

# Chapter 7

## Economic Models of Choice

Greg M. Allenby, Jaehwan Kim and Peter E. Rossi

### 7.1 Introduction

Models of choice are fundamental to the field of marketing because they represent the culmination of marketing efforts embedded in the 4P's (product, price, promotion and place). It is by understanding how people make purchase decisions that we can inform firms on the success of their efforts in each of the functional areas of marketing. Choice models quantify the process of exchange, allowing us to understand the origins of preference and the determinants of costs in a transaction, including the time, money and other resources needed to acquire and use an offering.

Choice is complex. It involves our resources, perceptions, memory and other factors as we acquire and use products to improve our lives. Our goal in this chapter is to provide a review of direct utility choice models that attempt to rationalize choice. We do this within an economic framework of demand where people are assumed to be constrained utility maximizers. We take this view because marketplace data supports the concept of constrained maximization as evidenced by the large proportion of zero's in disaggregate marketing data, coupled with the observation that people are sensitive to price and demands on their time. That is, marketing data is overwhelmingly characterized by sparse demand where most people don't purchase most of the

---

G.M. Allenby (✉)

Fisher College of Business, Ohio State University, Columbus, OH, USA  
e-mail: allenby.1@osu.edu

J. Kim

Korea University Business School, Seoul, Korea  
e-mail: jbayes@korea.ac.kr

P.E. Rossi

Anderson School of Management, UCLA, Los Angeles, CA, USA  
e-mail: perossichi@gmail.com

© Springer International Publishing AG 2017

B. Wierenga and R. van der Lans (eds.), *Handbook of Marketing Decision Models*, International Series in Operations Research & Management Science 254, DOI 10.1007/978-3-319-56941-3\_7

products available for sale, don't frequent most websites available to them, and don't read most of the literature published on topics of interest. Instead, they select what they consume in a manner that suggests people are resource conserving.

We acknowledge that our treatment of choice models is selective and related to our own research agenda. We believe it is not possible to provide a comprehensive survey of choice models in marketing, as evidenced by the presence of dedicated conferences and journal volumes to the issue of choice. Choice encompasses a vast domain of economic, psychologic and social subject matter, and our chapter provides a narrow emphasis in an area of choice which we hope to popularize and expand.

An advantage of rationalizing decisions within an economic framework is that it can lead to interventions and policy recommendations that improve the profitability of the firm. Measuring the impact of product quality on demand often requires models capable of dealing with more than simple discrete choices, where only one unit of one product is chosen. Multi-part pricing, time, space and other constraints likewise impact the attainable utility consumers can achieve. Firms considering changes to their product line require measures of consumer satisfaction and compensating values based on flexible patterns of substitution that are not pre-ordained by properties such as IIA (Allenby 1989). A potential disadvantage is that economic models of choice and demand can be too simplistic, as often pointed out by behavioral decision theorists. Our view is that much of the criticism of standard economic models of choice results from simplistic model assumptions rather than a defect in the fundamental paradigm of constrained choice. Our view is that the solution is to develop richer specifications of utility and constraints than to reject an economic formulation.

This chapter in the *Handbook of Marketing Decision Models* starts with a simple discrete choice model that has become a workhorse model in marketing over the last 25 years. The simple discrete choice model allows us to introduce terminology and basics of choice modeling. We then expand our discussion by considering direct utility models that allow the study of utility formation separate from the role of constraints, which provide insights into what people give up in an exchange. The direct utility formulation also allows us to model choices in the context of continuous demand where more than one item may be selected. Throughout our analysis, we offer a critical assessment of model assumptions and point to future directions for additional research.

## 7.2 A Simple Model of Discrete Choice

The discrete choice model is characterized by one and only one choice alternative being selected at a time. This model of choice is applicable to a wide variety of product categories, ranging from automobiles to smartphones to cordless power drills. Consumers are assumed to be endowed with a budget for the purchase that determines the upper limit of expenditure they are willing to make. If the good costs less than the budgeted amount, which we denote " $E$ " for expenditure, then the remainder of the unspent money ( $E - p$ ) can be used for other purposes. Thus, we conceive

of choice as between a number of different “inside” goods and an “outside” good that represents money unspent in the product category. The utility function for this situation can be represented as:

$$u(x, z) = \sum_{k=1} \psi_k x_k + \psi_z z$$

where  $x$  is a vector of demand for the inside goods,  $z$  is the demand for the outside good or non-purchase,  $\psi_j$  is the marginal utility of choosing the  $j^{th}$  good and  $\psi_z$  is the marginal utility for money. It is customary in choice models to include an error term to allow for unobserved factors affecting choice. We can then consider the utility from selecting each of the alternatives as:

$$\begin{aligned} u(x_1 = 1, z = E - p_1) &= \psi_1 + \psi_z (E - p_1) + \varepsilon_1 \\ u(x_2 = 1, z = E - p_2) &= \psi_2 + \psi_z (E - p_2) + \varepsilon_2 \\ &\vdots \\ u(x = 0, z = E) &= \psi_z (E) + \varepsilon_z \end{aligned}$$

The utility for each of the inside goods is comprised of three terms (i) the marginal utility for the inside good; (ii) utility for the unspent money that can be put to other use; and (iii) an error term. The utility for the outside good is just the utility for the budgeted allotment,  $E$ , plus error.

Consumers are assumed to select the choice alternative that provides them greatest utility. The choice model becomes:

$$\begin{aligned} \Pr(j) &= \Pr(\psi_j + \psi_z (E - p_j) + \varepsilon_j > \psi_k + \psi_z (E - p_k) + \varepsilon_k \text{ for any } k \neq j) \\ &= \Pr(V_j + \varepsilon_j > V_k + \varepsilon_k \text{ for any } k \neq j) \\ &= \Pr(\varepsilon_k < V_j - V_k + \varepsilon_j \text{ for any } k \neq j) \\ &= \int_{-\infty}^{+\infty} \left[ \int_{-\infty}^{V_j - V_1 + \varepsilon_j} \cdots \int_{-\infty}^{V_j - V_k + \varepsilon_j} \phi(\varepsilon_k) \cdots \phi(\varepsilon_1) \right] \phi(\varepsilon_j) d\varepsilon_k \cdots d\varepsilon_1 d\varepsilon_j \\ &= \int_{-\infty}^{+\infty} \prod_{k \neq j} \Phi(V_j - V_k + \varepsilon_j) \phi(\varepsilon_j) d\varepsilon_j \end{aligned}$$

where  $\Phi$  denotes the cdf and  $\phi$  denotes the pdf of the distribution of  $\varepsilon$ . Distributional assumptions play a role in determining the functional form of the choice model, with extreme value errors leading to the logit model and normally distributed errors giving rise to the probit model. Assuming standard extreme value errors (i.e.,  $EV(0, 1)$ ) results in the following logit expression for the choice probability:

$$\begin{aligned}
Pr(j) &= \frac{\exp [V_j]}{\exp [V_z] + \sum_k \exp [V_k]} \\
&= \frac{\exp [\psi_j + \psi_z (E - p_j)]}{\exp [\psi_z (E)] + \sum_k \exp [\psi_k + \psi_z (E - p_k)]} \\
&= \frac{\exp [\psi_j - \psi_z p_j]}{1 + \sum_k \exp [\psi_k - \psi_z p_k]} \tag{7.1}
\end{aligned}$$

Thus, the choice probability is a function of a choice-specific intercept ( $\psi_k$ ) and a price term with a coefficient that is common across the choice alternatives.

This simple model of choice is used extensively in marketing because of its computational simplicity. Demand is restricted to two points for each of the inside goods,  $\{0, 1\}$ , with only one good allowed to be chosen and the remaining budget allocated to the outside good,  $z$ . Moreover, because utility is specified as linear, the budgetary allotment,  $E$ , cancels out of the expression and does not figure into the choice probability specification in a meaningful way, other than to remove choice options from the denominator of the probability expression when the price is too high, i.e.,  $p_k > E$ .

