For the *Handbook of Choice Modelling*
(edited by Stephane Hess and Andrew Daly)

Multiple Discrete-Continuous Choice Models: A Reflective Analysis and a Prospective View

Chandra Bhat
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
301 E. Dean Keeton St. Stop C1761, Austin, TX 78712
Tel: 512-471-4535, Fax: 512-475-8744
Email: bhat@mail.utexas.edu

Abdul Pinjari
University of South Florida
Department of Civil and Environmental Engineering
4202 E. Fowler Ave., ENB 118, Tampa, FL 33620
Tel: 813-974-9671, Fax: 813-974-2957
Email: apinjari@usf.edu

ABSTRACT
Discrete choice models are widely used to analyze consumers' choice of a single, discrete alternative from a set of mutually exclusive choice alternatives. However, in numerous situations, consumers may choose multiple alternatives that are imperfect substitutes for one another, as opposed to a single alternative. Further, consumer choices typically involve decisions pertaining to "how much to consume" (*i.e.,* a continuous quantity dimension) along with the discrete choices of "what to choose." Such multiple discrete-continuous (MDC) choices, involving whether-to-choose and how-much-to-consume decisions for multiple goods, are ubiquitous in consumer decision-making and of interest for empirical enquiry in a variety of fields, including economics, marketing, and transportation. This paper reviews the econometric modeling structures used to model discrete-continuous choices, with a particular focus on MDC choices. A review of the recent advances is provided along with a discussion of the several developments on the horizon and the challenges that lie beyond.

1. BACKGROUND

Several consumer choices are characterized by a discrete dimension as well as a continuous dimension. Examples of such choice situations include vehicle type holdings and usage, appliance choice and energy consumption, housing tenure (rent or purchase) and square footage, brand choice and quantity, and activity type choice and duration of time investment of participation. Two broad model structures may be identified in the literature to handle such discrete-continuous choice situations. The first structure (sometimes referred to as the "reduced-form" structure) has a separate equation for the discrete choice and another separate equation for the continuous choice, with jointness introduced through the statistical correlation in the random stochastic components of each equation. That is, a discrete choice model and a continuous regression model are specified separately, and then simply statistically stitched together through the stochastic terms. This first structure has seen extensive use and has proved useful to handle many empirical situations, but it is not based off an underlying (and unifying) theoretical economic model (this structure does not include the class of indirect utility function-based models that are consistent with utility maximization, as discussed in the next section). The second structure to discrete-continuous choice modeling, and the one of interest in this paper, originates from the classical microeconomic theory of utility maximization. While much work in the context of consumer utility maximization has been focused on the case of a single discrete-continuous (SDC) choice situation (where the choice involves the selection of one of many alternatives and the continuous dimension associated with the chosen alternative), there has been increasing interest recently in the multiple discrete-continuous choice (MDC) situation (where the choice situation involves the selection of one or more alternatives, along with a continuous quantity dimension associated with the consumed alternatives). Such MDC choices are pervasive in the social sciences, including transportation, economics, and marketing. Examples include individuals' time-use choices (decisions to engage in different types of activities and time allocation to each activity), investment portfolios (where and how much to invest), and grocery purchases (brand choice and purchase quantity). Regardless of whether a choice situation belongs to an SDC case or an MDC case, at a basic level, the choice process faced by the consumer can be formulated using the theory of utility maximization as described next.

1.1. The Random Utility Maximization (RUM) Approach to Modeling Discrete-Continuous Choices

Consumers are assumed to maximize a direct utility function $U(\mathbf{x})$ over a set of non-negative consumption quantities $\mathbf{x} = (x_1, ..., x_k, ..., x_K)$ subject to a budget constraint, as below:

$$Max\ U(\mathbf{x}) \text{ such that } \mathbf{x}.\mathbf{p} = E \text{ and } x_k \geq 0 \qquad (1)$$

where $U(\mathbf{x})$ is a quasi-concave, increasing and continuously differentiable utility function with respect to the consumption quantity vector, $\mathbf{p}$ is the vector of unit prices for all goods, and E is the total expenditure (or income). Note that we are suppressing the index for the consumer in Equation (1) for presentation efficiency. The formulation above is equally applicable to cases with complete or incomplete demand systems (that is, the modeling of demand for all commodities that enter preferences or the modeling of demand for a subset of commodities that enter preferences).[1] The vector

[1] A complete demand system involves the modeling of the demands of all consumption goods that exhaust the consumption space of consumers. However, complete demand systems require data on prices and consumptions of all commodity/service items, and can be impractical when studying consumptions in finely defined

$\mathbf{x}$ in Equation (1) may or may not include an *outside* good. The outside good, when included, represents the part of the total budget (e.g., income) that is not spent on the K *inside* goods of interest to the analyst. Generally, the outside good is treated as a numeraire with unit price, implying that the prices and characteristics of all goods grouped into the outside category do not influence the choice and expenditure allocation among the inside goods (see Deaton and Muellbauer, 1980). The outside good allows for the overall demand for the inside goods to change due to changes in prices and other influential factors of the inside goods. Other assumptions typically made in the above utility maximization formulation are: (a) the direct utility contribution due to the consumption of different alternatives is additively separable, and (b) the constraint is linear in prices, and it is the only constraint governing consumers' decisions. We will return to these assumptions later.

The form of the utility function $U(\mathbf{x})$ in Equation (1) determines whether the formulation corresponds to a single discrete-continuous (SDC) model or a multiple discrete-continuous (MDC) model. The SDC case assumes that the choice alternatives are perfect substitutes; that is, the choice of one alternative precludes the choice of others. The MDC case accommodates imperfect substitution among goods, thus allowing for the possibility of consuming multiple alternatives. A linear utility form with respect to consumption characterizes the perfect substitutes (or SDC) case, while a non-linear utility form allowing diminishing marginal utility

---

commodity/service categories. Thus, it is common to use an incomplete demand system, typically in the form of a two stage budgeting approach or in the form of the use of a Hicksian composite commodity assumption. In the former two stage budgeting approach, separabilility of preferences is invoked, and the allocation is pursued in two independent stages. The first stage entails allocation between a limited number of broad groups of consumption items, followed by the incomplete demand system allocation of the group expenditure to elementary commodities/services within the broad consumption group of primary interest to the analyst (the elementary commodities/services in the broad group of primary interest are commonly referred to as "inside" goods). The plausibility of such a two stage budgeting approach requires strong homothetic preferences within each broad group and strong separability of preferences, or the less restrictive conditions of weak separability of preferences and the price index for each broad group not being too sensitive to changes in the utility function (see Menezes *et al.,* 2005). In the Hicksian composite commodity approach, the analyst assumes that the prices of elementary goods within each broad group of consumption items vary proportionally. Then, one can replace all the elementary alternatives within each broad group (that is not of primary interest) by a single composite alternative representing the broad group. The analysis proceeds then by considering the composite goods as "outside" goods and considering consumption in these outside goods as well as the "inside" goods representing the consumption group of main interest to the analyst. It is common in practice in this Hicksian approach to include a single outside good with the inside goods. If this composite outside good is not essential, then the consumption formulation is similar to that of a complete demand system. If this composite outside good is essential, then the formulation needs minor revision to accommodate the essential nature of the outside good. Please refer to von Haefen (2010) for a discussion of the Hicksian approach and other incomplete demand system approaches such as the one proposed by Epstein (1982) that we do not consider here. In this paper, we will consider incomplete demand systems in the form of the second stage of a two stage incomplete demand system with a finite, positive total budget as obtained from the first stage (for presentation ease, we will refer to this case as the "inside goods only" case in which at least one "inside" good has to be consumed and there are no essential outside goods) or in the form of a Hicksian composite approach with a single outside good that is essential and no requirement that at least one of the inside goods has to be consumed (for presentation ease, we will refer to this case simply as the "essential outside good" case or even more simply, as the outside good case; if the outside good is non-essential, the formulation becomes identical to the case of the "inside goods only" case, while if there are multiple outside goods, the situation is a very simple extension of the formulations presented here depending on whether the outside goods are all essential, all non-essential, or some combination of essential and non-essential). Finally, a complete demand system takes the same formulation as the "inside goods only" formulation.

with increasing consumption characterizes the imperfect substitutes (or MDC) case. An example SDC framework is Hanemann's (1984) specification:

$$U(\mathbf{x}) = U^*(\sum_{k=2}^{K} \psi_k x_k, x_1), \tag{2}$$

where U^* is a bivariate utility function and ψ_k (k = 2,…, K) represents the quality index (or baseline preference) specific to each inside good k, with the first good considered as the outside good. This functional form assures that, in addition to the outside good, exactly one inside good (k = 2, 3,…, K) is consumed. Hanemann (1984) refers to this as the "extreme corner solution". Examples of MDC frameworks will be discussed later.

Two approaches have been used to derive demand functions for the consumption quantities for the utility maximization problem in (1). The first approach, due to Hanemann (1978) and Wales and Woodland (1983), takes a direct approach to solving the constrained utility maximization problem in (1) via standard application of the Karush-Kuhn-Tucker (KKT) first-order necessary conditions of optimality. Considering the utility function $U(x)$ to be random over the population leads to stochastic KKT conditions, which form the basis for deriving probabilities for consumption patterns (including corner solutions). This approach is called the KKT approach due to the central role played by the KKT conditions (more popularly, the approach is referred to as the KT approach, but we use the label "KKT" to give credit to Karush who, in an unpublished manuscript, derived the first order optimality conditions in a constrained optimization setting even earlier than Kuhn and Tucker). The second approach, due to Hanemann (1984) and Lee and Pitt (1986), solves the maximization problem in Equation (1) by using "virtual prices" (a method that is dual to the KKT approach), which allows the analysis to start with the specification of a conditional indirect utility function. Subsequently, the implied Marshallian demand functions are obtained via Roy's identity (Roy, 1947).[2]

The vast majority of applications in the literature have involved single discrete or SDC choices. These use the indirect utility approach as opposed to the KKT approach (*i.e.,* the direct utility approach). This is mainly because the KKT approach has been perceived to be difficult to use until the past decade. This is primarily due to the absence of practical methods for estimating the structural parameters. In particular, the KKT conditions, in a stochastic setting, lead to a probability expression for the consumption vector that involves multidimensional integrals of the order of the number of goods in the analysis as discussed in Section 3.2 (and, until Bhat, 2005, this expression was thought to be analytically intractable). Further, simple and practically feasible prediction and welfare analysis methods were not available for models based on the KKT approach. However, recent interest in MDC problems has brought renewed attention to the KKT approach. Besides, the use of direct utility functions has some advantages: the relationship of the utility function to behavioral theory is more transparent, offering more interpretable parameters and better insights into identification issues. This is true even for the SDC case. For example, Bunch (2009) shows that the indirect utility function used by Chintagunta (1993) is in fact from the linear expenditure system, so the direct utility function is known. Applying the KKT approach yields the correct analytical expression for the reservation price in terms of parameters from the direct utility function, which has a clear behavioral interpretation. Over the

[2] Hanemann (1984) used this approach to derive a variety of SDC model forms consistent with Equation (2). Chiang (1991) and Chintagunta (1993) extend Hanemann's SDC formulation to include the possibility of no inside goods being selected by introducing a "reservation price". In their approach, an inside good is selected only if the quality adjusted price of at least one of the inside goods is below the reservation price. See Dubin and McFadden (1984) for another, slightly different, way of employing the (conditional) indirect utility approach for SDC choice analysis.

past decade, the field has witnessed significant strides in using the KKT approach for modeling MDC choices – both for estimation of the parameters for KKT models and for application of the models for forecasting and welfare analysis. Thus, in this paper, we focus on the KKT approach to modeling MDC choices. Specifically, we review the recent advances and outline an agenda for future research.

