Nevo (2000) "A Research Assistant's Guide for the Random Coefficients Models"
  - (b) Expliquem detalhadamente o funcionamento das rotinas do MATLAB presentes naquela página.

### Material para a segunda parte da Prova

No arquivo `flatdata.raw` temos 531\*6 linhas e seis colunas de dados. Cada linha se refere a uma empresa em um ano. Existem seis anos de dados, de forma que as linhas de 1-6 se referem aos seis anos para a primeira empresa, as seis seguintes para a segunda empresa, e assim por diante. As colunas são as seguintes: produto, idade, capital, trabalho e investimento. Se os valores para a empresa são zero em um determinado ano, isto significa que a firma ainda não existia, ou saiu do mercado ou ainda não entrou. Para checar os seus dados, a variável de idade deve variar de 1 até 112 nos seus dados, ser diferente de zero em 1616 observações e ter uma média de 7,84 e um desvio padrão de 11,84. Use a amostra para responder às seguintes perguntas:

- 3. Usando os dados, replique as colunas de 1 a 5 na tabela VI de Olley e Pakes (1996) para esta amostra de dados. Tenha certeza de passar os logaritmos nas variáveis que utilizar. Discuta detalhadamente quais são os possíveis vieses que podem surgir em resposta ao seu modelo econômico subjacente à estimação. Sob quais hipóteses as estimativas de cada um destes modelos é consistente? Existem evidências que estas hipóteses são violadas?
- 4. Agora, usando um polinomial de quarta ordem como em Olley e Pakes (1996), etapa 1, obtenha uma estimativa consistente de  $\beta_L$ . Tenha certeza em reportar exatamente como você fez isso (mas não precisa reportar os coeficientes do polinomial).

---

<sup>1</sup>[http://emlab.berkeley.edu/users/bhhall/e220c/rc\\_dc\\_code.htm](http://emlab.berkeley.edu/users/bhhall/e220c/rc_dc_code.htm)

- O valor do coeficiente estimado e o ajustamento do modelo mudam muito quando mudamos de terceira para quarta ordem?
  - O valor do coeficiente muda muito quando usamos diferentes polinômios em cada um dos anos dos dados?
  - Você confia pouco ou muito na sua estimativa e porquê? Quando comparado com os resultados dos itens anteriores, a sua nova estimativa está de acordo com os vieses que foram sugeridos pelo modelo teórico?
5. Estime os coeficientes  $\beta_K$  e  $\beta_A$  usando o método em três estágios de Olley e Pakes (1996) (coluna 8 na tabela VI). Quando estimar, o segundo estágio, o PROBIT para a sobrevivência da empresa, você deve utilizar um polinomial de quarta ordem – mais uma vez, reporte exatamente o que você fez, mas sem precisar fazer a alternativa de terceira ordem como comparação. Na terceira etapa (mínimos quadrados não lineares), também use um polinomial de terceira ordem. Discuta as suas estimativas comparadas com o que você observou no item 1.
6. Produtividade:
- Calcule o crescimento da produtividade para cada empresa conforme o discutido por Olley e Pakes (1996), seção 5. Qual seria uma forma alternativa de se calcular a produtividade? Mostre estatísticas descritivas para as duas medidas.
  - Calcule o crescimento da produtividade agregada para cada ano, usando a participação no produto como medida de agregação (faça a mesma medida de produtividade que Olley e Pakes (1996)).
  - Faça a mesma decomposição que eles e replique a tabela XI para a nossa amostra de dados. O que vocês podem concluir acerca de realocação de participação do produto versus crescimento das plantas?

#### **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; se vocês usarem o Eviews, ANEXEM O PROGRAMA (extensão .prg) . O workfile não é suficiente.