The discrete choice model requires an additional constraint to be statistically identified, or estimable, from a choice dataset. It is traditionally assumed that the scales of the error terms ( $\epsilon$ ) are set to one  $\sigma = 1$ . The likelihood function for the standard discrete choice model is therefore equal to the product of the individual choice probabilities:

$$\pi (y_t | \psi) = \prod_t \Pr (1)_{t}^{y_{1t}} \Pr (2)_{t}^{y_{2t}} \dots \Pr (k)_{t}^{y_{kt}} \Pr (z)_{t}^{y_{zt}}$$

where  $\psi$  represents the model parameters and  $y_{i,t}$  are the multinomial choices with one element equal to one and the rest equal to zero. Maximum likelihood estimates of the model parameters are the parameters that maximize the joint probability of the observed choices. Bayesian estimates of the model parameters introduce a prior distribution,  $\pi (\psi)$ , that is combined with the likelihood to derive the posterior distribution  $\pi (\psi | y) \propto \pi (y | \psi) \pi (\psi)$  (see Rossi et al. 2005). In a Bayesian analysis, point estimates of model parameters are typically taken as the mean of the posterior distribution.

### 7.2.1 Applications and Extensions

Thousands of research articles have been written that have extended and/or applied the logit choice model to choice data. Groundbreaking work on the logit model and transportation choice can be traced back to the work of McFadden (1973, 1986),

which was applied to marketing demand data by (Guadagni and Little 1983) and others. The linearity of the utility function results in linear indifference curves and corner solutions, where only one of the choice alternatives is selected.

The introduction of Bayesian statistical methods into marketing (Rossi et al. 2005) led to the incorporation of respondent heterogeneity in choice models, which greatly increased their popularity, especially in the context of conjoint analysis (Green and Srinivasan 1978). Allenby and Ginter (1995) examined the use of a binary logit model in market segmentation to understand which respondents are most likely to respond to a product reformulation, and (Lenk et al. 1996) studied the use of partial factorial designs and their ability to inform the parameters of this class of models. The results of these and subsequent studies demonstrated that choice models were useful for accurately representing preferences for a heterogeneous set of consumers. Moreover, empirical results over the years have supported the use of economic models to represent consumer preferences and sensitivities to variables like prices. Many models of choice specified without heterogeneity employed many interaction terms to represent demand. These interaction terms have largely disappeared from the published literature in the presence of heterogeneity.

The simple logit model of discrete choice described above has been extended in many ways. Allenby and Rossi (1991) propose a model that maintains linearity of the indifference curves but allows them to rotate in the positive orthant to represent goods of different levels of quality. As the budgetary allotment  $E$  is relaxed, their model predicts that consumers would trade-up to higher quality offerings. Berry et al. (1995) introduce demand shocks into a logit model to accommodate factors that shift the utility of all consumers in a market while parsimoniously representing a demand system. Their model has become a standard in the empirical industrial organization literature. The logit model has been generalized by Marshall and Bradlow (2002) to accommodate a variety of preference measures beside simple choice, Edwards and Allenby (2003) discuss multivariate extensions, and Chandukala et al. (2007) provide an review of choice models in marketing. Most recently, Allenby et al. (2014a) and Allenby et al. (2014b) discuss using a simple choice model to estimate the economic value of product features. Proceedings of the triennial Invitational Choice Symposium, in the journal *Marketing Letters*, provides an ongoing review of innovations and applications of the simple discrete choice model.

### 7.3 A General Model for Choice

We now consider a general model for choice that allows for the possibility that more than one offering may be selected, often referred to as models of multiple discreteness (Kim et al. 2002). The demand for multiple offerings is common in the purchase of goods offering different varieties, such as flavors of a good, and whenever more than one unit is purchased at a time. Allowing for the possibility of purchasing multiple units require us to employ a calculus-based approach to associate observed choices to constrained utility maximization. It is not feasible to search over a con-

tinuous demand space to find the utility maximizing solution. Instead, first-order conditions (i.e., setting derivatives to zero) are used to connect utility maximization to observed demand.

We begin with a utility specification that leads to a version of the standard discrete choice model discussed earlier. Consumers are assumed to be utility maximizers subject to a budgetary constraint. Utility is specified logarithmically for the inside goods, linearly for the outside good, and with a parameter  $\gamma$  that introduces flexibility in the rate of satiation (see Bhat 2005). The assumption of a linear outside good is almost universally made in quantitative choice models, and, as shown below, significantly degrades the fit of models relative to a non-linear specification:

$$\text{Max } u(x, z) = \sum_k \frac{\psi_k}{\gamma} \ln(\gamma x_k + 1) + z \quad \text{subject to} \quad p'x + z \leq E \quad (7.2)$$

where  $x$  is a vector of demand of dimension  $k$ ,  $\{\psi_k\}$  are baseline utility parameters,  $\gamma$  is a satiation parameter constrained to be positive,  $p$  is a vector of prices,  $z$  is an outside good with price equal to one, and  $E$  is the expenditure allocation. Equation (7.2) is additively separable and therefore assumes the goods are substitutes. The form of the utility function is selected because of the simplicity of the expression for marginal utility:

$$u_k = \frac{\partial u(x, z)}{\partial x_k} = \frac{\psi_k}{\gamma x_k + 1}$$

$$u_z = \frac{\partial u(x, z)}{\partial z} = 1$$

Marginal utility for the inside goods diminishes as quantity  $\{x_k\}$  increases, and is equal to  $\psi_k$  when  $x_k = 0$ . The rate of satiation, or the rate at which marginal utility decreases, is governed by the satiation parameter  $\gamma$ . A plot of marginal utility as a function of quantity ( $x_k$ ) is provided in Fig. 7.1.

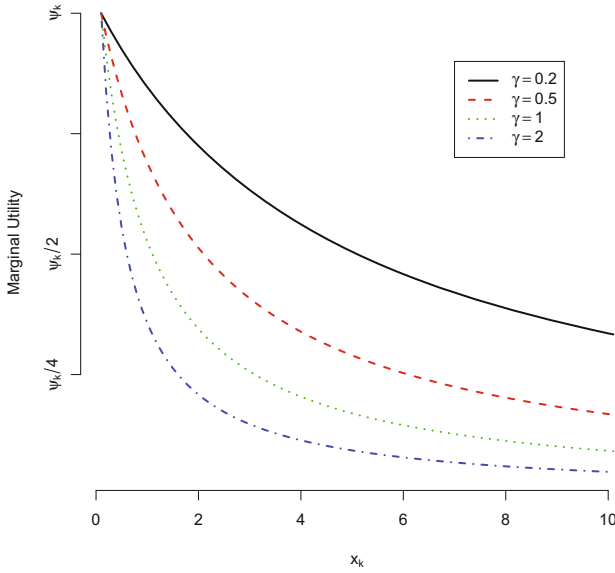
We solve for the utility maximizing solution by the method of Lagrangian multipliers that combines the constraint and utility function by introducing a parameter  $\lambda$  that ensures their slopes are proportional, or that the utility function and budget constraint are tangent, at the point of constrained maximization:

$$\text{Max } L = \sum_k \frac{\psi_k}{\gamma} \ln(\gamma x_k + 1) + z + \lambda(E - p'x - z)$$

Setting partial derivatives to zero we obtain the optimality conditions:

$$\frac{\partial L}{\partial x_k} = \frac{\psi_k}{\gamma x_k + 1} - \lambda p_k = 0$$

$$\frac{\partial L}{\partial z} = 1 - \lambda = 0$$



**Fig. 7.1** Marginal Utility

From the second equation we see that  $\lambda = 1$ , and we can substitute for  $\lambda$  in the first equation to obtain:

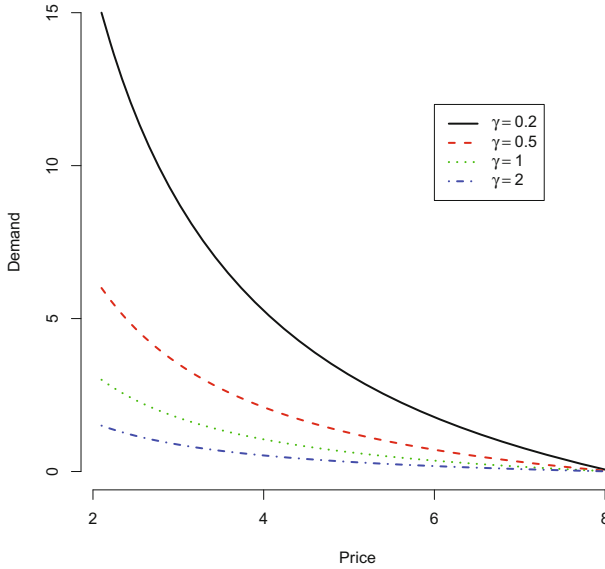
$$\frac{\Psi_k}{\gamma x_k + 1} = p_k$$

This expression holds whenever demand is positive, or  $x_k > 0$ , indicating that marginal utility is equal to price for positive demand, i.e., the “bang” is equal to the “buck.” When demand is observed to be zero, we have the condition that marginal utility is less than the price, or that the bang is less than the buck. Re-arranging terms results in an explicit expression for observed demand  $x$ :

$$x_k = \frac{\Psi_k - p_k}{\gamma p_k} \quad \text{for } \Psi_k > p_k \quad \text{else } x_k = 0$$

A plot of demand is provided in Fig. 7.2 for  $\Psi_k = 8$  and  $E = \$8.00$ .

Demand is increasing in the level of baseline marginal utility ( $\Psi_k$ ) and decreasing in price and the satiation parameter  $\gamma$ . An advantage of this expression is that demand is declining in prices, which is sometimes violated in regression-based models of demand. A disadvantage is that cross-effects are not present. Below we show that this is due to assuming that the utility in (7.2) is additively separable and the outside good ( $z$ ) does not satiate. Non-satiation of the outside good results in the utility maximizing solution unaffected by the budgetary allotment, or expenditure ( $E$ ), similar to that encountered with the simple model of discrete choice discussed earlier.



**Fig. 7.2** Demand Curves

The general form of the above solution is referred to as the Kuhn-Tucker (KT) conditions of utility maximizing demand:

$$\begin{aligned} \text{if } x_k > 0 \text{ and } x_j > 0 \text{ then } \lambda &= \frac{u_k}{p_k} = \frac{u_j}{p_j} \text{ for all } k \text{ and } j \\ \text{if } x_k > 0 \text{ and } x_j = 0 \text{ then } \lambda &= \frac{u_k}{p_k} > \frac{u_j}{p_j} \text{ for all } k \text{ and } j \end{aligned}$$

where  $\lambda$  is equal to the marginal utility of the outside good  $u_z$  because its price is normalized to equal one. We regard the above utility model as a basic structure from which to build various models of demand. The simplicity of the model leads to closed-form expressions for demand forecasting, and nests the standard discrete choice model where the data indicate the most preferred choice option instead of demand quantities. The KT condition for preference data, where respondents are asked to indicate their preference without reference to quantities (i.e.,  $x = 0$ ) reduces to:

$$\text{if } k \text{ preferred then } \frac{\psi_k}{p_k} > \frac{\psi_j}{p_j} \text{ for all } j$$

Taking logarithms leads to an expression similar to Equation (1) except that price is replaced by  $\ln(p_j)$ . The analysis of volumetric demand data (i.e.,  $x_k > 0$ ), however, is more informative of model parameters because the equality restrictions in the KT conditions are more informative than inequality restrictions.



### 7.3.1 Statistical Specification

Variation in observed demand for a respondent often requires the introduction of error terms to rationalize choice. It is convenient to introduce error terms in the baseline utility parameters by specifying a functional form that ensures that marginal utility is always positive:

$$\psi_{kt} = \exp [a'_{kt}\beta + \varepsilon_{kt}]$$

where  $a_{kt}$  is a vector of attributes of the  $k^{\text{th}}$  good and the error term is allowed to vary over time ( $t$ ). The parameters  $\beta$  are sometimes referred to as “part-worths” in conjoint analysis, reflecting the partial worth of product attributes and benefits. Substituting the expression for  $\psi_{kt}$  into the expression that equates the Lagrange multiplier ( $\lambda$ ) to the ratio of marginal utility to price, and recalling that  $\lambda = u_z/1 = 1$  results in the expression:

$$\frac{\exp [a'_{kt}\beta + \varepsilon_{kt}]}{\gamma x_{kt} + 1} = p_{kt}$$

Solving for  $\varepsilon_{kt}$  results in the following expression for the KT conditions:

$$\varepsilon_{kt} = g_{kt} \quad \text{if } x_{kt} > 0 \quad (7.3)$$

$$\varepsilon_{kt} < g_{kt} \quad \text{if } x_{kt} = 0 \quad (7.4)$$

where

$$g_{kt} = -a'_{kt}\beta + \ln(\gamma x_{kt} + 1) + \ln(p_{kt})$$

The assumption of *i.i.d.* extreme-value errors, i.e.,  $\text{EV}(0, \sigma)$ , results in a closed-form expression for the probability that  $R_t$  of  $N$  goods are chosen. Indexing the chosen goods by  $n_{1,t}$  and the remainder by  $n_{2,t}$  results in the following expression for the likelihood:

$$\begin{aligned} \Pr(x_t) &= \Pr(x_{n_{1,t}} > 0, \quad x_{n_{2,t}} = 0, \quad n_{1,t} = 1, \dots, R_t, \quad n_{2,t} = R_t + 1, \dots, N) \\ &= |J_{R_t}| \int_{-\infty}^{g_{R_t}} \dots \int_{-\infty}^{g_{R_t+1}} f(g_{1,t}, \dots, g_{R_t}, \varepsilon_{R_t+1}, \dots, \varepsilon_N) d\varepsilon_{R_t+1}, \dots, d\varepsilon_N \\ &= |J_{R_t}| \left\{ \prod_{i=1}^{R_t} \frac{\exp(-g_{it}/\sigma)}{\sigma} \exp(-e^{-g_{it}/\sigma}) \right\} \left\{ \prod_{j=R_t+1}^N \exp(-e^{-g_{jt}/\sigma}) \right\} \\ &= |J_{R_t}| \left\{ \prod_{i=1}^{R_t} \frac{\exp(-g_{it}/\sigma)}{\sigma} \right\} \exp \left\{ - \sum_{j=1}^N \exp(-g_{jt}/\sigma) \right\} \end{aligned}$$

where  $f(\cdot)$  is the joint density distribution for  $\varepsilon$  and  $|J_{R_t}|$  is the Jacobian of the transformation from random-utility error ( $\varepsilon$ ) to the likelihood of the observed data ( $x$ ). For this model, the Jacobian is equal to:

$$|J_{R_t}| = \prod_{i=1}^{R_t} \frac{\gamma}{\gamma x_{i,t} + 1}$$

The expression for the probability of the observed demand vector  $x_t$  is seen to be the product of  $R_t$  “logit” expressions multiplied by the Jacobian, where the purchased quantity,  $x_{it}$  is part of the value ( $g_{it}$ ) of the choice alternative. For the standard discrete choice model,  $g_{kt} = -a'_{kt}\beta + \ln(p_{kt})$  and the Jacobian is equal to one because demand ( $x$ ) enters the KT conditions through the conditions  $x_{kt} > 0$  or  $x_{kt} = 0$  only in Eqs. (7.3) and (7.4). The price coefficient in the standard choice model is the scale value of the Extreme Value error ( $1/\sigma$ ). Variation in the specification of the choice model utility function and budget constraint results in different values of ( $g_{kt}$ ) and the Jacobian  $|J_{R_t}|$ , but not to the general form of the likelihood, i.e.,

$$\Pr(x_t) = |J_{R_t}| \left\{ \prod_{i=1}^{R_t} f(g_{it}) \right\} \left\{ \prod_{j=R_t+1}^N F(g_{jt}) \right\}$$