1.2. Structure of the Chapter

The rest of this chapter is organized as follows. The next section provides an overview of the utility forms used to model MDC choices. Section 3 outlines the econometric structure and KKT conditions of optimality that form the basis for deriving the model structure and likelihood expressions. Section 4 outlines the specific model structures used in the literature based on different specifications of the utility form and the stochastic structure. Section 5 provides a brief discussion of the case where the choice alternatives comprise a combination of imperfect and perfect substitutes. Section 6 presents methods that enable the use of the KKT-based MDC models for forecasting and policy analysis purposes. Section 7 discusses several developments on the horizon and the challenges that lie beyond. Section 8 summarizes the book chapter.

2. UTILITY FORMS FOR MODELING MDC CHOICES

As discussed earlier, non-linear utility forms that allow diminishing marginal utility with increasing consumption can be used to model "multiple discreteness" in consumer choices. A number of different utility forms have been used in the literature. In this section, we discuss the following form used in Bhat (2008) as it subsumes a variety of utility forms used in previous studies as special cases:

$$U(\boldsymbol{x}) = \sum_{k=1}^{K} \frac{\gamma_k}{\alpha_k} \psi_k \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\} \tag{3}$$

In the above utility function, $U(\boldsymbol{x})$ is a quasi-concave, increasing, and continuously differentiable function with respect to the consumption quantity (Kx1)-vector $\boldsymbol{x}$ ($x_k \geq 0$ for all k), and ψ_k, γ_k and α_k are parameters associated with good k. The function in Equation (3) is a valid utility function if $\psi_k > 0$ and $\alpha_k \leq 1$ for all k. Further, for presentation ease, we assume temporarily that there is no Hicksian composite outside good that is consumed by all decision makers, so that corner solutions (*i.e.,* zero consumptions) are allowed for all the goods k. The possibility of corner solutions implies that the term γ_k, which is a translation parameter, should be greater than zero for all k. The reader will note that there is an assumption of additive separability of preferences in the utility form of Equation (1). More on this assumption later.

The form of the utility function in Equation (1) highlights the role of the various parameters ψ_k, γ_k, and α_k, and explicitly indicates the inter-relationships between these parameters that relate to theoretical and empirical identification issues. The form also assumes weak complementarity (see Mäler, 1974), which implies that the consumer receives no utility from a non-essential good's attributes if s/he does not consume it (*i.e.,* a good and its quality attributes are weak complements). The functional form proposed by Bhat (2008) in Equation (3) generalizes earlier forms used by Hanemann (1978), von Haefen *et al.* (2004), Phaneuf *et al.*

(2000) and others. Specifically, the utility form of Equation (3) collapses to the following linear expenditure system (LES) form when $\alpha_k \to 0 \, \forall \, k$:

$$U(\boldsymbol{x}) = \sum_{k=1}^{K} \gamma_k \psi_k \ln\left((x_k / \gamma_k) + 1\right) \tag{4}$$

2.1 Role of Parameters in the Utility Specification

2.1.1 Role of ψ_k: The role of ψ_k can be inferred by computing the marginal utility of consumption with respect to good *k*, which is:

$$\frac{\partial U(\boldsymbol{x})}{\partial x_k} = \psi_k \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k - 1} \tag{5}$$

It is clear from above that ψ_k represents the baseline marginal utility, or the marginal utility at the point of zero consumption of good *k*. Alternatively, the marginal rate of substitution between any two goods *k* and *l* at the point of zero consumption of both goods is ψ_k / ψ_l. This is the case regardless of the values of γ_k and α_k. Thus, a good *k* with a higher baseline marginal utility is more likely to be consumed than a good *k'* with a lower baseline marginal utility.

2.1.2 Role of γ_k: An important role of the γ_k terms is to shift the position of the point at which the indifference curves are asymptotic to the axes from (0,0,0…,0) to $(-\gamma_1, -\gamma_2, -\gamma_3, ..., -\gamma_K)$, so that the indifference curves strike the positive orthant with a finite slope. This, combined with the consumption point corresponding to the location where the budget line is tangential to the indifference curve, results in the possibility of zero consumption of good *k*. To see this, consider two goods 1 and 2 with $\psi_1 = \psi_2 = 1$, $\alpha_1 = \alpha_2 = 0.5$, and $\gamma_2 = 1$. Figure 1 presents the profiles of the indifference curves in this two-dimensional space for various values of γ_1 ($\gamma_1 > 0$). To compare the profiles, the indifference curves are all drawn to go through the point (0,8). The reader will also note that all the indifference curve profiles strike the y-axis with the same slope. As can be observed from the figure, the positive values of γ_1 and γ_2 lead to indifference curves that cross the axes of the positive orthant, allowing for corner solutions. The indifference curve profiles are asymptotic to the *x*-axis at $y = -1$ (corresponding to the constant value of $\gamma_2 = 1$), while they are asymptotic to the *y*-axis at $x = -\gamma_1$.

Figure 2 points to another role of the γ_k term as a satiation parameter. Specifically, the figure plots the sub-utility function for alternative *k* for $\alpha_k \to 0$ and $\psi_k = 1$, and for different values of γ_k. All of the curves have the same slope $\psi_k = 1$ at the origin point, because of the functional form used here. However, the marginal utilities vary for the different curves at $x_k > 0$. Specifically, the higher the value of γ_k, the less is the satiation effect in the consumption of x_k. Thus, different values of γ_k lead to different satiation effects, provided $\alpha_k < 1$.

2.1.3 Role of α_k: The express role of α_k is to reduce the marginal utility with increasing consumption of good *k*; that is, it represents a satiation parameter. When $\alpha_k = 1$ for all *k*, this represents the case of absence of satiation effects or, equivalently, the case of constant marginal utility. The utility function in Equation (1) in such a situation collapses to $\sum_k \psi_k x_k$, which represents the perfect substitutes case. This is the case of single discreteness. As α_k moves downward from the value of 1, the satiation effect for good *k* increases. When $\alpha_k \to 0$, the utility function collapses to the LES form, as discussed earlier. α_k can also take negative values and, when $\alpha_k \to -\infty$, this implies immediate and full satiation. Figure 3 plots the utility function for alternative *k* for γ_k = 1 and ψ_k = 1, and for different values of α_k. Again, all of the curves have the same slope ψ_k = 1 at the origin point, and accommodate different levels of satiation through different values of α_k for any given γ_k value.

2.2 Empirical Identification Issues Associated with Utility Form

The discussion in the previous section indicates that ψ_k reflects the baseline marginal utility, which controls whether or not a good is selected for positive consumption (or the extensive margin of choice). The role of γ_k is to enable corner solutions, though it also governs the level of satiation. The purpose of α_k is solely to allow satiation. The precise functional mechanism through which γ_k and α_k impact satiation are, however, different; γ_k controls satiation by translating consumption quantity, while α_k controls satiation by exponentiating consumption quantity. Clearly, both these effects operate in different ways, and different combinations of their values lead to different satiation profiles. However, empirically speaking, and as discussed in detail in Bhat (2008), it is very difficult to disentangle the two effects separately, which leads to serious empirical identification problems and estimation breakdowns when one attempts to estimate both γ_k and α_k parameters for each good. In fact, for a given ψ_k value, it is possible to closely approximate a sub-utility function profile based on a combination of γ_k and α_k values with a sub-utility function based solely on γ_k or α_k values. In actual application, it would behoove the analyst to estimate models based on both the α_k-profile (*i.e.,* a utility function based solely on α_k values) and the γ_k-profile (*i.e.,* a sub-utility function with based solely on values γ_k, with the α_k values set to zero), and choose a specification that provides a better statistical fit. Alternatively, the analyst can stick with one functional form *a priori*, but experiment with various fixed values of α_k for the γ_k-profile and γ_k for the α_k-profile.

2.3 Utility Form for Situations with an Outside Good

Thus far, the discussion has assumed that there is no outside numeraire good (*i.e.,* no essential Hicksian composite good). If an outside good is present, label it as the first good which now has a unit price of one. Then, the utility functional form needs to be modified as follows:

$$U(\boldsymbol{x}) = \frac{1}{\alpha_1}\psi_1 \left\{ (x_1+\gamma_1)^{\alpha_1} \right\} + \sum_{k=2}^{K} \frac{\gamma_k}{\alpha_k}\psi_k \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\} \tag{6}$$

In the above formula, we need $\gamma_1 \le 0$, while $\gamma_k > 0$ for k > 1. Also, we need $x_1 + \gamma_1 > 0$. The magnitude of γ_1 may be interpreted as the required lower bound (or a "subsistence value") for consumption of the outside good.

The identification considerations discussed for the "no-outside good" case carries over to the "with outside good" case. For example, as in the "no-outside good" case, the analyst will generally not be able to estimate both α_k and γ_k for the outside and inside goods. Another important normalization necessary for parameter identification, regardless of the presence or absence of the outside good, is that the coefficients of explanatory variables (including the constants) in the baseline utility parameters ψ_k (k = 1, 2,…, K) should be normalized (for example, to zero) for at least one alternative. In situations with a Hicksian composite outside good, the natural candidate for such normalization is the baseline marginal utility parameter of the outside good. This identification condition is similar to that in the standard discrete choice model, though the origin of the condition is different between standard discrete choice models and the multiple discrete-continuous models. In standard discrete choice models, individuals choose the alternative with the highest indirect utility, so that all that matters is relative utility. In multiple discrete-continuous models, the origin of this condition is the adding up (or budget) constraint associated with the quantity of consumption of each good.

3. ECONOMETRIC STRUCTURE AND KARUSH-KUHN-TUCKER (KKT) CONDITIONS OF OPTIMALITY

The KKT approach employs a direct stochastic specification by assuming the utility function $U(\boldsymbol{x})$ to be random over the population. In all recent applications of the KKT approach for multiple discreteness, a multiplicative random element is introduced to the baseline marginal utility of each good as follows:

$$\psi(z_k, \varepsilon_k) = \psi(z_k) \cdot e^{\varepsilon_k}, \tag{7}$$

where z_k is a set of attributes characterizing alternative k and the decision maker, and ε_k captures idiosyncratic (unobserved) characteristics that impact the baseline utility for good k. The exponential form for the introduction of the random term guarantees the positivity of the baseline utility as long as $\psi(z_k) > 0$. To ensure this latter condition, $\psi(z_k)$ is further parameterized as $\exp(\beta' z_k)$, which then leads to the following form for the baseline random utility associated with good k:

$$\psi(z_k, \varepsilon_k) = \exp(\beta' z_k + \varepsilon_k). \tag{8}$$

The z_k vector in the above equation includes a constant term. The overall random utility function of Equation (3) then takes the following form:

$$U(\boldsymbol{x}) = \sum_k \frac{\gamma_k}{\alpha_k} \left[\exp(\beta' z_k + \varepsilon_k)\right] \cdot \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\} \tag{9}$$

As indicated earlier, the part of β' (*i.e.,* the coefficients of explanatory variables) corresponding to at least one alternative must be normalized to zero.

