Modeling the Choice of Residential Location

Daniel McFadden, Department of Economics, Massachusetts Institute of Technology, Cambridge

The problem of translating the theory of economic choice behavior into concrete models suitable for analyzing housing location is discussed. The analysis is based on the premise that the classical, economically rational consumer will choose a residential location by weighing the attributes of each available alternative and by selecting the alternative that maximizes utility. The assumption of independence in the commonly used multinomial logit model of choice is relaxed to permit a structure of perceived similarities among alternatives. In this analysis, choice is described by a multinomial logit model for aggregates of similar alternatives. Also discussed are methods for controlling the size of data collection and estimation tasks by sampling alternatives from the full set of alternatives.

The classical, economically rational consumer will choose a residential location by weighing the attributes of each available alternative—accessibility to work place, shopping, and schools; quality of neighborhood life and availability of public services; costs, including price, taxes, and travel costs; and dwelling characteristics, such as age, number of rooms, type of appliances—and by choosing the alternative that maximizes

section on limiting the number of alternatives establishes that, if choice among a set of alternatives is described by a multinomial logit model, then the model can be estimated by sampling from the full set of alternatives, with appropriate adjustment in the estimation mechanism. Thus, estimation can be carried out with limited data collection and computation.

The solutions I give to the two problems above will be applied to empirical studies of housing location by Quigley (1) and Lerman (2).

THEORY OF HOUSING LOCATION CHOICE

Assume the classical model of the rational, utility-maximizing consumer. Suppose the consumer faces a residential location decision, with a choice of communities indexed $c = 1, \ldots, C$ and dwellings indexed $n = 1, \ldots, N_c$ in community c. The consumer will have a utility U_{cn} for alternative cn, which is a function of the

GENERALIZED EXTREME VALUE MODEL

I shall now introduce a family of choice models, derived from stochastic utility maximization, that includes multinomial and nested logit. This family allows a general pattern of dependence among the unobserved attributes of alternatives and yields an analytically tractable closed form for the choice probabilities. The following result characterizes the family.

Suppose $G(y_1, \ldots, y_J)$ is a nonnegative, homogeneous-of-degree-one function of $(y_1, \ldots, y_J) \geq 0$. Suppose $G \rightarrow \infty$ if $y_i \rightarrow \infty$ for each i, and for k distinct components $i_1, \ldots, i_k$, $\partial^k G/\partial y_{i_1} \ldots y_{i_k}$ is nonnegative if k is odd and nonpositive if k is even. Then

$$P_i = \{\exp(V_i)\, G_i[\exp(V_1), \ldots, \exp(V_J)]\}/G[\exp(V_1), \ldots, \exp(V_J)] \quad (12)$$

defines a probabilistic choice model from alternatives $i = 1, \ldots, J$, which is consistent with utility maximiza-

Let $Y_j = \exp(u_j)$. Consider function $G(Y_1, \ldots, Y_J) : \mathbb{R}^J \to \mathbb{R}$

Let $Y_j = \exp(u_j)$. Consider function $G(Y_1, \ldots, Y_J) : \mathbb{R}^J \to \mathbb{R}$

If G satisfies:

- $G \geq 0$
- Homogeneous of degree k
- $G \to \infty$ as $Y_j \to \infty$ for any j
- Cross partial derivatives weakly alternate in sign, beginning with $G_i \geq 0$

Then:

$$F(u_1, \ldots, u_J) = \exp[-G(Y_1, \ldots, Y_J)]$$

Then:

$$F(u_1, \ldots, u_J) = \exp[-G(Y_1, \ldots, Y_J)]$$

is the CDF of a multivariate extreme value distribution

Then:

$$F(u_1, \ldots, u_J) = \exp[-G(Y_1, \ldots, Y_J)]$$

is the CDF of a multivariate extreme value distribution

And the choice probabilities are:

$$P_i = \frac{Y_i G_i}{G}$$

where G_i is the derivative of G with respect to Y_i

Alternative formulation:

$$P_i = \frac{\partial \log(G)}{\partial u_i}$$

Alternative formulation:

$$P_i = \frac{\partial \log(G)}{\partial u_i}$$

Key insight: $\log(G)$ (plus Euler's constant) = expected utility **for any GEV model**

Multinomial Logit Case

$$G = \sum_{j=1}^{J} \exp(u_j)$$

Multinomial Logit Case

$$G = \sum_{j=1}^{J} \exp(u_j)$$

Taking derivative of $\log(G)$:

Multinomial Logit Case

$$G = \sum_{j=1}^{J} \exp(u_j)$$

Taking derivative of $\log(G)$:

$$\begin{aligned} P_i &= \frac{\partial \log(G)}{\partial u_i} \\ &= \frac{\exp(u_i)}{\sum_{j=1}^{J} \exp(u_j)} \end{aligned}$$

Multinomial Logit Case

$$G = \sum_{j=1}^{J} \exp(u_j)$$

Taking derivative of $\log(G)$:

$$\begin{aligned} P_i &= \frac{\partial \log(G)}{\partial u_i} \\ &= \frac{\exp(u_i)}{\sum_{j=1}^{J} \exp(u_j)} \end{aligned}$$

This gives multinomial logit probabilities

Nested Logit Case

Two nests (F, NF) and no-purchase option N:

$$G = \left(\sum_{j \in F} \exp\left(\frac{u_j}{\lambda_F} \right) \right)^{\lambda_F} + \left(\sum_{j \in NF} \exp\left(\frac{u_j}{\lambda_{NF}} \right) \right)^{\lambda_{NF}} + \exp(u_N)$$

Nested Logit Case

Two nests (F, NF) and no-purchase option N:

$$G = \left(\sum_{j \in F} \exp\left(\frac{u_j}{\lambda_F} \right) \right)^{\lambda_F} + \left(\sum_{j \in NF} \exp\left(\frac{u_j}{\lambda_{NF}} \right) \right)^{\lambda_{NF}} + \exp(u_N)$$

For $k \in F$:

Nested Logit Case

Two nests (F, NF) and no-purchase option N:

$$G = \left(\sum_{j \in F} \exp\left(\frac{u_j}{\lambda_F}\right)\right)^{\lambda_F} + \left(\sum_{j \in NF} \exp\left(\frac{u_j}{\lambda_{NF}}\right)\right)^{\lambda_{NF}} + \exp(u_N)$$

For $k \in F$:

$$P_k = \frac{\exp\left(\frac{u_k}{\lambda_F}\right)\left(\sum_{j \in F} \exp\left(\frac{u_j}{\lambda_F}\right)\right)^{\lambda_F - 1}}{\left(\sum_{j \in F} \exp\left(\frac{u_j}{\lambda_F}\right)\right)^{\lambda_F} + \left(\sum_{j \in NF} \exp\left(\frac{u_j}{\lambda_{NF}}\right)\right)^{\lambda_{NF}} + \exp(u_N)}$$