### 7.3.2 Non-linear Outside Good

The assumption that utility is linear in the outside good ( $z$ ) results in KT conditions that do not involve the budgetary allotment  $E$ . A non-linear specification for the outside good leads to a demand model where the budgetary allotment plays a role in identifying the utility maximizing solution. This is important as it allows for the presence of cross-price effects on the demand for each of the items. For example, the utility function with logarithmic specification for the quantity of the outside good leads to:

$$u(x, z) = \sum_k \frac{\psi_k}{\gamma} \ln(\gamma x_k + 1) + \ln(z)$$

has marginal utility for the outside good equal to:

$$u_z = \frac{\partial u(x, z)}{\partial z} = \frac{1}{z}$$

and the KT condition  $\lambda = u_z/1 = u_k/p_k$  leads to new expressions for  $g_{kt}$  and the Jacobian:

$$g_{kt} = -a'_{kt}\beta + \ln(\gamma x_{kt} + 1) + \ln\left(\frac{p_{kt}}{E - p'_t x_t}\right) \tag{7.5}$$

$$\begin{aligned}
 |J_{R_t}| &= \det \left[ \frac{\partial g_{R_t}}{\partial x_{R_t}'} \right] = \det \begin{bmatrix} \frac{\gamma}{\gamma x_{1t}+1} + \frac{p_{1t}}{E-p_t'x_t} & \frac{p_{2t}}{E-p_t'x_t} & \dots & \frac{p_{R_t}}{E-p_t'x_t} \\ \frac{p_{1t}}{E-p_t'x_t} & \frac{\gamma}{\gamma x_{2t}+1} + \frac{p_{2t}}{E-p_t'x_t} & \dots & \frac{p_{R_t}}{E-p_t'x_t} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{p_{1t}}{E-p_t'x_t} & \frac{p_{2t}}{E-p_t'x_t} & \dots & \frac{\gamma}{\gamma x_{R_t}+1} + \frac{p_{R_t}}{E-p_t'x_t} \end{bmatrix} \\
 &= \prod_{k=1}^{R_t} \left( \frac{\gamma}{\gamma x_{kt} + 1} \right) \left\{ \sum_{k=1}^{R_t} \frac{\gamma x_{kt} + 1}{\gamma} \cdot \frac{p_{kt}}{E - p_t'x_t} + 1 \right\}
 \end{aligned}$$

The off-diagonal elements of the Jacobian are non-zero because of the right-most term in the expression for  $g_{kt}$ . The expenditure allotment  $E$  can be treated as a parameter and is statistically identified through the KT condition associated with positive demand that result in the equality restrictions  $\varepsilon_{kt} = g_{kt}$ . However, its estimated value will depend on the degree of assumed concavity of the utility function for the outside good (e.g., logarithm versus a power function).

Predicted demand for the model with non-linear outside good does not have a closed form because the KT conditions leads to an implicit function for  $x$ :

$$x_k = \frac{\psi_k - \lambda p_k}{\lambda \gamma p_k} \quad \text{for } \psi_k > \lambda p_k \quad \text{else } x_k = 0$$

where  $\lambda = \frac{1}{E-p'x}$ . However, there are methods for obtaining demand estimates. A general solution is using standard constrained optimization routines such as `constrOptim` in the R statistical package that directly maximizes the utility function subject to the budget constraint.

### 7.3.3 Applications and Extensions

Volumetric (non-binary) demand is common in marketing, occurring in the consumption of most packaged goods and services. The quantity purchased is often not restricted to just a single unit of a good, and the development of a quantity-based model reduces the number of distinct choice alternatives that need to be modeled. For example, in the beverage category, soda is routinely sold as 6-packs and 12-packs of 12 ounce cans. The decision to purchase a 12-pack of Coke reflects both an item and quantity decision, and economic models used to rationalize this choice treat the demand quantities as the outcome of a constrained utility problem. The advantage of this is that it requires fewer parameters and error terms than in models that treat different package sizes as having unique intercept and error terms. Moreover, it can handle zero demand quantities and is especially well suited for sparse data environments.

Models of discrete and continuous demand were pioneered by Hanemann (1984) who coupled a discrete choice (logit) model with a conditional demand model. The

horizontal variety literature on multiple discreteness models (e.g., Kim et al. 2002; Bhat 2005, 2008) extends these models to allow for the selection of more than one choice alternative. In addition, the direct utility model described above is flexible with regard to the utility function that is employed. Lee et al. (2013), for example, study the presence of complements where goods have a super-additive effect on consumer utility.

The direct utility approach has been usefully extended to the areas where data are from non-consumer product categories. Luo et al. (2013), for example, investigated how consumers allocate time resources among leisure activities over time based on dynamic version of direct utility specifications above. In similar vein, Lin et al. (2013) study consumers' media consumptions such as TV, radio, and internet accounting for substitution and complementarities in multiplexing activities.

## 7.4 Constraints

An advantage of using a direct utility model to study consumer purchase decisions is that it separates what is gained in an exchange from what is given up. Consumers give up various resources for the right to acquire and use marketplace offerings that provide them with utility. These resources include money, time, attention, and any other constraint on their lives. Dieters, for example, pay attention to the caloric content of the food they consume and may not purchase items on sale if they are high in calories. Consumers constrained by space may not purchase large package sizes, and consumers may not purchase goods with large access costs, such as the fixed costs associated with learning to play a new sport. In many cases, consumer choice is governed more by constraints on available options than by the utility afforded by different offerings. We begin this section with a discussion of choice models with multiple constraints (Satomura et al. 2011) and then examine models with non-linear constraints (Howell et al. 2015; Howell and Allenby 2015).

### 7.4.1 Multiple Constraints

We develop our model of multiple constraints for consumers constrained by money and quantity. Quantity constraints arise when consumers have limited storage space in their homes that they wish to develop to a product category. Space constraints are represented as  $Q$  denoting the upper limit of quantity. Consumers are assumed to make choices that maximize utility subject to multiple constraints:

$$\begin{aligned} \text{Max } u(x, z, w) &= \sum_k \frac{\psi_k}{\gamma} \ln(\gamma x_k + 1) + \ln(z) + \ln(w) \\ \text{subject to } & p'x + z \leq E \quad \text{and} \quad q'x + w \leq Q \end{aligned}$$

The utility maximizing solution is found by forming the auxiliary function  $L$ , but this time with two multipliers  $\lambda$  and  $\mu$ :

$$\text{Max } L = u(x, z, w) + \lambda\{M - p'x - z\} + \mu\{Q - q'x - w\}$$

resulting in the following first-order conditions for constrained utility maximization:

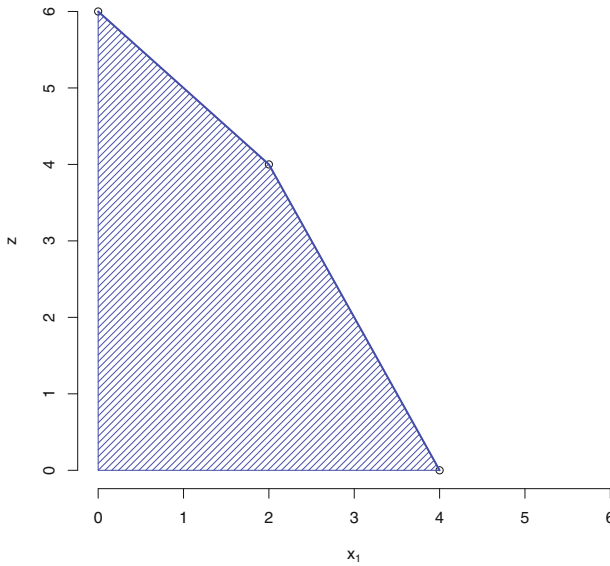
$$\begin{aligned} \varepsilon_{kt} &= g_{kt} & \text{if } x_{kt} > 0 \\ \varepsilon_{kt} &< g_{kt} & \text{if } x_{kt} = 0 \\ g_{kt} &= -a'_{kt}\beta + \ln(\gamma x_{kt} + 1) + \ln\left(\frac{p_{kt}}{E - p'_t x_t} + \frac{q_k}{Q - q'_t x_t}\right) \end{aligned}$$

that differs from the earlier expression in Equation (5) in that the last term involves both the budget and quantity restrictions. As either  $p'_t x_t$  approaches  $E$ , or  $q'_t x_t$  approaches  $Q$ , the last term on the right side becomes large, making it less likely to observe positive demand ( $x_{kt} > 0$ ) and more likely to observe zero demand ( $x_{kt} = 0$ ). Thus, goods that tend to exhaust either of the allocated budgets  $E$  and  $Q$  are less likely to be selected.

The Lagrangian multipliers  $\lambda$  and  $\mu$  can be shown to be the expected change in attainable utility for a unit change in the constraint (see Sydsæter et al. 2005, Chap. 14). Thus, one can evaluate the impact of the constraints on choice and utility, and determine which is more profitable for the consumer to relax. Budgets ( $E$ ) can be relaxed by endowing consumers with greater wealth by the use of coupons and other means of temporary price reductions, and quantity constraints ( $Q$ ) can be relaxed by improvements in packaging and other forms of space saving. By comparing the cost of these changes to the expected increase in utility allows firms to determine which constraint to relax.

### 7.4.2 Non-linear Constraints

Non-linear constraints arise when costs, viewed in a broad sense, do not scale in proportion to the quantity consumed. Examples include fixed costs that are incurred just once for the first unit of demand (Howell and Allenby 2015), access costs that arise as consumers transform market-place goods for consumption (Kim et al. 2015b), and when unit prices depend on the quantity purchased. An example of a fixed cost is the cost of a coffee maker, while an access cost for coffee involves the purchase of coffee beans and its daily preparation. Access costs might be shared among different choice alternatives that affects the variety of goods consumed (Kim et al. 2015a). Quantity-dependent pricing, often referred to as multi-part pricing, has been studied extensively in analytic models but are often difficult to implement in practice because of the different prices that consumers face. Non-linear pricing can result in irregular budget sets, where the budgetary constraint having kink points and possibly points



**Fig. 7.3** Irregular budget set

of discontinuity. When this occurs, it is not possible to use first-order conditions to find a global optimal quantity of demand that maximizes constrained utility.

Figure 7.3 displays the budget constraint for two goods,  $x_1$  and  $z$ , when  $E = \$6.00$  and the price for the first two units of  $x$  is \$1.00, and then the price rises to \$2.00 per unit. The budgetary constraint has a kink point at  $x_1 = 2$  units that will result in a build-up of mass in the likelihood of demand. That is, many consumers will want to purchase two units of  $x$  because of the low price. The solution to modeling demand when the budgetary constraint has a kink point is to employ first-order conditions to find optimal demand quantities within regions of the budget set that are linear, and then to compare the solutions to find the utility maximizing demand.

Consider the case where there are two inside goods,  $x_1$  and  $x_2$  with kink points at  $\tau_1$  and  $\tau_2$ . Price is assumed to take on a low value ( $p_\ell$ ) below the kink point and a higher value above the kink point ( $p_h$ ). We can partition the demand space into four regions:

$$\begin{aligned}
 \mathbb{P}_1 : & \text{Max } u(x_{1t}, x_{2t}, z_t) \\
 & \text{s.t. } p_\ell x_{1t} + p_\ell x_{2t} + z_t = E \\
 & \quad 0 \leq x_{1t} \leq \tau_1, 0 \leq x_{2t} \leq \tau_2 \\
 \mathbb{P}_2 : & \text{Max } u(x_{1t}, x_{2t}, z_t) \\
 & \text{s.t. } p_\ell \tau_1 + p_h(x_{1t} - \tau_1) + p_\ell x_{2t} + z_t = E \\
 & \quad \tau_1 < x_{1t}, 0 \leq x_{2t} \leq \tau_2
 \end{aligned}$$

$$\begin{aligned}
\mathbb{P}_3 : \quad & \text{Max } u(x_{1t}, x_{2t}, z_t) \\
& \text{s.t. } p_{\ell 1}\tau_1 + p_{\ell 2}x_{2t} + p_{h2}(x_{2t} - \tau_2) + z_t = E \\
& \quad 0 \leq x_{1t} \leq \tau_1, \tau_2 < x_{2t} \\
\mathbb{P}_4 : \quad & \text{Max } u(x_{1t}, x_{2t}, z_t) \\
& \text{s.t. } p_{\ell 1}\tau_1 + p_{\ell 2}\tau_2 + p_{h1}(x_{1t} - \tau_1) + p_{h2}(x_{2t} - \tau_2) + z_t = E \\
& \quad \tau_1 < x_{1t}, \tau_2 < x_{2t}.
\end{aligned}$$

Then, the first-order conditions associated with observed demand are:

$$\begin{aligned}
\varepsilon_{kt} < g_{\ell kt} & \quad \text{if } x_{kt} = 0 \\
\varepsilon_{kt} = g_{\ell kt} & \quad \text{if } 0 < x_{kt} < \tau_k \\
g_{\ell kt} < \varepsilon_k < g_{hkt} & \quad \text{if } x_{kt} = \tau_k \\
\varepsilon_{kt} = g_{hkt} & \quad \text{if } x_{kt} > \tau_k
\end{aligned}$$

where

$$\begin{aligned}
g_{\ell kt} &= -a'_k \beta + \ln(\gamma x_{kt} + 1) + \ln\left(\frac{p_{\ell kt}}{z_t}\right) \\
g_{hkt} &= -a'_k \beta + \ln(\gamma x_{kt} + 1) + \ln\left(\frac{p_{hkt}}{z_t}\right)
\end{aligned}$$

when the outside good is specified logarithmically. We define the set  $A = \{k : x_{kt} = 0\}$ , the set  $B = \{k : 0 < x_{kt} < \tau_k\}$ , the set  $C = \{k : x_{kt} = \tau_k\}$ , and the set  $D = \{k : x_{kt} > \tau_k\}$ . The likelihood is therefore:

$$\begin{aligned}
Pr(x_{kt}) &= Pr(x_{At} = 0, 0 < x_{Bt} < \tau_k, x_{Ct} = \tau_k, x_{Dt} > \tau_k) \\
&= |J_{BUD}| Pr(\varepsilon_{At} < g_{\ell At}, \varepsilon_{Bt} = g_{\ell Bt}, g_{\ell Ct} < \varepsilon_{Ct} < g_{hCt}, \varepsilon_{Dt} = g_{hDt}) \\
&= |J_{BUD}| \times F(g_{\ell At}) \times f(g_{\ell Bt}) \times (F(g_{hCt}) - F(g_{\ell Ct})) \times f(g_{hDt})
\end{aligned}$$

where  $f$  is the pdf of the error distribution and  $F$  is the CDF of that distribution.  $J_{B,D}$  is the Jacobian of  $\varepsilon_{CUD}$  and is defined as:

$$J_{ij} = \frac{\partial g_{\cdot i}}{\partial x_{jt}} = \frac{1}{x_{it} + 1} I(i = j) + \frac{p_{jt}}{z_t}$$

with  $\cdot = \ell$  if  $i \in B$ , and  $\cdot = h$  if  $i \in D$ .