In the presence of a Hicksian composite outside good, arbitrarily designating the first alternative as the outside good, the overall random utility function can be written as:

$$U(\boldsymbol{x}) = \frac{1}{\alpha_1}\exp(\varepsilon_1)\left\{(x_1+\gamma_1)^{\alpha_1}\right\} + \sum_k \frac{\gamma_k}{\alpha_k}\left[\exp(\beta' z_k + \varepsilon_k)\right]\cdot\left\{\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right\} \qquad (10)$$

Note that, for identification, $\psi(z_1,\varepsilon_1)$ is specified as e^{ε_1}, by normalizing the coefficients of z_1 to zero. But some studies (particularly those in the environmental economics literature) impose a stronger normalization by considering the utility of the outside good as being deterministic (*i.e.,* $\varepsilon_1 = 0$) and setting $\psi(z_1,\varepsilon_1)=1$. Then the overall random utility function becomes:

$$U(\boldsymbol{x}) = \frac{1}{\alpha_1}\left\{(x_1+\gamma_1)^{\alpha_1}\right\} + \sum_k \frac{\gamma_k}{\alpha_k}\left[\exp(\beta' z_k + \varepsilon_k)\right]\cdot\left\{\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right\} \qquad (11)$$

While the above normalization is not theoretically inappropriate, it is unnecessary. Further, it is arbitrary to set a good's utility contribution to be deterministic. This is particularly a problem in situations with no Hicksian composite outside good, where the analyst has to arbitrarily choose the utility contribution of any one alternative to be deterministic. Further, as demonstrated in Bhat (2008), the probability expressions and probability values for the consumption pattern depend on which choice alternative is chosen for this normalization. Finally, in contexts with an outside good, including the stochastic term on the outside good ε_1 helps in capturing correlation among the random utilities of the inside goods. Such correlation helps in inducing greater competition among the consumptions of the inside goods, when compared to the competition between the inside goods and the outside good. Thus, we prefer the specification with stochasticity in the utility contribution of all choice alternatives, including that of the outside good in situations with an outside good.

3.1 Optimal Consumptions

The analyst can solve for the optimal expenditure allocations by forming the Lagrangian and applying the Karush-Kuhn-Tucker (KKT) conditions. For the utility form in Equation (10), the Lagrangian function for the problem is:[3]

$$L = \frac{1}{\alpha_1}\exp(\varepsilon_1)\left\{(x_1+\gamma_1)^{\alpha_1}\right\} + \sum_{k=2}^{K} \frac{\gamma_k}{\alpha_k}\left[\exp(\beta' z_k + \varepsilon_k)\right]\left\{\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right\} - \lambda\left[\sum_{k=1}^{K} p_k x_k - E\right], \qquad (12)$$

where λ is the Lagrangian multiplier associated with the budget constraint (that is, it can be viewed as the marginal utility of total expenditure or income). The KKT first-order conditions for the optimal consumptions (the x_k^* values) are given by:

$$\frac{\exp(\varepsilon_1)}{p_1}\left(x_1^*+\gamma_1\right)^{\alpha_1-1} = \lambda \text{, since } x_1^* > 0$$

[3] Note that the subsequent discourse is for the case with a Hicksian composite outside good that is essential. However, the derivations carry over to the case without an outside good in a straightforward manner.

$$\frac{\exp(\beta' z_k + \varepsilon_k)}{p_k} \left(\frac{x_k^*}{\gamma_k} + 1\right)^{\alpha_k - 1} = \lambda \text{, if } x_k^* > 0 \text{, } k = 2, \ldots, K \tag{13}$$

$$\frac{\exp(\beta' z_k + \varepsilon_k)}{p_k} \left(\frac{x_k^*}{\gamma_k} + 1\right)^{\alpha_k - 1} < \lambda \text{, if } x_k^* = 0 \text{, } k = 2, \ldots, K$$

In the above KKT conditions, the first condition is for the outside good, while the next two sets of conditions are for the inside goods (k = 2, 3,…, K). Note that the price of the Hicksian outside numeraire good p_1 is unity.

The optimal demand satisfies the conditions in Equation (13) plus the budget constraint $\sum_k p_k x_k^* = E$. Substituting for the expression of λ from the KKT condition for the outside good into the KKT conditions for the inside goods, and taking logarithms, one can rewrite the KKT conditions as:

$$\begin{aligned} &V_k + \varepsilon_k = V_1 + \varepsilon_1 \text{ if } x_k^* > 0 \;\; (k = 2, 3, \ldots, K) \\ &V_k + \varepsilon_k < V_1 + \varepsilon_1 \text{ if } x_k^* = 0 \;\; (k = 2, 3, \ldots, K), \end{aligned} \tag{14}$$

where

$V_1 = (\alpha_1 - 1)\ln\left(x_1^* + \gamma_1\right) - \ln p_1$, and

$V_k = \beta' z_k + (\alpha_k - 1)\ln\left(\frac{x_k^*}{\gamma_k} + 1\right) - \ln p_k$ (k = 2, 3,…, K).

3.2 General Econometric Model Structure and Identification

To complete the model structure, the analyst needs to specify the error structure. In the general case, let the joint probability density function of the ε_k terms be $f(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_K)$. Then, the probability that the individual consumes the first M of the K goods is:

$$\begin{aligned} P(x_1^*, x_2^*, x_3^*, \ldots, x_M^*, 0, 0, \ldots, 0) = &\; |J| \int_{\varepsilon_1=-\infty}^{+\infty} \int_{\varepsilon_{M+1}=-\infty}^{V_1 - V_{M+1} + \varepsilon_1} \int_{\varepsilon_{M+2}=-\infty}^{V_1 - V_{M+2} + \varepsilon_1} \cdots \int_{\varepsilon_{K-1}=-\infty}^{V_1 - V_{K-1} + \varepsilon_1} \int_{\varepsilon_K=-\infty}^{V_1 - V_K + \varepsilon_1} \\ &f(\varepsilon_1, V_1 - V_2 + \varepsilon_1, V_1 - V_3 + \varepsilon_1, \ldots, V_1 - V_M + \varepsilon_1, \varepsilon_{M+1}, \varepsilon_{M+2}, \ldots, \varepsilon_{K-1}, \varepsilon_K) \\ &d\varepsilon_K d\varepsilon_{K-1} \ldots d\varepsilon_{M+2} d\varepsilon_{M+1} d\varepsilon_1 \;, \end{aligned} \tag{15}$$

where J is the Jacobian whose elements are given by (see Bhat, 2005):

$$J_{ih} = \frac{\partial[V_1 - V_{i+1} + \varepsilon_1]}{\partial x_{h+1}^*} = \frac{\partial[V_1 - V_{i+1}]}{\partial x_{h+1}^*};\; i, h = 1, 2, \ldots, M - 1. \tag{16}$$

The probability expression in Equation (15) is a (K-M+1)-dimensional integral. The dimensionality of the integral can be reduced by one by noticing that the KKT conditions can also be written in a differenced form. To do so, define $\tilde{\varepsilon}_{k1} = \varepsilon_k - \varepsilon_1$, and let the implied multivariate distribution of the error differences be $g(\tilde{\varepsilon}_{21}, \tilde{\varepsilon}_{31}, \ldots, \tilde{\varepsilon}_{K1})$. Then, Equation (11) may be written in the equivalent (K-M)-integral form shown below:

$$P(x_1^*,\ x_2^*,\ x_3^*,\ ...,\ x_M^*,\ 0,\ 0,\ ...,\ 0) = \ |J| \int_{\tilde{\varepsilon}_{M+1,1}=-\infty}^{V_1-V_{M+1}} \int_{\tilde{\varepsilon}_{M+2,1}=-\infty}^{V_1-V_{M+2}} \cdots \int_{\tilde{\varepsilon}_{K-1,1}=-\infty}^{V_1-V_{K-1}} \int_{\tilde{\varepsilon}_{K,1}=-\infty}^{V_1-V_K} \tag{17}$$

$$g(V_1-V_2,\ V_1-V_3,\ ...,\ V_1-V_M,\ \tilde{\varepsilon}_{M+1,1},\ \tilde{\varepsilon}_{M+2,1},\ ...,\ \tilde{\varepsilon}_{K,1})d\tilde{\varepsilon}_{K,1}d\tilde{\varepsilon}_{K-1,1}...d\tilde{\varepsilon}_{M+1,1}$$

The equation above indicates that the probability expression for the observed optimal consumption pattern of goods is completely characterized by the (K-1) error terms in the differenced form. Thus, all that is estimable is the (K-1)×(K-1) covariance matrix of the error differences. In other words, it is not possible to estimate a full covariance matrix for the original error terms $(\varepsilon_1, \varepsilon_2, ..., \varepsilon_K)$ because there are infinite possible densities for f(.) that can map into the same g(.) density for the error differences (see Train, 2003, page 27, for a similar situation in the context of standard discrete choice models). There are many possible ways to normalize f(.) to account for this situation. For example, one can assume an identity covariance matrix for f(.), which automatically accommodates the normalization that is needed. Alternatively, one can estimate g(.) without reference to f(.).

In the general case when the unit prices p_k vary across goods, it is possible to estimate $K*(K-1)/2$ parameters of the full covariance matrix of the error differences, as just discussed (though the analyst might want to impose constraints on this full covariance matrix for ease in interpretation and stability in estimation). However, when the unit prices are not different among the goods, an additional scaling restriction needs to be imposed. A typical way to do is by normalizing the scale of the random error terms (*i.e.,* the scale of the ε_k terms) to one.

4. SPECIFIC MODEL STRUCTURES

4.1 The Multiple Discrete-Continuous Extreme-Value (MDCEV) Model

Following Bhat (2005, 2008), consider an extreme value distribution for ε_k and assume that ε_k is independent of z_k (k = 1, 2,…, K). The ε_k 's are also assumed to be independently distributed across alternatives with a scale parameter of σ (σ can be normalized to one if there is no variation in unit prices across goods). Let V_k be defined as follows:

$$V_1 = (\alpha_1 - 1)\ln\left(x_1^* + \gamma_1\right) - \ln p_1$$

$$V_k = \beta' z_k + (\alpha_k - 1)\ln\left(x_k^* + 1\right) - \ln p_k \ \ (k = 2, 3, ..., K), \text{ when the } \alpha\text{-profile is used, and} \tag{18}$$

$$V_k = \beta' z_k - \ln\left(\frac{x_k^*}{\gamma_k} + 1\right) - \ln p_k \ \ (k = 2, 3, ..., K), \text{ when the } \gamma\text{-profile is used.}$$

As discussed earlier, it is generally not possible to estimate the V_k form in Equation (14), because the α_k terms and γ_k terms serve a similar satiation role.

