ELSEVIER

Journal of Business Research 51 (2001) 115–125

JOURNAL OF BUSINESS RESEARCH

Random utility models in marketing research: a survey

George Baltas[a,*], Peter Doyle[b]

[a]Athens University of Economics and Business, Athens, Greece
[b]Warwick Business School, University of Warwick, Coventry CV4 7AL, England, UK

Accepted 1 February 2000

Abstract

Random utility (RU) models are well-established methods for describing discrete choice behavior. Recently, there has been a strong upsurge in interest driven by advances in data gathering and estimation technology. This review paper describes the principles and issues, and develops a taxonomy of three major families of models. The paper summarizes and classifies the different approaches. The advantages and limitations of the various alternatives are outlined. Practical issues in implementing the models are also discussed.

Keywords: Random utility models; Marketing research; Discrete choice behavior

1. Introduction

With an ever increasing importance of market intelligence, the need to understand advanced methods of market research has never been greater. Random utility (RU) models have been developed to describe choice among mutually exclusive discrete alternatives and received considerable academic and industry attention. This paper surveys RU models, discusses key issues and develops a structurally meaningful synthesis of the different formulations. It is intended to help managers and researchers keep informed of a fast-changing and important area which may not directly fall within their own specific professional or research interests. The paper is organized as follows. The subsequent section illustrates the underlying principles of RU models and explains how probabilistic choice flows from utility maximization. The third section discusses the behavior of RU models and issues that arise from restrictive stochastic assumptions. The fourth section introduces a taxonomy and describes the different models. The fifth section is concerned with implementation and experimental data. The sixth section discusses some practical issues and the last section concludes.

* Corresponding author.
E-mail address: gb@aueb.gr (G. Baltas).

2. Overview

Consider an individual agent choosing a single option among a finite set of alternatives, for example, a consumer deciding which brand to buy. This is the realm of behavior that is considered in RU modeling. In RU models, preferences for such discrete alternatives are determined by the realization of latent indices of attractiveness, called product utilities. Utility maximization is the objective of the decision process and leads to observed choice in the sense that the consumer chooses the alternative for which utility is maximal. Individual preferences depend on characteristics of the alternatives and the tastes of the consumer. An RU model defines a mapping from observed characteristics into preferences. The analyst however cannot observe all the factors affecting preferences and the latter are treated as random variables. By its abstraction from various idiosyncratic factors, the model uses stochastic assumptions to describe unmeasured variation in preferences. An operational way to allow for maximization of latent preferences is to consider a utility function that is decomposable into two additively separable parts, (1) a deterministic component specified as a function of measured attributes of the alternatives and/or the individual, and (2) a stochastic component representing unobserved attributes affecting choice, interindividual differences in utilities depending upon the heterogeneity in tastes, measurement errors, and functional misspecification (Manski, 1977). In the next sections, we shall consider

0148-2963/00/$ – see front matter
PII: S0148-2963(99)00058-2

different models based on alternative hypotheses about the "unknown." Proceeding further in the same vein, let

$$U_{ij} = V_{ij} + \varepsilon_{ij} \tag{1}$$

be the utility of alternative j for consumer i, where V_{ij} is the deterministic component and ε_{ij} the random component. Typically, the deterministic component V_{ij} has been assumed to have an additively separable linear form $V_{ij} = \boldsymbol{x}_{ij}{}^T\boldsymbol{\beta}$, where $\boldsymbol{x}_{ij}$ and $\boldsymbol{\beta}$ are the vectors of exogenous variables and parameters, respectively. In the hypothetical case that V contains perfect information about the determinants of utility, the consumer would simply choose the product with the highest V_{ij}. The stochastic terms ε_{ij} shaping the true and latent utility in Eq. (1), introduce uncertainty regarding the choice and therefore, choice probabilities are invoked to describe choice behavior. The probabilistic description of choice has been introduced not to reflect behavior that is probabilistic. Rather, it is the lack of information that leads the analyst to treat utility as a random variable and consequently to describe choice in a probabilistic fashion. In fact, the properties of RU models can be attributed to the specific assumptions that each model implies about the stochastic terms. Under the utility maximization rule, a consumer facing a set of available products $C = \{1,2,3, \ldots, M\}$ will choose a product j with probability $P(j) = P(U_j > U_k)$ for all $k \in C,\ k \neq j$, or as it follows from Eq. (1):

$$P(j) = P(\varepsilon_k < V_j - V_k + \varepsilon_j) \text{ for all } k \in C,\ k \neq j. \tag{2}$$

