Demand Estimation 4
PhD Industrial Organization

Nicholas Vreugdenhil

Plan

1. Estimation algorithm
2. Instrumental variables
3. Extensions
4. Applications

Plan

1. **Estimation algorithm**
2. Instrumental variables
3. Extensions
4. Applications

Review: setup of the problem

- (Conditional, indirect) utility:

$$u_{ijt} = x_{jt}\beta_{it} + \alpha_{it}p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

  - (Note: first element of x_{jt} is 1 $\rightarrow$ absorbs mean of ξ_{jt})

- What are the parameters we need to estimate?

- Linear parameters:
  - Parameters from the mean utility equation: (α_0, β_0)

- Nonlinear parameters
  - Γ: coefficients on (observed) demographics
  - Σ: idiosyncratic "taste for characteristics"

- So, full parameter vector to estimate: $\theta = (\alpha_0, \beta_0, \Gamma, \Sigma)$.

Review: estimation algorithm: overview

- Step 1: For a guess of Γ and Σ, and a vector of mean utilities $\boldsymbol{\delta}_t$, compute model-predicted market shares.

- Step 2: For a guess of Γ and Σ do an **inversion**: find $\boldsymbol{\delta}_t$ where the model-predicted market shares match the empirical market shares s_t.
  - This step will repeatedly call the function from Step 1.

- Step 3: Use the computed $\boldsymbol{\delta}_t$ from Step 2 to compute $\xi_{jt} = \delta_{jt}(\Gamma, \Sigma) - x_{jt}\beta_0 - \alpha_0 p_{jt}$.
  - Interact with IVs to get the GMM objective function.
  - Search over all parameters θ to minimize objective function using non-linear optimization.

Estimation algorithm: step 3

- Denote the mean utilities from step 2: $\delta_{jt}(\Gamma, \Sigma)$
- Compute $\xi_{jt}(\theta) = \delta_{jt}(\Gamma, \Sigma) - x_{jt}\beta_0 - \alpha_0 p_{jt}$
  - Above equation is why we called Γ, Σ 'nonlinear' variables, and β_0, α_0 the 'linear variables'
- Interact with the instrumental variables to get the GMM objective function (denoting W as the GMM weight matrix):

$$\xi(\theta)' ZWZ' \xi(\theta)$$

- Solve for the parameters using nonlinear optimization.

$$\hat{\theta} = arg \min_{\theta} \xi(\theta)' ZWZ' \xi(\theta)$$

- Note: since this is just a GMM problem, can also get standard errors using standard GMM methods

Numerical issues (documented by Knittel and Metaxoglou (2014))

- (See Conlon and Gortmaker (2020) for latest updates on best practices.)
- 1. Objective function is highly nonlinear with many local minima
  - Numerical results can be sensitive to starting values or choice of optimizer method
  - (Partial) solution: test results with different starting values and optimizer methods
  - (Partial) solution: choose an optimizer that is used for commercial purposes (e.g. Knitro)
  - (Partial) solution: Conlon and Gortmaker (2020) make some suggestions of free optimizer methods that work well in Scipy (a Python library)
  - Solution (probably not yet computationally feasible): use a global optimizer like 'differential evolution'
- 2. Need to choose very tight convergence tolerances for the inversion ($< 10^{-12}$)
- **These are common issues in structural models, so be on the lookout in other contexts.**

Alternative algorithm: MPEC ('Mathematical Programming With Equilibrium Constraints')

$$\min_{\theta,\xi} \quad \xi' Z W Z' \xi$$

$$\text{subject to} \quad \tilde{\sigma}(\delta(\xi); x, p, \hat{F}, \theta) = s$$

- Notice minimization is over both θ and ξ here.
- Advantage of this approach:

Alternative algorithm: MPEC ('Mathematical Programming With Equilibrium Constraints')

$$\min_{\theta,\xi} \quad \xi' Z W Z' \xi$$

$$\text{subject to} \quad \tilde{\sigma}(\delta(\xi); x, p, \hat{F}, \theta) = s$$

- Notice minimization is over both θ and ξ here.
- Advantage of this approach: No need for inversion step
- Disadvantage of this approach:

Alternative algorithm: MPEC ('Mathematical Programming With Equilibrium Constraints')

$$\min_{\theta,\xi} \quad \xi' Z W Z' \xi$$
$$\text{subject to} \quad \tilde{\sigma}(\delta(\xi); x, p, \hat{F}, \theta) = s$$

- Notice minimization is over both θ and ξ here.
- Advantage of this approach: No need for inversion step
- Disadvantage of this approach: Many more parameters to solve for (ξ)
- Dube et al (2012): claim this approach results in a speedup.
  - However, can be complicated to program, and some have found it slow for large problems

Plan

1. Estimation algorithm
2. **Instrumental variables**
3. Extensions
4. Applications

Instruments

- We used the following moment conditions restriction:

$$E(\xi_{jt}|\mathbf{Z}_t) = 0$$

- Here, $\mathbf{Z}_t$ is a vector of instruments
- Note that above assumption implies a large set of potential IVs $z_{jt} = A_j(\mathbf{Z}_t)$ for which the unconditional moment restriction holds:

$$E(z_{jt}\xi_{jt}) = 0$$

- What instruments should we use for $\mathbf{Z}_t$?
  - Recall the dual role for instruments in the model.