From Equation (17), the probability that the individual allocates expenditure to the first first M of the K goods ($M \geq 1$) with a corresponding consumption vector $x^* = (x_1^*, x_2^*, x_3^*, ..., x_M^*, 0, 0, ..., 0)$ is:

$$P\left(x_1^*, x_2^*, x_3^*, \ ..., \ x_M^*, \ 0, \ 0, \ ..., \ 0\right)$$

$$= \ |J| \int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty} \left\{ \left(\prod_{i=2}^{M} \frac{1}{\sigma} \lambda \left[\frac{V_1 - V_i + \varepsilon_1}{\sigma} \right] \right) \right\} \times \left\{ \prod_{s=M+1}^{K} \Lambda \left[\frac{V_1 - V_s + \varepsilon_1}{\sigma} \right] \right\} \frac{1}{\sigma} \lambda \left(\frac{\varepsilon_1}{\sigma} \right) d\varepsilon_1 \ , \tag{19}$$

where λ is the standard extreme value density function, Λ is the standard extreme value cumulative distribution function, and $|J|$ is the determinant of the Jacobian matrix obtained from applying the change of variables calculus between the stochastic KKT conditions and the consumptions, given by the following expression (Bhat, 2008):

$$|J| \ = \frac{1}{p_1} \left(\prod_{i=1}^{M} f_i \right) \left(\sum_{i=1}^{M} \frac{p_i}{f_i} \right), \text{ where } f_i = \left(\frac{1-\alpha_i}{x_i^* + \gamma_i} \right) \tag{20}$$

The integral in Equation (19) collapses to a surprisingly simple closed form expression providing the following overall expression (Bhat, 2008):

$$P\left(x_1^*, x_2^*, x_3^*, \ ..., \ x_M^*, \ 0, \ 0, \ ..., \ 0\right) = \frac{1}{p_1} \cdot \frac{1}{\sigma^{M-1}} \cdot \left(\prod_{i=1}^{M} f_i \right) \left(\sum_{i=1}^{M} \frac{p_i}{f_i} \right) \left[\frac{\prod_{i=1}^{M} e^{V_i/\sigma}}{\left(\sum_{k=1}^{K} e^{V_k/\sigma} \right)^M} \right] (M-1)! \tag{21}$$

The reader will note that the above probability expression can be used even in contexts without an essential Hicksian composite outside good. The only difference in the probability expressions between the two contexts is in how V_1 is defined. Specifically, in situations without an essential Hicksian composite outside good, V_1 is defined in the same fashion as V_k (k = 2, 3,…, K) are defined in Equation (18). Further, the expression in Equation (21) is dependent on the unit price of the good that is used as the first one (see the $1/p_1$ term in front). In particular, different probabilities of the same consumption pattern arise depending on the good that is labeled as the first good (note that any good that is consumed may be designated as the first good).[4] In terms of the likelihood function, the $1/p_1$ term can be ignored, since it is simply a constant in each individual's likelihood function. Thus, the same parameter estimates will result independent of the good designated as the first good for each individual.

In the case when $M = 1$ (*i.e.,* only one alternative is chosen), there are no satiation effects (α_k=1 for all k) and the Jacobian term drops out (that is, the continuous component drops out, because all expenditure is allocated to good 1). Then, the model in Equation (21) collapses to the standard MNL model. Thus, the MDCEV model is a multiple discrete-continuous extension of the standard MNL model.

4.2 Closed form extensions of the Multiple Discrete-Continuous Extreme-Value (MDCEV) Model

Thus far, we have assumed that the ε_k terms are independently and identically extreme value distributed across alternatives k. The analyst can extend the model to allow correlation across

[4] This is not an issue in contexts with a numeraire Hicksian composite outside good because p_1= 1.

alternatives using a generalized extreme value (GEV) error structure. The advantage of the GEV structure is that it results in closed-form probability expressions for any and all consumption patterns.

4.2.1 The MDCNEV Model: Pinjari and Bhat (2010) formulate a special two-level nested case of the MDCGEV model with a nested extreme value distributed error structure that has the following joint cumulative distribution:

$$F(\varepsilon_1, \varepsilon_2, .., \varepsilon_K) = \exp\left[-\sum_{\eth=1}^{S_K}\left\{\sum_{i \in \eth^{\text{th}}\text{nest}} \exp\left(-\varepsilon_i / \theta_{\eth}\right)\right\}^{\theta_{\eth}}\right] \tag{22}$$

In the above expression, $\eth\ (=1,2,...,S_K)$ is the index to represent a nest of alternatives, S_K is the total number of nests the K alternatives belong to, and $\theta_{\eth}(0 < \theta_{\eth} \leq 1; \eth = 1,2,...,S_K)$ is the (dis)similarity parameter introduced to induce correlations among the stochastic components of the utilities of alternatives belonging to the $\eth^{th}$ nest. This error structure assumes that the nests are mutually exclusive and exhaustive (*i.e.,* each alternative can belong to only one nest and all alternatives are allocated to one of the S_K nests).

Without loss of generality, let $1,2,...,S_M$ be the nests the M chosen alternatives belong to, and let $q_1, q_2, ..., q_{S_M}$ be the number of chosen alternatives in each of the S_M nests (thus, $q_1 + q_2 + ... + q_{S_M} = M$). Using the nested extreme value error distribution assumption specified in Equation (22) (and the above-identified notation), Pinjari and Bhat (2010) derived the following expression for the multiple discrete-continuous nested extreme value (MDCNEV) model:

$$P(x_1^*, x_2^*, ... x_M^*,\ 0,...,0) =$$

$$|J| \frac{\prod_{i \in \{\text{chosen alts}\}} e^{\frac{V_i}{\theta_i}}}{\prod_{\eth=1}^{S_M}\left(\sum_{i \in \eth^{\text{th}}\text{nest}} e^{\frac{V_i}{\theta_{\eth}}}\right)^{q_{\eth}}} \sum_{r_1=1}^{q_1} \ldots \sum_{r_{\eth}=1}^{q_{\eth}} \ldots \sum_{r_{S_M}=1}^{q_{S_M}} \left\{ \prod_{\eth=1}^{S_M} \left[\frac{\left(\sum_{i \in \eth^{\text{th}}\text{nest}} e^{\frac{V_i}{\theta_{\eth}}}\right)^{\theta_{\eth}}}{\sum_{\eth=1}^{S_k}\left\{\left(\sum_{i \in \eth^{\text{th}}\text{nest}} e^{\frac{V_i}{\theta_{\eth}}}\right)^{\theta_{\eth}}\right\}} \right]^{q_{\eth} - r_{\eth} + 1} \left(\prod_{\eth=1}^{S_M} sum(X_{r\eth})\right)\left(\sum_{\eth=1}^{S_M}(q_{\eth} - r_{\eth} + 1) - 1\right)! \right\} \tag{23}$$

In the above expression, $sum(X_{r\eth})$ is the sum of elements of a row matrix $X_{r\eth}$ (see Appendix A of Pinjari and Bhat, 2010 for a description of the form of the matrix $X_{r\eth}$).

4.2.2 The MDCGEV Model: More recently, Pinjari (2011) formally proved that the existence of, and derived, the closed form probability expressions for MDC models with error structure based on McFadden's (1978) GEV structure. To do so, he expressed the probability expression in Equation (15) as an integral of an M^{th} order partial derivative of the K-dimensional joint cumulative distribution function (CDF) of the error terms $(\varepsilon_1, \varepsilon_2, .., \varepsilon_K)$:

$$P(x_1^*,...,x_M^*,0,...,0) = |J| \int_{\varepsilon_1=-\infty}^{+\infty} \left[\left\{ \frac{\partial^M}{\partial\varepsilon_1 ... \partial\varepsilon_M} F(\varepsilon_1, \varepsilon_2, ..., \varepsilon_K) \right\} \Bigg|_{\varepsilon_i = V_1 - V_i + \varepsilon_1, \forall i=1,2,...,K} \right] d\varepsilon_1 \qquad (24)$$

where $F(\varepsilon_1, \varepsilon_2, ..., \varepsilon_K)$ is the joint CDF of the error terms $(\varepsilon_1, \varepsilon_2, .., \varepsilon_K)$ specified based on McFadden's (1978) GEV form as below:

$$F_{GEV}(\varepsilon_1, \varepsilon_2, .., \varepsilon_K) = \exp\left[-G\left(e^{-\varepsilon_1}, e^{-\varepsilon_2}, ..., e^{-\varepsilon_K} \right) \right] \qquad (25)$$

where G is a non-negative function with the following properties:

1. $G\left(y_1,...y_i...,y_K\right) \geq 0, \quad \forall\, y_i > 0 \ (i = 1,2,...,K)$
2. G is homogeneous of degree $\mu > 0$, that is $G\left(ay_1,...ay_i...,ay_K\right) = a^{\mu} G(y_1,...y_i...,y_K)$,
3. $\lim_{y_i \to +\infty} G\left(y_1,...y_i...,y_K\right) = +\infty, \forall i = 1,2,...,K$, and
4. $\left(-1\right)^M \dfrac{\partial^M G(y_1,...y_K)}{\partial y_1 ... \partial y_M} \leq 0, \quad \forall\, y_i > 0 \ (i = 1,2,...,K)$.

He then derived a general, closed form for the probability expressions as below:

$$P(x_1^*,...,x_M^*,0,...,0) = |J| \prod_{i=1}^{M} e^{V_i} \times \left[\begin{array}{l} \frac{(M-1)!}{H^M} \left\{ \pm(H_1 H_2 .. H_M) \right\} + \\ \frac{(M-2)!}{H^{M-1}} \left\{ \pm(H_{12}^2 H_3 .. H_M) \pm (H_1 H_{23}^2 .. H_M) \pm ... \pm (H_1 H_2 .. H_{(M-1)M}^2) \right\} + \\ \frac{(M-3)!}{H^{M-2}} \left\{ \pm(H_{123}^3 H_4 .. H_M) \pm ... \pm (H_{12}^2 H_{34}^2 .. H_M) \pm ... \right\} + \\ ... \\ \frac{1!}{H^2} \left\{ \pm(H_{123...M-1}^{M-1} H_M) \pm ... \pm (H_1 H_{234...M}^{M-1}) \right\} + \\ \frac{0!}{H} \left\{ \pm(H_{123...M}^{M}) \right\} \end{array} \right] \qquad (26)$$

where $H_i = \dfrac{\partial H(e^{V_1},..,e^{V_K})}{\partial e^{V_i}}$, $H_{123...n}^n = \dfrac{\partial^n H(e^{V_1},..,e^{V_K})}{\partial e^{V_1} .. \partial e^{V_n}}$, and all other terms are defined similarly[5].

Recognizing that working with the above general form of probability expressions becomes difficult in situations with complex covariance structures and a large set of choice alternatives (because of the sheer number of terms in the expression), Pinjari (2011) derived

[5] G and H are similar functions, but with different arguments; G represents $G(e^{-\varepsilon_1},...,e^{-\varepsilon_n})$, whereas H represents $G(e^{V_1},..e^{V_i}..,e^{V_n})$. Note from the $\pm$ signs that the sign in front of each mixed partial derivative term depends on the number of partial derivatives in the term and the number of chosen alternatives M. Also note that the model structures for MDCNEV and MDCGEV are derived for the case without price variation across choice alternatives. One can extend these structures for situations with price variation in a straightforward fashion.

compact probability expressions for a variety of cross-nested error structures. The reader is referred to this paper for further details.

4.3 The Mixed MDCEV Model

The MDCGEV structure is able to accommodate flexible correlation patterns. However, it is unable to accommodate random taste variation, and it imposes the restriction of equal scale of the error terms. Incorporating a more general error structure is straightforward through the use of a mixing distribution, which leads to the Mixed MDCEV (or MMDCEV) model. Specifically, the error term, ε_k, may be partitioned into two components, ζ_k and η_k. The first component, ζ_k, can be assumed to be independently and identically Gumbel distributed across alternatives with a scale parameter of σ. The second component, η_k, can be allowed to be correlated across alternatives and to have a heteroscedastic scale. Let $\eta = (\eta_1, \eta_2, ..., \eta_K)'$, and assume that η is distributed multivariate normal, $\eta \sim N(0, \Omega)$.

