Discrete Choice Models With Multiple Unobserved Choice Characteristics

Susan Athey - Stanford University

&

Guido Imbens - UC Berkeley

"Economics to Econometrics"
A Conference in Honor of Daniel McFadden
Berkeley, May 6th, 2005

Outline

- Motivation
- Model: Unobserved Product Characteristics, Unobserved Individual Heterogeneity
- Literature
- Identification: Why Do We Need Multiple Unobserved Product Characteristics
- Implementation: Gibbs Sampler
- Application: Scanner Data for Yogurt Purchases

Motivation

- Large literature on discrete response models, going back to McFadden (1974, 1981).

- What is (non-parametrically) identified in those models?

- Results on binary response models with random coefficients (Ichimura & Thompson, 1998)

- Looking for Matzkin (2003) type identification results:
  What models are just-identified?

Model I: Single Market

Utility for choice j for individual i is U_{ij}:

$$U_{ij} = g(X_j, \xi_j, Z_i, \nu_i) + \epsilon_{ij}.$$

$\epsilon_{ij} \sim \mathcal{N}(0, 1)$, independent of everything.

$$g(X_j, \xi_j, Z_i, \nu_i) = X_j' \beta_i + \xi_j' \gamma_i,$$

where

$$\theta_i = \begin{pmatrix} \beta_i \\ \gamma_i \end{pmatrix} = \begin{pmatrix} \Delta_o \\ \Delta_u \end{pmatrix} Z_i + \begin{pmatrix} \nu_{io} \\ \nu_{iu} \end{pmatrix}.$$

$$\nu_i = \begin{pmatrix} \nu_{io} \\ \nu_{iu} \end{pmatrix} | \mathbf{X}, \mathbf{Z} \sim \mathcal{N}(0, \Omega),$$

so that

$$U_{ij} = X_j' \Delta_o Z_i + X_j' \nu_{io} + \xi_j \Delta_u Z_i + \xi_j' \nu_{iu} + \epsilon_{ij}.$$

Dimension of ξ possibly greater than one.

X_j are exogenous.

Market share is $s_j(z) = \text{pr}(Y_i = j | Z_i = z)$ (function of individual characteristics, e.g., income, demographics)

Literature

- Berry-Levinsohn-Pakes (1995, 1998):

  $$U_{ij} = X_j' \Delta_o Z_i + X_j' \nu_{io} + \xi_j + \epsilon_{ij}.$$

  random coeff., single unobs. prod. char.,

- Elrod & Keane (1995)

  $$U_{ij} = \xi_j' \Delta_u + \xi_j' \nu_{iu} + \epsilon_{ij}.$$

  multiple unobserved prod char, panel data.

- Goettler & Shachar (2001):

  $$U_{ij} = \xi_{1j} + \left(\xi_{2j} - \nu_i\right)' A_i \left(\xi_{2j} - \nu_i\right) + \epsilon_{ij}.$$

  multiple unobserved prod char, panel data.

- Hedonic Models: Berry & Pakes (this morning), Song (2004)

  $$U_{ij} = X_j' \Delta_o Z_i + X_j' \nu_{io} + \xi_j.$$

  No ϵ_{ij}

Identification

Simple setting: single market, J products, large number of individuals.

No unobserved product characteristics ξ_j,
no unobserved individual characteristics ν_i.

I. If there are no ϵ_{ij}, then all individuals with the same observed characteristics $Z_i = z$ make the same choice: $s_j(z) \in \{0, 1\}$.

II. With ϵ_{ij} extreme value we can rationalize any set of market shares:

$$s_j(z) = \frac{\exp(g(X_j, z)}{\sum_{k=1}^{J} \exp(g(X_k, z)},$$

by setting $g(X_j, z) = \log(s_j(z)/s_1(z))$ and $g(X_1, z) = 0$.
(no need for unobserved product charactistics or unobserved individual characteristics in this case)

Identification (ctd)

III. Suppose there are enough products so that we have multiple products with similar (identical) observed product characteristic, and $s_j(z)$ satisfying $0 < \underline{c} < J \cdot s_j(z) < \bar{c} < 1$ for all j and z.

Consider two products j and k with the same observed product characteristics $X_j = X_k = x$.

Then the absence of unobserved product characteristics implies equal market shares $s_j(z) = s_k(z)$.

