Coping with Insufficient Data: The case of Household Automobile Holding Modelling

Ryuichi Kitamura (Kyoto Univ.) and
Toshiyuki Yamamoto (Nagoya Univ.)

Motivation: Insufficient data

- It is often the case that typically available data do not contain all the variables that are desired for the analysis of the behavior of interest.

- In case of the analysis of household automobile holding behavior, information on the cost of holding an automobile is rarely available in data
  - make and model of the automobile, acquired new or used, purchase price, fuel consumption rate, or insurance costs—is typically unavailable.

Objective: A possible approach

A theoretical model

- based on external principles
- embodies relationships among observed variables.

In this study,

A utilitarian model of household automobile holding is developed

- based on the assumption that a household holds an optimum number of automobiles at the time of observation.

Key: Base auto ownership cost

- Purchase price is not available

But

- minimum fixed cost per unit time to hold an automobile can be assumed
- Each household is assumed to spend a nonnegative amount of money in addition to the base cost to hold a better automobile that offers more amenities

Utility model

Assumption

- Household optimizes its vehicle holding and use

Model

- expresses the utility of automobile holding in terms of income and household size
- without requiring variables that can hardly be measured, e.g., unit cost of auto and transit travel

$$U = U(M_A, M_T, A, X \mid n_A) = \left(M_A \left(\frac{n_A}{n_H} \right)^{\eta} \right)^{\alpha} M_T^{\beta} A^{\gamma} X^{\delta}$$

Modifier that represents effect of auto availability

n_A = number of automobiles

n_H = number of adult household members

M_A = mobility by automobile per adult household member (person-km)

M_T = mobility by public transit per adult household member (person-km)

A = auto amenities expenditure per automobile

X = expenditure per adult household member for other goods

Subject to

- Income constraint
- Minimum mobility requirements be met
- Ceiling on the use of household automobiles in terms of total vehicle kilometers.

$$n_H p_A M_A + n_H p_T M_T + n_A(qA + \tilde{C}) + n_H pX = Y$$

$$n_H (M_A + M_T) \geq \underline{M}(Z)$$

$$\frac{n_H M_A}{n_A} \leq \overline{M}(Z)$$

Income constraint

$$n_H p_A M_A + n_H p_T M_T + n_A (qA + \tilde{C}) + n_H pX = Y$$

p_A = auto variable cost (per person-km)

p_T = transit variable cost (per person-km)

q = unit auto amenity cost

$\tilde{C}$ = base auto cost per unit time

p = price of other goods

Minimum mobility requirements

$$n_H (M_A + M_T) \geq \underline{M}(Z)$$

Household has a larger mobility than the minimum mobility $\underline{M}(Z)$

Ceiling on use of automobiles

$$\frac{n_H M_A}{n_A} \leq \bar{M}(Z)$$

Mobility by automobile can be served by household fleet $\bar{M}(Z)$

Unbounded solution

- Unbounded solution ignore 2nd and 3th constraints
- By introducing random error terms
- We get indirect utility below

$$U_i^*(n_A) = \begin{cases} (\beta+1)\ln Y + \alpha(1+\eta)\ln n_H + C + \varepsilon_i(0) & \text{if } n_A = 0 \\ (\alpha+\beta+\gamma+1)\ln(Y - n_A\tilde{C}) - (\gamma-\alpha\eta)\ln n_A + \varepsilon_i(n_A) & \text{if } n_A = 1,2,\ldots \end{cases}$$

- which doesn't require variable costs, nor purchase price

Future tasks

- Estimation of discrete choice model
  - Base automobile cost is unknown. It is proposed that alternative values be postulated for when estimating the model
- Boundary solutions must be incorporated into the estimation process