For given values of the vector η, one can follow the discussion of the earlier section and obtain the usual MDCEV probability that the first M of the k goods are consumed. The unconditional probability can then be computed as:

$$P\left(x_1^*, x_2^*, ..., x_M^*, 0, ..., 0\right) = \int_{\eta} \frac{1}{\sigma^{M-1}} |J| \left[\frac{\prod_{i=1}^{M} e^{(V_i + \eta_i)/\sigma}}{\left(\sum_{k=1}^{K} e^{(V_k + \eta_k)/\sigma} \right)^M} \right] (M-1)! \; dF(\eta). \qquad (27)$$

where F is the multivariate cumulative normal distribution.

Other distributions may also be used for η. Note that the distribution of η can arise from an error components structure or a random coefficients structure or a combination of the two, similar to the case of the usual mixed logit model. Thus, the model in Equation (27) can be extended in a conceptually straightforward manner to also include random coefficients on the independent variables z_k, and random-effects (or even random coefficients) in the α_k satiation parameters (if the α profile is used) or the γ_k parameters (if the γ profile is used).

4.4 The Multiple Discrete-Continuous Probit (MDCP) Model

The choice of extreme value (either EV or GEV) stochastic specification is driven by convenience (of analytical tractability) rather than theory. A multivariate normally (MVN) distributed stochasticity assumption leads to complex likelihood functions, one reason why the KKT approach did not gain traction for empirical analysis until recently. Attempts have been made to address this issue by using simulation methods such as the GHK simulator (see Kim *et al*., 2002) and Bayesian estimation methods. However, the GHK and other such simulators become computationally impractical as the dimensionality of integration increases with the number of alternatives. Bayesian estimation methods can also be computationally intensive and saddled with convergence-determination issues. Thus, no study has been able to estimate KKT demand systems with multivariate normal (MVN) distributions beyond a small number of alternatives.

Notwithstanding the estimation difficulties, there are notable advantages of using an MVN error distribution. First, the MVN error kernel makes it easy to incorporate general covariance structures as well as random coefficients, as long as the number of choice alternatives is not too large. Second, an appealing feature of MVN errors is the possibility of negative correlations among the utilities of different alternatives (as opposed to MEV errors, which allow only positive dependency). This can potentially be exploited to capture situations where the choice of one alternative may reduce (if not preclude) the likelihood of choosing another, where the pattern of substitution is fundamentally different from the substitution due to satiation effects. Given these advantages, we show below that the probability expression of the MDCP model involves the evaluation of a multivariate normal cumulative distribution function (MVNCDF).

Equation (17) provides the general expression for consumption probabilities for an MDC model based on KKT conditions of random utility maximization. One can rewrite the probability expression using a *differenced-errors* form as below:

$$P(x_1^*, x_2^*, ..., x_M^*, 0, ..., 0) = |J| \times P\begin{bmatrix} \left((\tilde{\varepsilon}_{2,1} = V_1 - V_2), (\tilde{\varepsilon}_{3,1} = V_1 - V_3), ..., (\tilde{\varepsilon}_{M,1} = V_1 - V_M)\right), \\ \left((\tilde{\varepsilon}_{M+1,1} < V_1 - V_{M+1}),\ (\tilde{\varepsilon}_{M+2,1} < V_1 - V_{M+2}),\ ...,\ (\tilde{\varepsilon}_{K,1} < V_1 - V_K)\right) \end{bmatrix} \quad (28)$$

In the above expression, $\left(\tilde{\varepsilon}_{2,1}, \tilde{\varepsilon}_{3,1}, ..., \tilde{\varepsilon}_{M,1}, \tilde{\varepsilon}_{M+1,1}, ..., \tilde{\varepsilon}_{K,1}\right)$ is a K-1 dimensional vector of error differences following a multivariate normal distribution with a zero mean vector $\boldsymbol{\mu}$ (all elements in $\boldsymbol{\mu}$ are zeros), and a variance-covariance matrix Σ. For later use, partition this K-1 dimensional vector into two smaller vectors $\boldsymbol{A}$ and $\boldsymbol{B}$, where $\boldsymbol{A} = \{\tilde{\varepsilon}_{2,1}, \tilde{\varepsilon}_{3,1}, ..., \tilde{\varepsilon}_{M,1}\}$ and $\boldsymbol{B} = \{\tilde{\varepsilon}_{M+1,1}, \tilde{\varepsilon}_{M+2,1}, ..., \tilde{\varepsilon}_{K,1}\}$. Thus, the $\boldsymbol{\mu}$ and Σ matrices can also be partitioned as: $\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu_1} \\ \boldsymbol{\mu_2} \end{bmatrix}$, and $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$. In the partition of Σ, Σ_{11} and Σ_{22} are the variance-covariance matrices of $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively, while Σ_{12} and Σ_{21} contain the covariance terms between the elements in $\boldsymbol{A}$ and those in $\boldsymbol{B}$.

Now, express the MDCP probability expression in Equation (28) as:

$$P(x_1^*, x_2^*, ..., x_M^*, 0, ..., 0) = |J| \times \mathrm{P}\left[\boldsymbol{A} = \mathbf{a}, \boldsymbol{B} < \mathbf{b}\right] \quad (29)$$

Where, $\mathbf{a} = \{V_1 - V_2, V_1 - V_3, ..., V_1 - V_M\}$ and $\mathbf{b} = \{V_1 - V_{M+1}, V_1 - V_{M+2}, ..., V_1 - V_K\}$.

One can express the above expression as a product of marginal and conditional probabilities:

$$P(x_1^*, x_2^*, ..., x_M^*, 0, ..., 0) = |J|.\mathrm{P}\left(\boldsymbol{A} = \mathbf{a}\right).\mathrm{P}\left(\boldsymbol{B} < \mathbf{b} \mid \boldsymbol{A} = \mathbf{a}\right) \quad (30)$$

To simplify the conditional probability expression in the above expression, we utilize a property of multivariate normal (MVN) distribution that the distribution of **B** conditional on **A** = **a**, is another MVN distribution as given below (Tong, 1990. pp 35):

$$(\boldsymbol{B} \mid \boldsymbol{A} = \mathbf{a}) \sim N\left(\bar{\boldsymbol{\mu}}, \bar{\Sigma}\right), \text{ where } \bar{\boldsymbol{\mu}} = \boldsymbol{\mu}_2 + \Sigma_{21}\Sigma_{11}^{-1}\left(\mathbf{a} - \boldsymbol{\mu}_1\right), \text{ and } \bar{\Sigma} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \quad (31)$$

In the above expression, since $\boldsymbol{\mu_1}$ and $\boldsymbol{\mu_2}$ are zero-vectors, one can write $\bar{\boldsymbol{\mu}} = \Sigma_{21}\Sigma_{11}^{-1}\mathbf{a}$.

Using the above result, the conditional probability expression in Equation (30) can be expressed as $\Pr(\boldsymbol{B} < \mathbf{b} \mid \boldsymbol{A} = \mathbf{a}) = \Pr(\boldsymbol{C} < \mathbf{b})$ where $\boldsymbol{C}$ is an MVN distribution as described above. Then, the MDCP consumption probability can be expressed as:

$$P(x_1^*, x_2^*, ..., x_M^*, 0, ..., 0) = |J| . \mathrm{P}\left(\boldsymbol{A} = \mathbf{a}\right) . \mathrm{P}\left(\boldsymbol{C} < \boldsymbol{c}\right) \tag{32}$$

In the above joint probability expression, the marginal probability $\mathrm{P}\left(\boldsymbol{A} = \mathbf{a}\right)$ is a multivariate normal probability distribution function (pdf) with a simple closed form expression, where as the MVNCDF $\mathrm{P}\left(\boldsymbol{C} < c\right)$ does not have a closed form.

Next, write the MVNCDF $\Pr(\boldsymbol{C} < \mathbf{b})$ in standardized form as below:

$$\begin{aligned} \Pr(\boldsymbol{C} < \mathbf{b}) &= \mathrm{P}\left(\frac{\boldsymbol{C} - \bar{\boldsymbol{\mu}}}{\boldsymbol{\sigma}} < \frac{\mathbf{b} - \bar{\boldsymbol{\mu}}}{\boldsymbol{\sigma}}\right) \\ &= \mathrm{P}\left(W_1 < w_1,\ W_2 < w_2, ...,\ W_{K-M} < w_{K-M}\right) \end{aligned} \tag{33}$$

where, $\left(W_1, W_2, ..., W_{K-M}\right)$ is a vector of standardized, normally distributed random variables in $\frac{\boldsymbol{C} - \bar{\boldsymbol{\mu}}}{\boldsymbol{\sigma}}$ and $\left(w_1, w_2, ..., w_{K-M}\right)$ is a vector of scalars in $\frac{\mathbf{b} - \bar{\boldsymbol{\mu}}}{\boldsymbol{\sigma}}$. Similarly, $\bar{\boldsymbol{\mu}} = (\mu_1, ..., \mu_i, ..., \mu_{K-M})$ is a vector of means and $\boldsymbol{\sigma} = (\sigma_1, ..., \sigma_i, ..., \sigma_{K-M})$ is a vector of standard deviations[6] of the normally distributed random variables in $\boldsymbol{C}$.

The problem now boils down to approximating the MVNCDF in Equation (33). In the recent past, there has been some evidence that using analytical approximations (as opposed to simulation) for evaluating the MVN cumulative distribution function can help in easier estimation of single discrete choice models (e.g., the multinomial probit model; see Bhat and Sidharthan, 2011). Bhat *et al.* (2012) show that such analytical approximation methods can help in the estimation of MDCP models as well (*i.e.,* MDC models with MVN errors). The performance of different analytical approximation methods to evaluate the MVNCDF to estimate the parameters of the MDCP models is an open avenue for further research.

5. THE JOINT MDCEV-SINGLE DISCRETE CHOICE MODEL

The MDCEV model and its extensions discussed thus far are suited for the case when the alternatives are imperfect substitutes, as recognized by the use of a non-linear utility that accommodates a diminishing marginal utility as the consumption of any alternative increases. However, there are many instances where the choice situation is characterized by a combination of imperfect and perfect substitutes in the choice alternative set. The MDCEV model needs to be modified to handle such a combination of a multiple discrete-continuous choice among the imperfect substitutes, as well as a single choice of one alternative each from each subset of perfect substitutes. We do not discuss this case here due to space constraints, but the reader is referred to Bhat *et al*. (2009) and Bhat *et al*. (2006) for such formulations. Both these studies by Bhat and co-authors assume the absence of price variation across the perfect substitutes. Formulation of KKT model systems to consider price variation across imperfect substitutes as well as perfect substitutes is a potentially fruitful avenue for further research.

---

[6] σ_i is the square root of the ii^{th} element of the covariance matrix $\bar{\Sigma}$.

6. PREDICTION AND WELFARE ANALYSIS

Thanks to the above advances, several empirical applications have appeared in the recent literature using the KKT approach to model MDC choices. These applications cover a wide range of empirical contexts, including individuals' time-use analysis, household expenditure patterns, household vehicle ownership and usage, household energy consumption, recreational demand choices, and valuation of a variety of environmental goods (e.g., fish stock, air quality, water quality). One reason why the KKT approach did not gain much attention until the recent decade was the difficulty of estimating the model parameters. But we are now able to easily estimate KKT demand systems with a large number of choice alternatives (see Van Nostrand *et al*., 2013 for a model with 211 choice alternatives). Another reason why the KKT approach has not gained popularity is the lack of simple methods to *apply* the models for forecasting and policy analysis purposes. This section reviews the recent advances aimed to fill that gap.