Instruments: BLP instruments

- Use **characteristics of products in the market**.

$$E(\xi_{jt}|\mathbf{x}_t) = 0$$

  - $\mathbf{x}_t$ is the vector of product characteristics in market t
  - (Note: don't include price in characteristics)
  - 'Observed characteristics mean independent of unobserved characteristics'
- Can form many moment conditions from above assumption. BLP use:
  - characteristics of own product
  - sum of characteristics of other products produced by the firm
  - sum of characteristics of competitiors

Instruments: BLP instruments

- Power: closeness in characteristic space affects markups which affects price.

- Justification for this instrument:

Instruments: BLP instruments

- Power: closeness in characteristic space affects markups which affects price.
- Justification for this instrument: **product characteristics are set before ξ_{jt} known**.
  - Do you think this is a reasonable assumption in the car market?

Instruments: BLP instruments

- Power: closeness in characteristic space affects markups which affects price.
- Justification for this instrument: **product characteristics are set before ξ_{jt} known**.
  - Do you think this is a reasonable assumption in the car market?
- What if firms are forward looking and anticipate ξ_{jt} when choosing product characteristics?
  - Possible solution: use panel data.
  - E.g. Sweeting (2013) assumes $\xi_{jt} = \rho\xi_{jt-1} + u_{jt}$, where u_{jt} unanticipated at time $t-1$
  - Implies moments: $E(\xi_{jt} - \rho\xi_{jt-1}|x_{t-1}) = 0$.
  - Comment: many connections here to the dynamic panel / production function literatures.

Instruments: cost-based / Hausman instruments

- Ideal instrument: cost-shifters

- But, cost data are rarely observed in practice.

- Hausman (1996) and Nevo (2001) use indirect cost measures: **prices in other markets**
  - i.e. $p_{jt'}$ for $t' \neq t$
  - Validity condition: conditional on x_t and x_t', pricing is independent across markets and ξ_{jt} and $\xi_{jt'}$ are independent.
  - In words: "IVs exploit common cost shocks across markets"

- Problems (example):
  - Unobserved advertising campaigns

Instruments: Waldfogel-Fan instruments

- Used in Waldfogel (2003), Fan (2013)
- Use **demographics in other counties where the product is sold**
  - Fan (2013): newspapers sold in multiple counties, uses demographics in other counties as IVs.
  - Idea: rely on consumption/preference externalities
  - E.g. Product offered in multiple counties $\rightarrow$ characteristics of product impacted by the attributes (like demographics) of the other counties.
  - Validity: conditional on variables in model, ξ_{jt} not correlated across counties (same assumption in Hausman instruments)
  - Additional concern: set of counties where product is offered is not exogenous

Plan

1. Estimation algorithm
2. Instrumental variables
3. **Extensions**
4. Applications

Extensions/other useful data sources

- Second choice data
- (e.g. from a survey)
- e.g. see Berry, Levinsohn and Pakes: "Differentiated products demand systems from a combination of micro and macro data: the new vehicle market" (2004)
- Micro-moments

Plan

1. Estimation algorithm
2. Instrumental variables
3. Extensions
4. **Applications**

Application: distinguishing between models of competition

- Nevo "Measuring Market Power in the Ready-to-eat Cereal Industry" (Econometrica, 2001)
- RTE Cereal
  - Concentrated market
  - High margins
  - High advertising-to-sales ratios
  - Aggressive introduction of new products
- "Used as a classic example of a concentrated differentiated-products industry in which price competition is approximately cooperative and rivalry is channeled into advertising and new product innovation"

Application: distinguishing between models of competition

- **Research questions**
- 1. Is there collusive pricing?
- 2. Decompose price-cost margins (PCM) into:
  - i. Product differentiation
  - ii. Portfolio effects (firms offer multiple products)
  - iii. Price collusion
- **Main takeaways for this class:**
  - See demand estimation 'in action'
  - Inclusion of a supply side that allows for horizontal competition
  - (Application/question is of general interest - but very "traditional IO")

Application: distinguishing between models of competition

- **Overall strategy:**
- 1. Estimate demand
- 2. Use estimates + pricing rules implied by different models of firm conduct to get price-cost margins (PCM)
    - Challenge: costs not observed
- 3. Compare predicted PCM from different models of conduct to true PCM
    - See which model of firm conduct best matches the data.
- **Main finding:**
- Nash-Bertrand pricing best matches observed PCM