The above framework can be extended to an arbitrary number of kink points and an arbitrary number of goods. The challenge in applying the model is in keeping track of the number of possible outcome regions (e.g., regions A–D above) and correctly computing the likelihood.

More generally, constraints limit the region over which utility is maximized and often results in the need for estimators that do not rely solely on first-order condi-

tions to associate observed demand and choices to model parameters. Demand, for example, could take the form of a mixture of demand types, with it being discrete for some decision variables (e.g., the decision to stream music) and continuous for others (the number of songs to download) (Kim et al. 2015a). Another example involves the use of screening rules and the presence of consideration sets in consumer decision making (Kim et al. 2015c) that involves a subset of the items available for sale. Although such constraints complicate the estimation of direct utility models, many creative procedures have been proposed to relate observed demand to constrained utility maximization for model estimation.

## 7.5 Error Specification

We considered the form of the utility function and the nature of constraints in developing variants of direct utility choice models in the discussion above. We now consider the influence of the error term in the model. The error term plays an important role in models of choice because of what it implies about factors that are not explicitly present in the utility function or the budget constraint. In general, if an iid (independent and identically distributed) additive error term is used in the model, there are some realizations of the error that guarantee that a particular good will be chosen. That is, there will not exist dominated alternatives for which demand is zero for any respondent. While this assumption might seem harmless for the analysis of small choice sets, it is problematic whenever the choice set becomes large and every alternative is assigned some positive purchase probability that is independent of a brand's attributes. Most product categories consist of dozens, and sometimes hundreds, of goods for sale. Assuming iid error terms for each alternatives leads to demand predictions that are not sufficiently sensitive to changes in prices and other product features, particularly when there are some brands that compete at a heightened level with other brands.

Another problem with standard error assumptions is that observed demand is often discrete, not continuous, and is often constrained by packaging decisions made by firms. Consumers, for example, may not be able to purchase five cans of soda, or eight eggs at a grocery store. Marketplace demand must therefore be viewed as a censored outcome of a constrained model of choice, where the offerings made available reside on a coarse grid dictated by the packaging. Choice models must therefore be modified to acknowledge this restriction when inferring about model parameters and when predicting future demand.

### 7.5.1 *Correlated Errors*

One solution to recognizing groups of similar products is to allow the random utility errors to be correlated. The nested logit model (Hausman and McFadden 1984), for



example, is motivated by the presence of correlated errors for goods that are similar. Correlated errors are most flexibly introduced into a choice model with Normally distributed error terms. Dotson et al. (2015) discuss a parsimonious model that allows for error covariances that are related to product features. The effect of the model is to relax the assumption of IIA associated with the standard logit model, where the ratio of choice probabilities for any two alternatives does not involve any of the other choice alternatives. With IIA, the introduction of a good similar to an existing good draws share proportionally to the baseline shares of all other goods and does not favor other, similar goods.

It is desirable that the correlations among the error terms of similar goods be large, and the correlation among goods that are dissimilar be small. In the extreme, if two goods are identical, the errors should be perfectly collinear and would split the demand associated with them. A simple model of correlated errors for  $N$  choice alternatives is:

$$\Sigma = \begin{bmatrix} 1 & \cdots & \sigma_{1N} \\ \vdots & \ddots & \vdots \\ \sigma_{N1} & \cdots & 1 \end{bmatrix}$$

where

$$\sigma_{kj} = \exp \left[ \frac{-d_{kj}}{\theta} \right]$$

and  $d_{kj}$  is a measure of perceptual distance between goods  $k$  and  $j$ . Thus, if two goods are perceived to be nearly alike, then  $d_{kj}$  is close to zero and  $\sigma_{kj}$  is close to one. They investigate alternative parameterizations of the distance measure, and find that one based on baseline utility consistently fits the data best:

$$d_{kj} = \left| \psi_k - \psi_j \right|$$

where  $\psi$  is the marginal utility of the offering. Thus, the  $\psi$  parameters appear in the mean and covariance of the utility specification, which necessitates the need for customized software for model estimation. Dotson et al. (2015) discuss estimation of this model as a hierarchical Bayes model.

### 7.5.2 Indivisible Demand

Restriction on demand caused by the discreteness of packaging results in an additional constraint on choice that censors the model error term to produce integer demand:

$$x_{kt} \in \{0, 1, 2, \dots\}, \quad \forall k \in \{1, \dots, N\}$$

Thus, instead of consumers purchasing the alternative with greatest utility among those that satisfy the budgetary allotment, they are assumed to choose from among the available alternatives that maximize utility.

Lee and Allenby (2014) show how to deal with this constraint in model estimation and prediction. The model likelihood become a product of mass points, as there are multiple realizations of each model error term that can correspond to observed demand on the package grid. A variant of Bayesian data augmentation (Tanner and Wong 1987) is used to evaluate the likelihood at feasible points on the package grid:

$$U^*(x_{1t}^*, \dots, x_{nt}^*) \geq \max \{ U^*(x_{1t}^* + \Delta_1, \dots, x_{nt}^* + \Delta_n) \mid (x_{1t}^* + \Delta_1, \dots, x_{nt}^* + \Delta_n) \in F \}_{\Delta_i \in \{-1, 0, 1\}} \quad (7.6)$$

and use the inequality relationship to determine ranges of the error term that are consistent with the observed demand being utility maximizing. Their analysis indicates that data corresponding to the corner solutions are most affected by the presence of packaging constraints, where zero demand should be interpreted as not liking an offering enough to buy one unit, as opposed to not liking it enough to buy any.

## 7.6 Indirect Utility Models

Models of discrete/continuous choice have a long history in the marketing and economics literature beginning with the work of Hanemann (1984) and extended by many that build on a framework where demand is positive for just one of many different choice alternatives. Krishnamurthi and Raj (1988) discuss the estimation of models where the continuous component of demand is driven by a set of covariates different from the covariates used to form utility in the discrete choice portion of the model, and rely on an indirect utility specification to provide a theoretical basis for demand quantities. Similarly, Harlam and Lodish (1995) introduce variables into their model specification that provide summary measures of merchandising activity that is difficult to interpret in terms of a direct utility model. Chintagunta (1993), Dubé (2004) and Song and Chintagunta (2007) also motivate their model specification using the concept of an indirect utility function without explicitly relating it to a specific direct utility model. An indirect utility function is defined as the maximal attainable utility as a function of prices and expenditure ( $E$ ). Indirect utility models, however, are not amenable to disaggregate demand analysis in marketing where the attributes of products can change and there exist mass points of demand at particular prices.

For example, products in a conjoint analysis change across choice sets as product attributes and levels are experimentally manipulated. As the attribute-levels of the choice alternatives change, so should their degree of substitution and level of price interaction, indicating that many of the parameters of an indirect utility model would also need to be functions of product characteristics. It is difficult to implement a characteristics model of demand within an indirect utility model because

indirect utility models often have many parameters. In addition, it is not clear how to incorporate model error into an indirect utility specification to allow for mass points of demand that occur at corners, kink points and due to packaging constraints. As discussed above, these issues can be addressed through a direct utility specification where corner solutions give rise to inequality constraints in the likelihood through the Kuhn-Tucker conditions.