Once the model parameters are estimated, prediction exercises or welfare analyses with KKT-based MDC models involve solving the constrained, non-linear random utility maximization problem in Equation (1) (or its dual form) for each consumer. In the presence of corner solutions (*i.e.,* multiple discreteness), there is no straight-forward analytic solution to this problem. The typical approach is to adopt a constrained non-linear optimization procedure at each of several simulated values drawn from the distribution of the stochastic error terms (*i.e.,* the ε_k terms). The constrained optimization procedure itself has been based on either enumerative or iterative techniques. The enumerative technique (used by Phaneuf *et al*., 2000) involves an enumeration of all possible sets of alternatives that the consumer can potentially choose. This brute-force method becomes computationally impractical as the number of choice alternatives increases.

von Haefen *et al*. (2004) proposed a numerical bisection algorithm based on the insight that, with additively separable utility functions, the optimal consumptions of all goods can be derived if the optimal consumption of the outside good is known. Specifically, conditional on unobserved heterogeneity, they iteratively solve for the optimal consumption of the outside good (and that of other goods) using a bisection procedure. They begin their iterations by setting the lower bound for the consumption of the outside good to zero and the upper bound to be equal to the budget. The average of the lower and upper bounds is used to obtain the initial estimate of the outside good consumption. Based on this, the amounts of consumption of all other inside goods are computed using the KKT conditions. Next, a new estimate of consumption of the outside good is obtained by subtracting the budget on the consumption of the inside goods from the total budget available. If this new estimate of the outside good is larger (smaller) than the earlier estimate, the earlier estimate becomes the new lower (upper) bound of consumption for the outside good, and the iterations continue until the difference between the lower and upper bounds is within an arbitrarily designated threshold. To circumvent the need to perform predictions over the entire distribution of unobserved heterogeneity (which can be time-consuming), von Haefen *et al.* condition on the observed choices.

In a recent paper, Pinjari and Bhat (2011) undertook analytic explorations with the KKT conditions of optimality that shed new light on the properties of Bhat's MDCEV model with additive utility functions. Specifically, they derive a property that the price-normalized baseline marginal utility (*i.e.,* ψ_k / p_k) of a chosen alternative must be greater than the price-marginalized baseline marginal utility of an alternative that is not chosen. Further, they discuss a fundamental property of several KKT demand model systems in the literature with additively separable utility form and a single linear binding constraint. Specifically, the choice alternatives can always be arranged in the descending order of a specific measure that depends on the functional form of the utility function. Consequently, when all the choice alternatives are arranged in the descending order of their baseline marginal utility, and the number of chosen alternatives (*M*) is known, it is a trivial task to identify the chosen alternatives as the first *M* alternatives in the arrangement. Based on this insight, Pinjari and Bhat (2011) propose computationally efficient prediction algorithms for different forms of the utility function in Equation (3). One such forecasting algorithm, for the utility form with equal α_k parameters across all choice alternatives (*i.e.,* $\alpha_k = \alpha \ \forall k = 1, 2, ..., K$) for choice situations with an outside good is outlined in four broad steps below. For predictions algorithms for other additively separable utility forms, the reader is referred to Pinjari and Bhat (2011).

Step 0: Assume that only the outside good is chosen and let the number of chosen goods $M = 1$.

Step 1: Given the input data (z_k, p_k), model parameters (β, γ_k, α), and the simulated error term (ε_k) draws, compute the price-normalized baseline utility values (ψ_k / p_k) for all alternatives. Arrange all the *K* alternatives available to the consumer in the descending order of the (ψ_k / p_k) values (with the outside good in the first place).

Step 2: Compute the value of λ using the following equation. Go to step 3.

$$\lambda = \left(\frac{E + \sum_{k=2}^{M} p_k \gamma_k}{p_1 (\psi_1 / p_1)^{\frac{1}{1-\alpha}} + \sum_{k=2}^{M} p_k \gamma_k (\psi_k / p_k)^{\frac{1}{1-\alpha}}} \right)^{\alpha - 1} \tag{34}$$

Step 3: If $\lambda > (\psi_{M+1} / p_{M+1})$ (this condition represents the KKT condition for the *M+1*th alternative)

Compute the optimal consumptions of the first *M* alternatives in the above descending order using the following expressions. Set the consumptions of other alternatives as zero and stop.

$$x_1^* = \frac{(\psi_1 / p_1)^{\frac{1}{1-\alpha}} \left(E + \sum_{k=2}^{M} p_k \gamma_k \right)}{p_1 (\psi_1 / p_1)^{\frac{1}{1-\alpha}} + \sum_{k=2}^{M} p_k \gamma_k (\psi_k / p_k)^{\frac{1}{1-\alpha}}} \tag{35}$$

$$x_k^* = \gamma_k . \left(\frac{(\psi_k / p_k)^{\frac{1}{1-\alpha}} . \left(E + \sum_{k=2}^{M} p_k \gamma_k \right)}{p_1 (\psi_1 / p_1)^{\frac{1}{1-\alpha}} + \sum_{k=2}^{M} p_k \gamma_k (\psi_k / p_k)^{\frac{1}{1-\alpha}}} - 1 \right); \ \forall k = (2, 3, ..., M) \tag{36}$$

Else, (if $\lambda \le (\psi_{M+1} / p_{M+1})$, set $M = M+1$ and go to step 4.

Step 4: If ($M = K$), Compute the optimal consumptions using Equations (35) and (36) and stop.

Else, (if $M < K$), go to step 2.

The algorithm outlined above can be applied a large number of times with different simulated values of the ε_k terms to sufficiently cover the simulated distribution of unobserved heterogeneity (*i.e.,* the ε_k terms) and obtain the distributions of the consumption forecasts.

The above discussion is primarily oriented toward using KKT-based MDC models for prediction, but does not extend the discussion to include welfare analysis. For a discussion of how such prediction algorithms can be used for welfare analysis, see von Haefen and Phaneuf (2005).

7. FUTURE DIRECTIONS

In the recent past, there has been an increasing recognition of the need to extend the basic formulation of consumer's utility maximization in Equation (1) in the following directions:

(1) Flexible functional forms for the utility specification,

(2) Flexible stochastic specifications for the utility functions,

(3) Flexibility in the specification of constraints faced by the consumer,

Each of these directions is discussed next.

7.1 Flexible, Non-additive Utility Forms

Most KKT models in the literature assume that the direct utility contribution due to the consumption of different alternatives is additively separable. Mathematically, this assumption implies that: $U(x_1,\ldots,x_K) = U_1(x_1) +\ldots+ U_K(x_K)$, and greatly simplifies the task of model estimation and welfare analysis. However, this assumption imposes strong restrictions on preference structures and consumption patterns. First, the marginal utility of one alternative is independent of the consumption of another alternative. This assumption, with an increasing and quasi-concave utility function, implies that goods can be neither inferior nor complementary; they can only be substitutes. Thus, for example, one cannot model a situation where the consumption of one good (*e.g.*, a new car) may increase the consumption of other goods (e.g., gasoline). Third, even flexible substitution patterns in the consumption of different goods can be achieved only by correlating the stochastic utility components of different goods, but not through an explicit functional form. To overcome the restrictions identified above, it is critical to develop tractable estimation methods with flexible, non-additively separable utility functions.

There have been a handful of recent efforts in this direction. For example, building on Bhat's additively separable linear Box-Cox utility form, Vasquez-Lavin and Hanemann (2009) presented a general utility form with interaction terms between sub-utilities, as below:

$$U(\boldsymbol{x}) = \sum_{k=1}^{K} \psi_k \frac{\gamma_k}{\alpha_k}\left\{\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right\} + \frac{1}{2}\sum_{k=1}^{K}\sum_{m=1}^{K}\left\{\theta_{km}\frac{\gamma_k}{\alpha_k}\left[\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right]\frac{\gamma_m}{\alpha_m}\left[\left(\frac{x_m}{\gamma_m}+1\right)^{\alpha_m} - 1\right]\right\} \quad (37)$$

In the above expression, the second term induces interactions between pairs of goods (m,k) and includes quadratic terms (when $m = k$). These interaction terms allow the marginal utility of a good (k) to depend on the consumption of other goods (m). Specifically, a positive (negative) value for θ_{mk} implies that m and k are complements (substitutes). However, the quadratic nature of the utility form does not maintain global consistency (over all consumption bundles) of the strictly increasing and quasi-concave property. Specifically, for certain parameter values and consumption patterns, the utility accrued can *decrease* with increasing consumption, or the marginal utility can *increase* with increasing consumption, which is theoretically inconsistent. Bhat and Pinjari (2010) show how a simple normalization by setting $\theta_{km} = 0$ when $m = k$ in Equation (37) can resolve the issues of theoretical (in)consistency and parameter (un)identification. Other efforts on accommodating complementarity in consumption include Lee *et al*. (2010) who propose simpler interaction terms using log(quantities), and Gentzkow (2007) who accommodates interactions in indirect utility functions.

Despite the above efforts, there are still unresolved conceptual and methodological issues pertaining to: (1) the form of non-additive utility functions, (2) the specification of stochasticity in non-additive utility functions, (3) estimation of parameters with increasing number of choice alternatives, and (4) interpretation of the resulting dependency patterns in consumption. Resolving these issues will be a big step forward in enhancing the behavioral realism of KKT-based RUM MDC models. Further, within the context of non-additively separable preferences, it is important to recognize asymmetric dependencies in consumption. For example, the purchase of a new car may lead to increased gasoline consumption, but not the other way round.

7.2 Flexible Stochastic Specifications

The above discussion was in the context of the form of the utility function. But there is potential for improving the stochastic specification as well. For example, most studies assume IID extreme value random error terms in the utility function. Recent advances on relaxing the IID assumption, specifically via employing multivariate extreme value (MEV) distributions, have been discussed in Section 4.2. Although we are now able to estimate KKT-based RUM MDC models with general MEV stochastic distributions, no clear understanding exists on how different stochastic specifications and utility functional forms influence the properties of KKT models. Examining the substitution patterns implied by the different stochastic assumptions in KKT-based MDC models is a useful avenue for research. Further, the estimation of the MDCP model with MVN distributed stochasticity (as discussed in Section 4.4) is an important avenue for investigation.

7.3 Multiple Constraints

Most MDC model applications to date consider only a single linear binding constraint as governing the consumption decisions (*e.g.*, the linear constraint in Equation 1). This stems from an implicit assumption that only a single resource is needed to consume goods. However, in numerous empirical contexts, multiple types of resources, such as time, money and space, need to be expended to acquire and consume goods. While the role of multiple constraints has been long recognized in microeconomic theory (see Becker, 1965), the typical approach to accommodating the different constraints has been to convert them all into a single effective constraint. For example, the time constraint has been collapsed into the money constraint using a monetary value of time. In many situations, however, it is important to consider the different constraints in their own right, because resources may not always be freely exchangeable with each other. To address this issue, a handful of recent studies (Satomura *et al*., 2011; Castro *et al*., 2012; Pinjari and Sivaraman, 2012) have provided model formulations to accommodate multiple linear constraints with additive utility functional forms. Satomura *et al*. (2011) provided a formulation to account for the role of money and space constraints in consumers' decisions on soft drink purchases. Castro *et al*. (2012) provide a general treatment of the issue by providing formulations for different scenarios such as complete demand systems (*i.e.,* a case without the need of a Hicksian composite good), and incomplete demand systems (a case with the Hicksian composite good). Pinjari and Sivaraman (2012) provide a time- and money-constrained formulation in the context of households' annual vacation travel destination and mode choices.