The probability that j is chosen is then obtained by making assumptions about the form of the distribution of the random variables and integrating Eq. (2) over a continuum of all possible values for ε_j. From Eq. (2), we can write the selection probability for, say, the first alternative as

$$P(1) = \int_{-\infty}^{\infty} \int_{-\infty}^{V_1-V_2+\varepsilon_1} \int_{-\infty}^{V_1-V_3+\varepsilon_1} \cdots \times \int_{-\infty}^{V_1-V_M+\varepsilon_1} f(\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots \varepsilon_M) d\varepsilon_M \ldots d\varepsilon_3 d\varepsilon_2 d\varepsilon_1. \tag{3}$$

In words, Eq. (3) states that the choice probability of alternative 1 is the probability of over all possible values of ε_1, all the other random terms being less than $V_1 - V_j + \varepsilon_1$, $\forall j \in C$. The popular multinomial logit (MNL) model is derived by assuming that the random terms are independently identically distributed (IID) according to the double exponential distribution with mode zero and variance $\mu^2\pi^2/6$, where μ is a positive scale parameter. The choice probability in Eq. (3) then takes the compact form,

$$P(j) = \frac{\exp\left(\frac{V_j}{\mu}\right)}{\sum_{k \in C} \exp\left(\frac{V_k}{\mu}\right)}. \tag{4}$$

The analytic form of the MNL probabilities has greatly contributed to the popularity of the MNL model. The expression in Eq. (4) can be derived in a great number of ways (McFadden, 1973; Train, 1986; Anderson et al., 1992). Having laid out the necessary background, we turn to the stochastic assumptions of the models.

3. Stochastic assumptions of RU models

Suppose we observe members of a population of consumers, each member i of which has a utility function $U_{ij} = V_{ij} + \varepsilon_{ij}$ for each product j of a set $C = [1,2,3, \ldots, M]$. V_{ij} is the non-stochastic function mapping attributes into utility and ε_{ij} accounts for factors not included in V_{ij}.

The simple MNL model accounts for unobserved determinants of choice by IID random terms. That is they are assumed to have the same distribution, with the same mean and variance and also to be uncorrelated across and within individuals. An interesting property is the effect of increasing unexplained stochastic variation on the identified coefficients. Since the variance (assumed the same for all ε_j) is related with the parameter μ, it is obvious in Eq. (4) that the variance discounts the value of the estimated parameters in the non-stochastic function V. Since the variables in V are exogenous, the estimated coefficients absorb the variance effect. Intuitively, high variance implies limited ability of the observed variables to explain choices and therefore leads to smaller values of the coefficients. Since μ is a transformation of the variation of the random disturbances, it can be seen as an index of unobserved variation in preferences that cannot be explained by the variables in the non-stochastic function V. In the MNL, the price coefficient reflects the response of choice probabilities to prices and its magnitude is related to the variance parameter μ. As unobserved variation decreases, the value of the identified price parameter in V increases and vice versa. Therefore, the identified price coefficient can be regarded as an index of average substitutability among alternatives that is related to stochastic variation. Although the MNL accommodates varying rates of symmetric substitution, the assumption of IID random components remains restrictive and imposes the independence of irrelevant alternatives (IIA) property (Ben-Akiva and Lerman, 1985). Under this structural restriction, the odds of the consumer choosing j over k remain the same regardless of the composition of the choice set. An analogous and possibly more important drawback is that the model cannot postulate any pattern of differential substitutability between products. An improvement in an alternative's systematic utility will have a proportionally equal impact on the selection probabilities of all other alternatives. Thus, an implication of the IIA property is that the cross-elasticity of the probability of brand j with respect to a change in V_k is the same for all j with $j \neq k$.

The assumption of independent preferences is restrictive. In reality, alternatives may not be equally dissimilar. Differential similarities among products due to shared characteristics lead to correlated utilities. When these conditions