To illustrate, consider a constrained maximization problem involving the utility function similar to (2):

$$\max u(x) = \sum_k \frac{\psi_k}{\gamma} \ln(\gamma x + 1) \quad \text{subject to } p'x \leq E$$

where the  $\psi$ 's are assumed to sum to one (i.e.,  $\sum \psi_k = 1$ ). We can solve for the utility maximizing quantities,  $x^*$  and obtain the optimal demand function (see appendix):

$$x_k^* = \frac{1}{\gamma} \left( \frac{\psi_k}{\lambda p_k} - 1 \right)$$

where  $\lambda = \frac{1}{\gamma E + \sum_k p_k}$  is a Lagrangian multiplier. By substituting the demand function ( $x^*$ ) into the utility function, one can obtain the expression for indirect utility ( $V$ ) as follows:

$$V \equiv u(x^*) = \sum_k \frac{\psi_k}{\gamma} \left( \ln \psi_k - \ln p_k + \ln \left( \gamma E + \sum_k p_k \right) \right)$$

Details are provided in the appendix. While this formulation provides an elegant solution to optimal demand and indirect utility, it depends critically on equality constraints between the  $\psi$  parameters and optimal demand quantities  $x^*$  that are associated with interior solutions, not corner solutions. Corner solutions result in inequality restrictions, not equality restrictions, in the Kuhn-Tucker conditions.

The indirect utility function is often expressed as in terms of a Taylor series approximation to an unspecified utility function such as translog, and includes pairwise price interactions among the choice options so that a flexible pattern of substitution can be achieved (Pollak and Wales 1992). For example, consider a generalized quadratic indirect utility function often referred to as a translog indirect utility (Christensen et al. 1975):

$$\ln V = \alpha_0 + \sum_k \alpha_k \ln \frac{p_k}{E} + \frac{1}{2} \sum_k \sum_j \beta_{kj} \ln \frac{p_k}{E} \ln \frac{p_j}{E}$$

where  $(\alpha, \beta)'$  are parameters that capture the substitution among product offerings. Recently, Mehta (2015) proposes an indirect utility model for a general demand model based on Kuhn-Tucker conditions (Wales and Woodland 1983) that employs

virtual prices to deal with corner solutions (see Lee and Pitt 1987). Virtual prices are the prices at which demand is expected to be exactly equal to zero given the parameters of the indirect utility function. The virtual prices are then substituted into the demand system as if they were observed. The problem with this formulation is that it assumes more than what is observed, by conditioning on latent quantities, and overstates the value of information coming from zero demand by assuming a density contribution to the likelihood rather than a mass contribution.

The generalized quadratic indirect utility function is over-parameterized and, in general, not valid unless monotonicity and concavity constraints are imposed. There is no general theory of the extent to which this quadratic approximation provides a uniform functional approximation. Therefore, there is no reason to believe that the even a “regular” generalized quadratic indirect utility function can approximate any indirect utility for the purpose of demand specification. This would require a proof of global approximation not only of the indirect utility function but also of its derivatives. But, our principal objection to the indirect utility formulation is that it obscures the process of formulated direct utility models and associated constraints that embody the reality of the consumer purchase process. For example, if the consumer faces a non-linear budget set due to pack-size discounts, the researcher should write down the direct utility model and constraints rather than choosing an arbitrary indirect utility function which may not be consistent with this situation.

For virtually all situations in consumer choice modeling, there will be no closed form expression for the indirect utility function. Moreover, given corners and kinks, the indirect utility function may not be differentiable everywhere, eliminating the convenience of Roy’s identity for deriving demand. Indirect utility functions are useful for welfare computations but are not of practical value in the specification of consumer choice models.

## 7.7 Conclusion

This chapter has reviewed an emerging area of analysis in marketing decision models that rationalizes choice from principles of constrained utility maximization. We advocate for a direct utility specification of choice models for a variety of reasons. Models of constrained utility maximization reflect goal-directed behavior on the part of consumers, which is overwhelmingly supported in disaggregate marketing data by the prevalence of zero demand for most offerings. By far, the most frequently observed number in disaggregate marketing datasets is the number zero, implying that consumers are resource conserving and not acting randomly.

The direct utility formulation separates that which is gained in an exchange (utility) from that which is given up (constraint)—i.e., the resources needed to acquire and use a marketplace offering. Understanding the determinants of utility is useful for product strategy as marketers advocate for making what people will want to buy. Quantifying the relationship between product attributes and the benefits, and the resulting utility afforded by a competitive set of products, is one of the most

important tasks of marketing research. Likewise, understanding the impediments to acquiring and using a product is useful for driving sales and effectively communicating with prospects.

We examine three aspects of direct utility models—the utility function, constraints and error—that are combined to form the likelihood for the data. Our treatment of these constructs is structural in nature, as we avoid the temptation of simply adding an error term to a flexible model that combines brands, attributes and prices. The problem with taking this flexible approach is that it is not consistent with the lumpiness of marketing data, where corner solutions are prevalent, where demand is often constrained to lie on a grid of available package sizes, and where pricing discounts lead to a mass buildup in the likelihood even for interior solutions. Throughout our development above we stress the importance of deriving the likelihood function from principles of constrained optimization, and show how these realities of marketing data can be accommodated within the framework of constrained utility maximization.

We also advocate against the use of indirect utility models for data that contain corner solutions, mass buildup at specific demand quantities, or multiple constraints. While an indirect utility function can always be defined in these nonstandard situations, it may be difficult, if not impossible, to express in closed form. More importantly, the indirect utility function will not be useful in deriving the associated demand system and its associated likelihood.

Additional research is needed to develop and apply a broader class of utility functions, constraints, and error specifications for marketing analysis. For more than 50 years, marketing has embraced the notion of extended models of behavior where needs, wants, beliefs, attitudes, consideration and perceptions have been shown to be determinants of demand (Howard and Sheth 1969). Often, a large battery of variables is used to represent each of these constructs. Mapping these variables to one another and to marketplace demand within a principled structure is a worthy endeavor.

## Appendix

Consider the direct utility model:

$$\max u(x) = \sum_k \frac{\psi_k}{\gamma} \ln(\gamma x_k + 1) \quad \text{subject to } p'x \leq E$$

where the  $\psi$ 's are assumed to sum to one (i.e.,  $\sum \psi_k = 1$ ) and  $\gamma > 0$ . Solving for the utility maximizing quantities,  $x^*$ , by the Lagrangian method will give rise to the following objective function:

$$L = \sum_k \frac{\psi_k}{\gamma} \ln(\gamma x_k + 1) - \lambda(p'x - E)$$

The FOC's for optimal demand ( $x^*$ ) are:

$$\begin{aligned} 0 &= \frac{\partial L}{\partial x_1} = \frac{\psi_1}{\gamma x_1 + 1} - \lambda p_1 && \Leftrightarrow && \psi_1 = \lambda p_1 (\gamma x_1 + 1) \\ 0 &= \frac{\partial L}{\partial x_2} = \frac{\psi_2}{\gamma x_2 + 1} - \lambda p_2 && \Leftrightarrow && \psi_2 = \lambda p_2 (\gamma x_2 + 1) \\ &&& && \vdots \\ 0 &= \frac{\partial L}{\partial x_k} = \frac{\psi_k}{\gamma x_k + 1} - \lambda p_k && \Leftrightarrow && \psi_k = \lambda p_k (\gamma x_k + 1) \end{aligned}$$

Using  $\sum \psi_k = 1$ , we solve for  $\lambda$ :

$$1 = \sum_{k=1}^K \psi_k = \sum_{k=1}^K \lambda p_k (\gamma x_k + 1)$$

or

$$\lambda = \frac{1}{\sum_{k=1}^K p_k (\gamma x_k + 1)} = \frac{1}{\gamma E + \sum_{k=1}^K p_k}$$

Substituting  $\lambda$  back into FOCs yields optimal demand equations:

$$\begin{aligned} x_k^* &= \frac{1}{\gamma} \left( \frac{\psi_k}{\lambda p_k} - 1 \right) \\ &= \frac{1}{\gamma} \left( \frac{\psi_k}{p_k} \left( \gamma E + \sum_k p_k \right) - 1 \right) \end{aligned}$$