7.4 Beyond Simple, Linear Constraints

The above discussion suggests that we have just begun to move toward models with multiple constraints. It is worth noting, however, that most of the literature on MDC modeling is geared toward simple, linear constraints that do not represent the complexity of situations consumers face in reality. There are several reasons why linear constraints do not hold. First, linear constraints represent a constant price per unit consumption (or a constant rate of resource-use). In many situations, however, prices vary with the amount of consumption leading to non-linear budget constraints. A classic example of such non-linear budgets is block pricing typically used in energy markets (e.g., electricity pricing). While the issue has long been recognized in the classical econometric literature on estimating demand functions, it is yet to be given due consideration in MDC choice studies. Second, linear constraints do not accommodate fixed costs (or setup costs) which cannot be converted into a constant price per unit consumption. For example, travel cost to a vacation destination is a *fixed* cost, unlike the lodging costs at the destination which can be treated as *variable* with a constant price per night.

Solving the consumer's direct utility maximization problem with non-linear constraints can become rather tedious, because the KKT conditions alone may not be sufficient anymore. In a recent study, Parizat and Shachar (2010) employ an enumeration approach to solve a direct utility maximization problem in the context of individuals' weekly leisure time allocation with fixed costs (e.g., ticket costs of going to a movie, the price of a meal). They acknowledge rather large computation times to estimate the parameters for their 12-alternative case. Thus, an alternative approach to incorporate non-linear constraints may be to work with the dual problem using indirect utility functions. Lee and Pitt (1987) provide a methodological treatment of incorporating block pricing with the dual approach. Further studies exploring this approach may enhance our ability to incorporate block prices. Another approach is to econometrically "treat" the inherent endogeneity between prices and consumption due to the dependency of prices on consumption, for example, by estimating price functions simultaneously with the consumer preferences (*i.e.,* utility functions). This approach can potentially help in dealing with demand-supply interactions in the market as well (see Berry *et al*., 1995).

7.5 Prediction and Welfare Analysis with Flexible Model Structures

Thanks to recent advances, we now have simple and computationally efficient methods to apply KKT models with additive utility forms for forecasting and welfare analysis purposes. As the field moves forward with the specification and estimation of more flexible MDC models, it is important to develop methods to apply these models as well. The prediction procedures proposed by von Haefen *et al*. (2004) and Pinjari and Bhat (2011) based on Karush-Kuhn-Tucker conditions of optimality can potentially be extended to the case with multiple linear binding constraints as well, although with additional layers of computational effort (as many as the number of constraints). However, these procedures fall apart in situations with non-additive utility functions, as they are critically hinged upon the additive utility assumption. Similarly, the presence of non-linear constraints can make it difficult to apply KKT conditions alone for solving the utility maximization problem. Resolving each of these issues is a welcome research direction.

Another useful direction of research is in the context of additive utility functions with a simple linear constraint. While we are now able to exploit the KKT conditions for obtaining the conditional predictions (given specific values of the random terms), we have not been able to characterize the unconditional distributions of the demand functions. In the presence of corner solutions, it is difficult to arrive at closed form expressions for the demand functions from Equation (1). Perhaps this is why we are not aware of successful attempts to arrive at analytical expressions for price elasticities and sensitivities to explanatory variables. Besides, application of these models requires the simulation of the stochastic terms. In some cases, such as the case with MEV stochastic distributions, the stochastic terms themselves are difficult to simulate. Thus development of fast methods to simulate MEV distributions can aid in the application of KKT models with such stochastic specifications.

8. SUMMARY

There has been an increasing recognition of the "multiple discrete-continuous (MDC)" nature of consumer choices. Over the past decade, the field has witnessed exciting developments in modeling MDC choices, especially with the advancement of the Karush-Kuhn-Tucker (KKT) approach to modeling consumer behavior based on random utility maximization (RUM). Notable developments include:

(a) Clever specifications with distributional assumptions that lead to closed-form probability expressions enabling easy estimation of the structural parameters (*e.g.*, the MDCEV model),

(b) Application of the KKT approach to model MDC choices in a variety of empirical contexts,

(c) Formulation of computationally efficient prediction/welfare analysis methods with KKT models,

(d) Extension of the basic RUM specification in Equation (1) to accommodate richer patterns of heterogeneity in consumer preferences and to allow flexibility in distributional assumptions. Most of these extensions have been "econometrically" oriented, akin to the extensions of the multinomial logit model in the traditional discrete choice analysis literature.

In the recent past, there has been an increasing recognition of the need to extend the basic formulation of consumer's utility maximization in Equation (1) in the following directions:

(a) Flexible functional forms for the utility specification, such as non-additive utility forms

(b) Flexible stochastic specifications for the utility functions, such as MVN distributions

(c) Flexibility in the specification of constraints faced by the consumer, including multiple inter-related constraints, and non-linear constraints.

Given the pace of recent developments, we optimistically look forward to seeing model formulations, estimation methods, and prediction/welfare analysis procedures for a general framework with non-additive utility forms, flexible stochastic distributional assumptions, and general forms of constraints.

ACKNOWLEDGEMENTS

This book chapter draws heavily in some places from papers by Bhat and colleagues, and a recent workshop report by Abdul Pinjari, Chandra Bhat, and David Bunch.

REFERENCES

Becker, G.S. (1965) A theory of the allocation of time. *The Economic Journal*, 75(299), 493-517.

Berry, S., J. Levinsohn, and A. Pakes (1995) Automobile prices in market equilibrium. *Econometrica*, 60(4), 841-890.

Bhat, C.R. (2005) A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions. *Transportation Research Part B*, 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 Part B*, 42(3), 274-303.

Bhat, C.R., M. Castro, and M. Khan (2012), A new estimation approach for the multiple discrete-continuous probit (MDCP) choice model, Technical paper, Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin.

Bhat, C.R., and A.R. Pinjari (2010) The generalized multiple discrete-continuous extreme value (GMDCEV) model: Allowing for non-additively separable and flexible utility forms. Working paper, Department of Civil, Architectural and Environmental Engineering, The University of Texas at Austin.

Bhat, C.R., and R. Sidharthan (2011) A simulation evaluation of the maximum approximate composite marginal likelihood (MACML) estimator for mixed multinomial probit models. *Transportation Research Part B*, 45(7), 940-953.

Bhat, C.R., S. Srinivasan, and S. Sen (2006) A joint model for the perfect and imperfect substitute goods case: application to activity time-use decisions. *Transportation Research Part B*, 40(10), 827-850.

Bhat, C.R., S. Sen, and N. Eluru (2009) The impact of demographics, built environment attributes, vehicle characteristics, and gasoline prices on household vehicle holdings and use. *Transportation Research Part B*, 43(1), 1-18.

Bunch, D.S. (2009) Theory-based functional forms for analysis of disaggregated scanner panel data. Working paper, Graduate School of Management, University of California-Davis.

Castro, M., C.R. Bhat, R.M. Pendyala, S.R. Jara-Diaz (2012) Accommodating multiple constraints in the multiple discrete-continuous extreme value (MDCEV) choice model. *Transportation Research Part B*, 46(6), 729-743.

Chiang J. (1991) The simultaneous approach to the whether, what, and how much to buy questions. *Marketing Science*, 10(4), 297-315.

Chintagunta, P. (1993) Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Science*, 12(2), 184-208.

Deaton, A., and J. Muellbauer (1980) *Economics and Consumer Behavior*. Cambridge University Press, Cambridge.

Dubin, J., and D. McFadden (1984) An econometric analysis of electricity appliance holdings and consumption. *Econometrica*, 52(2), 345-362.

Epstein, L.G. (1982) Integrability of incomplete systems of demand functions. *Review of Economic Studies* 49(3), 411-425.

Gentzkow, M. (2007) Valuing new goods in a model with complementarity: Online newspapers. *American Economic Review*, 97(3), 713-744.

Hanemann, W.M. (1978) A methodological and empirical study of the recreation benefits from water quality improvement. Ph.D. dissertation, Department of Economics, Harvard University.

Hanemann W.M. (1984) Discrete/continuous models of consumer demand. *Econometrica*, 52(3), 541-561.

Kim, J., G.M. Allenby, and P.E. Rossi (2002) Modeling consumer demand for variety. *Marketing Science*, 21(3), 229-250.

Lee, L.F., and M.M. Pitt (1986) Microeconometric demand systems with binding nonnegativity constraints: The dual approach. *Econometrica*, 54(5), 1237-1242.

Lee, L.F., and M.M. Pitt (1987) Microeconometric models of rationing, imperfect markets, and non-negativity constraints. *Journal of Econometrics*, 36(1-2), 89-110.

Lee, S., J. Kim, and G. Allenby (2010) A direct utility model for asymmetric complements. Working paper, Ohio State University.

Mäler, K-G. (1974) *Environmental Economics: A Theoretical Inquiry*. The Johns Hopkins University Press for Resources for the Future, Baltimore, MD.

McFadden, D. (1978) Modelling the choice of residential location, in A. Karlquist *et al.* (ed.), *Spatial Interaction Theory and Residential Location*, North-Holland, Amsterdam, 75-96.

Menezes, T.A., F.G. Silveira, C.R. Azzoni (2005) Demand elasticities for food products: a two-stage budgeting system. *NEREUS-USP*, São Paulo, 2005 (TD Nereus 09-2005).

Parizat, S., and R. Shachar (2010) When Pavarotti meets Harry Potter at the Super Bowl Working paper, Tel Aviv University.

Pinjari, A.R. (2011) Generalized extreme value (GEV)-based error structures for multiple discrete-continuous choice models. *Transportation Research Part B*, 45(3), 474-489.

Pinjari, A.R., and C.R. Bhat (2010) A multiple discrete-continuous nested extreme value (MDCNEV) model: Formulation and application to non-worker activity time-use and timing behavior on weekdays. *Transportation Research Part B*, 44(4), 562-583.

Pinjari, A.R., and C.R. Bhat (2011) An efficient forecasting procedure for Kuhn-Tucker consumer demand model systems: application to residential energy consumption analysis. Working paper, University of South Florida.

Pinjari, A.R., and V. Sivaraman (2012) A time and money budget constrained model of long-distance vacation travel demand. Working paper, University of South Florida.

Phaneuf, D.J., C.L. Kling, and J.A. Herriges (2000) Estimation and welfare calculations in a generalized corner solution model with an application to recreation demand. *The Review of Economics and Statistics*, 82(1), 83-92.

Roy, R. (1947) La distribution du revenu entre les divers biens. *Econometrica*, 15(3), 205-225.

Satomura, S., J. Kim, and G. Allenby (2011) Multiple constraint choice models with corner and interior solutions. *Marketing Science*, 30(3), 481-490.

Tong, Y.L. (1990) *The Multivariate Normal Distribution*. Springer-Verlag, New York.

Train, K. (2003) *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge, United Kingdom.

Van Nostrand, C., V. Sivaraman, and A.R. Pinjari (2013) Analysis of annual, long-distance, vacation travel demand in the United States: A multiple discrete-continuous choice framework. *Transportation*, 40(1), 151-171.