If the market shares differ there must be unobserved product characteristics.

Identification (ctd)

IV. Suppose there is a single unobserved product characteristic ξ_j.

Suppose also that utility $g(x, \xi, z)$ is monotone in ξ. (like BLP, Matzkin (2003), Petrin (2002), Elrod & Keane 91995), but unlike Goettler & Shachar (2001))

Compare two products j and k with $X_j = X_k = x$.

If $\xi_j > \xi_k$, then $s_j(z) \geq s_k(z)$ for all z.

Hence crossing of market share functions ($s_j(z_0) > s_k(z_0)$ and $s_j(z_1) < s_k(z_1)$) is incompatible with these assumptions. Presence of unobserved individual characteristics does not change this.

Identification (ctd)

Possibility A: utility is very non-monotone in scalar unobserved characteristic. If utility is allowed to be very non-monotone (many sign changes for derivative), a single unobserved choice characteristic is sufficient to rationalize data.

Possibility B: utility is inverse U-shaped in scalar unobserved characteristic. Combined with unobserved individual heterogeneity (e.g., unobserved heterogeneity in ideal point) this can rationalize data. With many observations per individual this possibility can be ruled out: the individual heterogeneity can be controlled for.

Possibility C: there are two unobserved product characteristics. With $g(x, \xi_1, \xi_2, z)$ monotone in both unobserved characteristics we can rationalize any countable set of market share functions $s_j(z)$.

Illustration of Identification Result

- Estimate Market Share Functions $s_j(z)$ for 8 Yogurt Brands as Function of Income (no observed product characteristics)

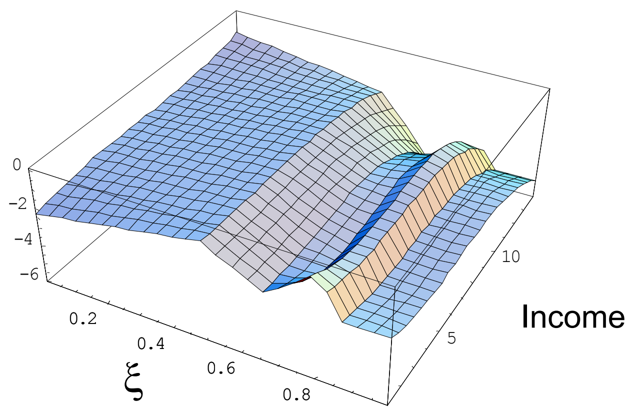
- Convert to utilities $g_j(z) = \log(s_j(z)/s_1(z))$, relative to Dannon brand.

- Find $g(\xi_j, z)$ s.t. $g_j(z) = g(\xi_j, z) \ \forall \ j, z$.
  - See Figure 1.
  - Note: $g(\xi, z)$ is non-monotone in ξ.

- Find $g(\xi_{j1}, \xi_{j2}, z)$ such that $g_j(z) = g(\xi_{j1}, \xi_{j2}, z)$ for all j and z.
  - See Fig 2 for $z = 1$, Fig 3 for $z = 13$.
  - Note: $g(\xi_1, \xi_2, z)$ is monotone in ξ_1, ξ_2.

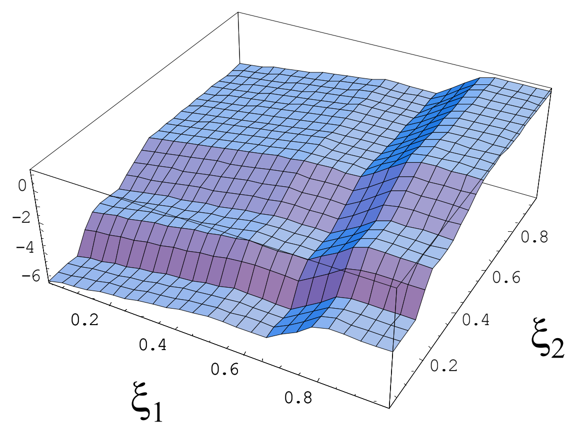
Case Study: Utility as a Function of Income for 8 Brands of Yoghurt

Utility (Relative to Dannon) as a Function of Income and a Single Unobserved Characteristic