Substituting  $x^*$  into the direct utility function allows us to obtain the expression for indirect utility ( $V$ ):

$$\begin{aligned} V &\equiv u(x^*) \\ &= \sum_k \frac{\psi_k}{\gamma} \ln(\gamma x_k^* + 1) \\ &= \sum_k \frac{\psi_k}{\gamma} \ln \left( \gamma \left( \frac{1}{\gamma} \left( \frac{\psi_k}{p_k} \left( \gamma E + \sum_k p_k \right) - 1 \right) \right) + 1 \right) \\ &= \sum_k \frac{\psi_k}{\gamma} \ln \left( \frac{\psi_k}{p_k} \left( \gamma E + \sum_k p_k \right) \right) \\ &= \sum_k \frac{\psi_k}{\gamma} \left( \ln \psi_k - \ln p_k + \ln \left( \gamma E + \sum_k p_k \right) \right) \end{aligned}$$

## References

- Allenby, Greg M. 1989. A unified approach to identifying, estimating and testing demand structures with aggregate scanner data. *Marketing Science* 8 (3): 265–280.
- Allenby, Greg M., Jeff Brazell, John R. Howell, and Peter E. Rossi. 2014a. Valuation of patented product features. *Journal of Law and Economics* 57 (3): 629–663.
- Allenby, Greg M., Jeff D. Brazell, John R. Howell, and Peter E. Rossi. 2014b. Economic valuation of product features. *Quantitative Marketing and Economics* 12 (4): 421–456.
- Allenby, Greg M., James L. Ginter. 1995. Using extremes to design products and segment markets. *Journal of Marketing Research* 392–403.
- Allenby, Greg M., and Peter E. Rossi. 1991. Quality perceptions and asymmetric switching between brands. *Marketing science* 10 (3): 185–204.
- Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* 63 (4): 841–890.
- Bhat, C.R. 2005. A multiple discrete-continuous extreme value model: Formulation and application to discretionary time-use decisions. *Transportation Research* 39 (8): 679–707.
- Bhat, C.R. 2008. The multiple discrete-continuous extreme value (mdcev) model: Role of utility function parameters, identification considerations, and model extensions. *Transportation Research* 42 (3): 274–303.
- Chandukala, Sandeep, Jaehwan Kim, Thomas Otter, Peter Rossi, and Greg Allenby. 2007. Choice models in marketing: Economic assumptions, challenges and trends. *Foundations and Trends in Marketing* 2 (2): 97–184.
- Chintagunta, Pradeep K. 1993. Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Science* 12 (2): 184–208.
- Christensen, L.R., D.W. Jorgenson, and L.J. Lau. 1975. Transcendental logarithmic utility functions. *The American Economic Review* 3: 367–383.
- Dotson, Jeff P., John R. Howell, Jeff D. Brazell, Thomas Otter, Peter J. Lenk, Steven N. MacEachern, and Greg M. Allenby. 2015. A probit model with structured covariance for similarity effects and source of volume calculations. *Working paper, Ohio State University*.
- Dubé, Jean-Pierre. 2004. Multiple discreteness and product differentiation: Demand for carbonated soft drinks. *Marketing Science* 23 (1): 66–81.
- Edwards, Yancy D., and Greg M. Allenby. 2003. Multivariate analysis of multiple response data. *Journal of Marketing Research* 40 (3): 321–334.
- Green, Paul E., and Venkatachary Srinivasan. 1978. Conjoint analysis in consumer research: issues and outlook. *Journal of consumer research* 103–123.
- Guadagni, Peter M., and John D.C. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing science* 2 (3): 203–238.
- Hanemann, W.M. 1984. Discrete/continuous models of consumer demands. *Econometrica* 52: 541–561.
- Harlam, Bari A., and Leonard M. Lodish. 1995. Modeling consumers' choices of multiple items. *Journal of Marketing Research* 404–418.
- Hausman, Jerry, Daniel McFadden. 1984. Specification tests for the multinomial logit model. *Econometrica: Journal of the Econometric Society* 1219–1240.
- Howard, John A, Jagdish N Sheth. 1969. *The theory of buyer behavior*, vol. 14. Wiley New York.
- Howell, John R, Greg M Allenby. 2015. Choice models with fixed costs. *Fisher College of Business Working Paper* .
- Howell, John R, Sanghak Lee, and Greg M. Allenby. 2015. Price promotions in choice models. *Marketing Science* forthcoming.
- Kim, Dongsoo, Sanghak Lee, Jaehwan Kim, and Greg M. Allenby. 2015a. A direct utility model for economies of scope and access. *Working paper, Ohio State University* .
- Kim, Hyowon, Sanghak Lee, Jaehwan Kim, and Greg M. Allenby. 2015b. A direct utility model with discrete and continuous decision variables. *Working paper, Ohio State University* .

- Kim, Jaehwan, Greg M. Allenby, and Peter E. Rossi. 2002. Modeling consumer demand for variety. *Marketing Science* 21 (3): 229–250.
- Kim, Youngju, Jaehwan Kim, and Greg M. Allenby. 2015c. Conjunctive screening in models of multiple discreteness. *working paper, Ohio State Univeristy*.
- Krishnamurthi, Lakshman, and S.P. Raj. 1988. A model of brand choice and purchase quantity price sensitivities. *Marketing Science* 7 (1): 1–20.
- Lee, Lung-Fei, and Mark M. Pitt. 1987. Microeconomic models of rationing, imperfect markets, and non-negativity constraints. *Journal of Econometrics* 36 (1): 89–110.
- Lee, Sanghak, and Greg M. Allenby. 2014. Modeling indivisible demand. *Marketing Science* 33 (3): 364–381.
- Lee, Sanghak, Jaehwan Kim, and Greg M. Allenby. 2013. A direct utility model for asymmetric complements. *Marketing Science* 32 (3): 454–470.
- Lenk, Peter J., Wayne S. DeSarbo, Paul E. Green, and Martin R. Young. 1996. Hierarchical bayes conjoint analysis: Recovery of partworth heterogeneity from reduced experimental designs. *Marketing Science* 15 (2): 173–191.
- Lin, Chen, Sriram Venkataraman, and Sandy D. Jap. 2013. Media multiplexing behavior: Implications for targeting and media planning. *Marketing Science* 32 (2): 310–324.
- Luo, Lan, Brian T. Ratchford, and Botao Yang. 2013. Why we do what we do: A model of activity consumption. *Journal of Marketing Research* 50 (1): 24–43.
- Marshall, Pablo, and Eric T. Bradlow. 2002. A unified approach to conjoint analysis models. *Journal of the American Statistical Association* 97 (459): 674–682.
- McFadden, Daniel. 1973. Conditional logit analysis of qualitative choice behavior. *Frontiers of econometrics*. New York: Academic Press.
- McFadden, Daniel. 1986. The choice theory approach to market research. *Marketing Science* 5 (4): 275–297.
- Mehta, N. 2015. A flexible yet globally regular multi-good demand system. *Marketing Science* forthcoming.
- Pollak, R.A., and T.J. Wales. 1992. *Demand System Specification and Estimation*. Oxford, UK.
- Rossi, P.E., G.M. Allenby, and R. McCulloch. 2005. *Bayesian Statistics and Marketing*. Wiley.
- Satomura, Takuya, Jaehwan Kim, and Greg M. Allenby. 2011. Multiple constraint choice models with corner and interior solutions. *Marketing Science* 30 (3): 481–490.
- Song, I., and P.K. Chintagunta. 2007. A discrete-continuous model for multicategory purchase behavior of households. *J. Marketing Res.* 44: 595–612.
- Sydsæter, Knut, Arne Strøm, Peter Berck, Arne Strøm, and Peter Berck. 2005. *Economists' mathematical manual*, vol. 3. Springer.
- Tanner, Martin A., and Wing Hung Wong. 1987. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association* 82 (398): 528–540.
- Wales, Terence J., and Alan Donald Woodland. 1983. Estimation of consumer demand systems with binding non-negativity constraints. *Journal of Econometrics* 21 (3): 263–285.