Vasquez Lavin, F., and M. Hanemann (2009) Functional forms in discrete/continuous choice models with general corner solution. Working paper, University of California Berkeley.

von Haefen, R.H. (2010) Incomplete demand systems, corner solutions, and welfare measurement. *Agricultural and Resource Economics Review* 39(1), 22-36.

von Haefen, R.H., and D.J. Phaneuf (2005) Kuhn-Tucker demand system approaches to nonmarket valuation. In *Applications of Simulation Methods in Environmental and Resource Economics*, R. Scarpa and A. Alberini (Eds.), pp. 135-158, Springer, Dordrecht, The Netherlands.

von Haefen, R.H., D.J. Phaneuf, and G.R. Parsons (2004) Estimation and welfare analysis with large demand systems. *Journal of Business and Economic Statistics*, 22(2), 194-205.

Wales, T.J., and A.D. Woodland (1983) Estimation of consumer demand systems with binding non-negativity constraints. *Journal of Econometrics*, 21(3), 263-285.

LIST OF FIGURES

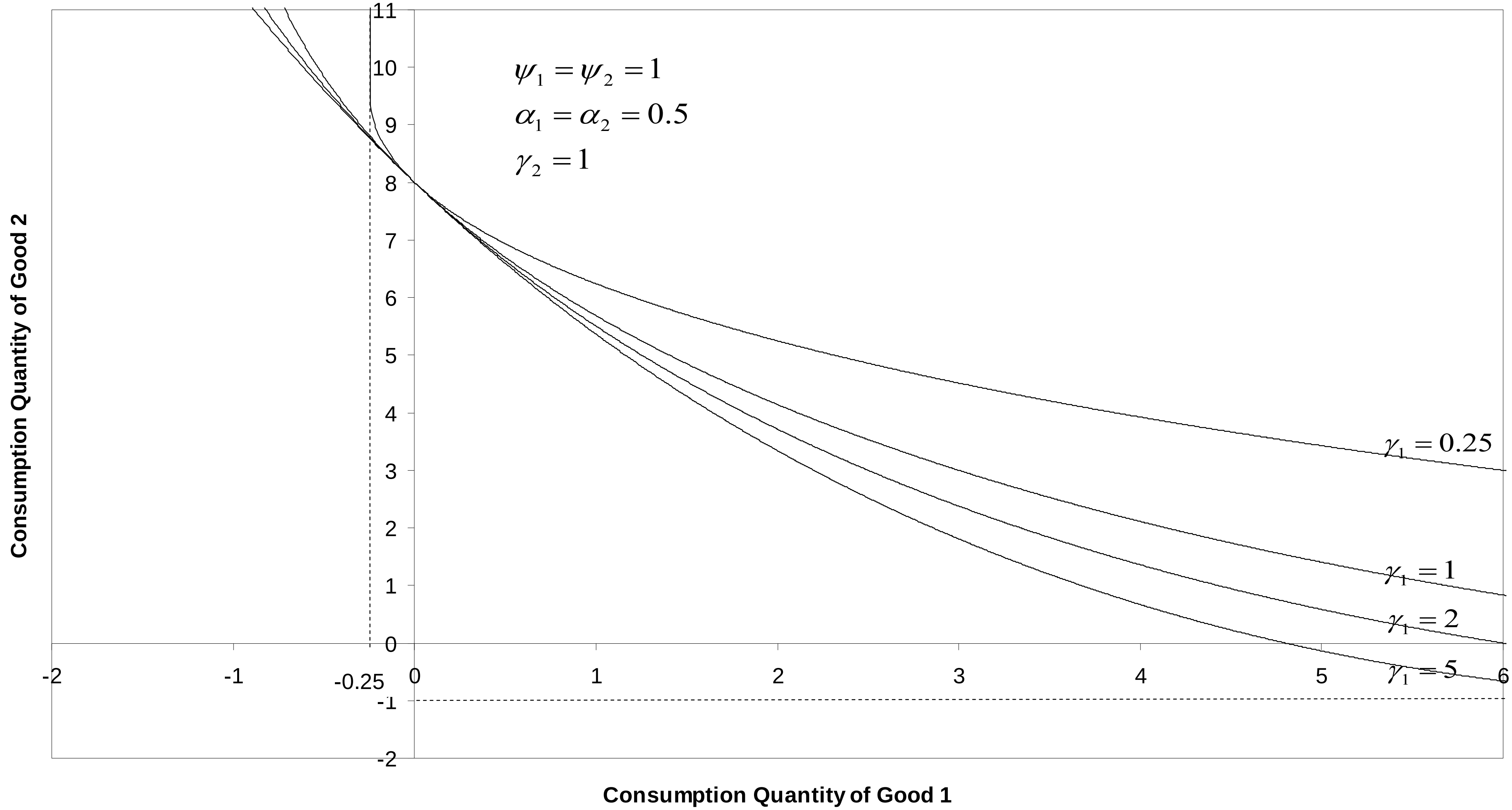

Figure 1. Indifference Curves Corresponding to Different Values of γ_1

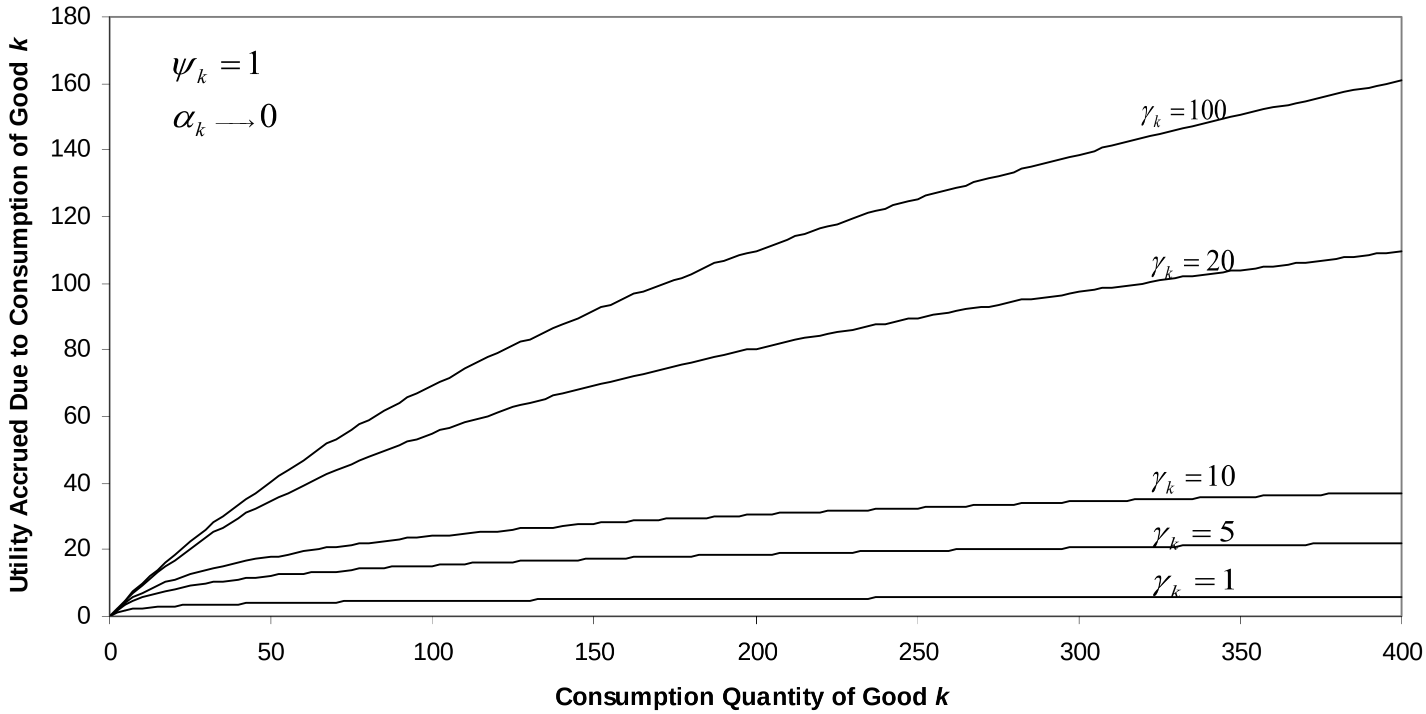

Figure 2. Effect of γ_k Value on Good *k*'s Subutility Function Profile

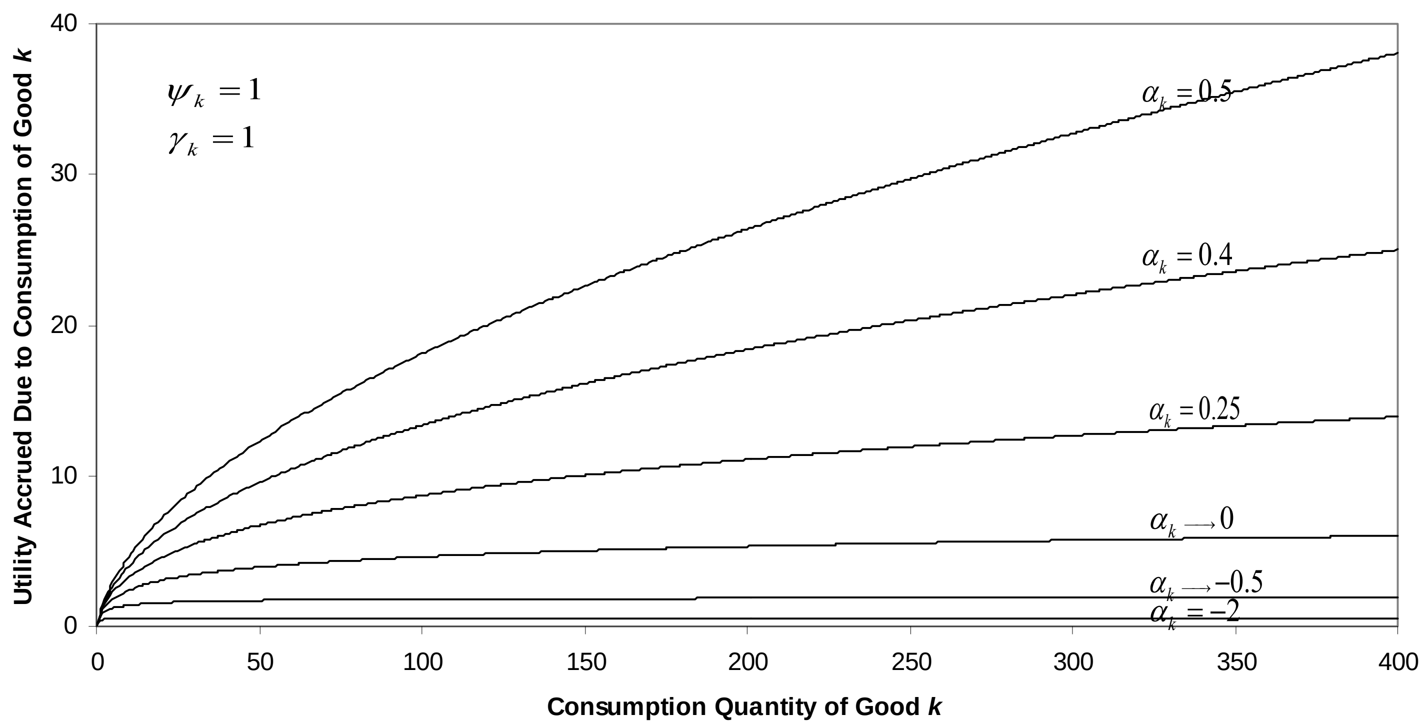

Figure 3. Effect of α_k Value on Good *k*'s Subutility Function Profile