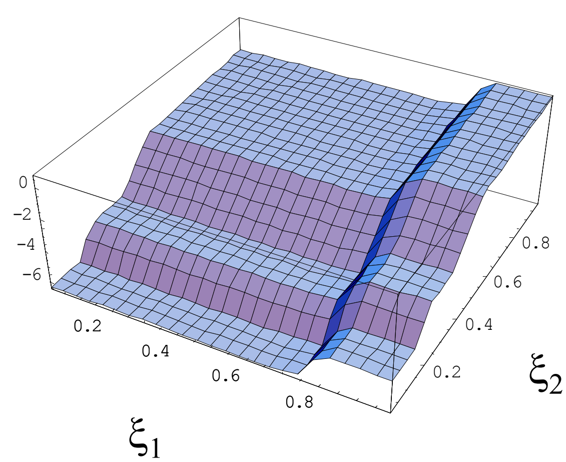


Utility (Relative to Dannon) as a Function of Two Unobserved Characteristics, Evaluated at Two Particular Levels of Income

Income Category 1



Income Category 13



Model II: Multiple Markets

Utility for choice j for individual i in market m is U_{ijm}:

$$U_{ijm} = g(X_{jm}, \xi_j, Z_i, \nu_i) + \epsilon_{ijm}.$$

$\epsilon_{ijm} \sim \mathcal{N}(0, 1)$, independent of everything.

Observed product characteristics now potentially vary by market: X_{jm}

Unobserved product characteristics do not vary by market: ξ_j

Product characteristics are still exogenous.

Market share is $s_{jm}(z) = \text{pr}(Y_{im} = j | Z_i = z)$.

Identification (ctd)

V. Consider markets m and n. Compare market share functions for product j, k and l in the two markets, and suppose that $X_{jm} \neq X_{jn}$, but $X_{km} = X_{kn}$ and $X_{lm} = X_{ln}$. The model with logistic ϵ_{ijm} implies that

$$\frac{s_{km}(z)}{s_{km}(z) + s_{jm}(z)} = \frac{s_{kn}(z)}{s_{kn}(z) + s_{jn}(z)}$$

by IIA property. Suppose this does not hold.

Possibility A: unobserved product characteristics ξ_{jm} that differ by market. (this can explain anything)

Possibility B: unobserved individual heterogeneity.

(Presence of unobserved individual heterogeneity can also be established with panel data.)

Identification: Summary

In order to rationalize choice data with a model that has utility a function of observed and unobserved product characteristics and observed and unobserved individual characteristics plus a idiosyncratic error we need:

with many products: multiple unobserved product characteristics.

with variation in observed product characteristics by market: unobserved heterogeneity in taste parameters.

Implementation: the Gibbs Sampler

For implementation we use a Bayesian approach with MCMC methods.

We divide the unobserved random variables into 5 groups.

1. The latent utilities U_{ik} for all individuals and all choices.

2. The individual taste parameters $\theta_i = (\beta_i, \gamma_i)$.

3. The common taste parameter Δ.

4. The unobserved choice characteristics ξ_k.

5. The covariance matrix of the individual taste parameters Ω.

Implementation: the Gibbs Sampler

Step I: Latent Utilities U_{ik}: draw sequentially from truncated normal distributions

Step II: Individual Coefficients β_i and γ_i: normal distribution

Step III: Common Regression Coefficients Δ: normal distribution

Step IV: Latent Choice Characteristics ξ_k: normal distribution

Step V: Covariance Matrix of Individual Taste Parameters Ω: wishart distribution

Application: The Demand for Yogurt

Based on scanner data. 8 brands, panel of households, 16824 observations

Four Models:

I. No Unobserved Product Characteristics, No Unobserved Individual Heterogeneity

II. No Unobserved Product Characteristics, Unobserved Individual Heterogeneity

III. Single Unobserved Product Characteristic, Unobserved Individual Heterogeneity

IV. Two Unobserved Product Characteristics, Unobserved Individual Heterogeneity

Own and Cross Price Elasticities for two leading brands, A and B (with market shares of 48% and 16%)

Model	# of Unobs prod char	Brand A own	Brand A cross	Brand B own	Brand B cross
I	0	-0.63	0.60	-0.61	0.12
II	0	1.42	-0.09	-0.74	-0.12
III	1	-0.54	0.58	-1.92	0.20
IV	2	-0.69	0.66	-2.25	0.21