The Multiple Discrete-Continuous Extreme Value (MDCEV) Model: Role of Utility Function Parameters, Identification Considerations, and Model Extensions

Chandra R. Bhat*
The University of Texas at Austin
Department of Civil, Architectural & Environmental Engineering
1 University Station C1761, Austin, Texas 78712-0278
Tel: 512-471-4535, Fax: 512-475-8744,
Email: bhat@mail.utexas.edu

* This paper was completed when the author was a Visiting Professor at the Institute of Transport and Logistics Studies, Department of Economics and Business, University of Sydney.

ABSTRACT

Many consumer choice situations are characterized by the simultaneous demand for multiple alternatives that are imperfect substitutes for one another. A simple and parsimonious Multiple Discrete-Continuous Extreme Value (MDCEV) econometric approach to handle such multiple discreteness was formulated by Bhat (2005) within the broader Kuhn-Tucker (KT) multiple discrete-continuous economic consumer demand model of Wales and Woodland (1983). This paper examines several issues associated with the MDCEV model and other extant KT multiple discrete-continuous models. Specifically, the paper proposes a new utility function form that enables clarity in the role of each parameter in the utility specification, presents identification considerations associated with both the utility functional form as well as the stochastic nature of the utility specification, extends the MDCEV model to the case of price variation across goods and to general error covariance structures, discusses the relationship between earlier KT-based multiple discrete-continuous models, and illustrates the many technical nuances and identification considerations of the multiple discrete-continuous model structure through empirical examples. The paper also highlights the technical problems associated with the stochastic specification used in the KT-based multiple discrete-continuous models formulated in recent Environmental Economics papers.

Keywords: Discrete-continuous system, Multiple discreteness, Kuhn-Tucker demand systems, Mixed discrete choice, Random Utility Maximization.

1. INTRODUCTION

Multiple discreteness (*i.e.*, the choice of multiple, but not necessarily all, alternatives simultaneously) is a rather ubiquitous characteristic of consumer decision-making.[1] Examples of multiple discreteness include situations where an individual may decide to participate in multiple kinds of maintenance and leisure activities within a given time period (Bhat, 2005), or a household may own a mix of different kinds of vehicles (such as a sedan and a pick-up truck or a sedan and a minivan; see Bhat and Sen, 2006). Such multiple discrete situations may be modeled using the traditional random utility-based (RUM) single discrete choice models by identifying all combinations or bundles of the "elemental" alternatives, and treating each bundle as a "composite" alternative (the term "single discrete choice" is used to refer to the case where a decision-maker chooses only one alternative from a set of alternatives). A problem with this approach, however, is that the number of composite alternatives explodes with the number of elemental alternatives. Another approach is to use the multivariate probit (logit) methods of Manchanda *et al.* (1999), Baltas (2004), Edwards and Allenby (2003), and Bhat and Srinivasan (2005). But this approach is not based on a rigorous underlying utility-maximizing framework of multiple discreteness; rather, it represents a statistical "stitching" of univariate utility maximizing models. In both the approaches discussed above to handle multiple discreteness, there is also no explicit way to accommodate the diminishing marginal returns (*i.e.*, satiation) in the consumption of an alternative. Additionally, and related to the above point, it is very cumbersome, even if conceptually feasible, to include a continuous dimension of choice (for example, modeling the durations of participation in the chosen activity purposes, in addition to the choice of activity purpose).[2]

---

[1] A brief history of the term "multiple discreteness" is in order here. Traditional discrete choice models focus on the selection of a single alternative from the set of available alternatives on a purchase occasion. That is, they consider the "extreme corner solution problem". Hanemann, in his 1978 dissertation, used the term "generalized corner solution problem" to refer to the situation where multiple alternatives may be chosen simultaneously. Hendel (1999) appears to have been the first to coin the term "multiple discreteness" to refer to the choice of multiple alternatives. This term is also used by Dube (2004).

[2] Another approach for multiple discreteness is the one proposed by Hendel (1999) and Dube (2004). These researchers consider the case of "multiple discreteness" in the purchase of multiple varieties within a particular product category as the result of a stream of expected (but unobserved to the analyst) future consumption decisions between successive shopping purchase occasions (see also Walsh, 1995). During each consumption occasion, the standard discrete choice framework of perfectly substitutable alternatives is invoked, so that only one product is consumed. Due to varying tastes across individual consumption occasions between the current shopping purchase and the next, consumers are observed to purchase a variety of goods at the current shopping occasion. A Poisson distribution is assumed for the number of consumption occasions and a normal distribution is assumed regarding varying tastes to complete the model specification. Such a "vertical" variety-seeking model, of course, is different

Wales and Woodland (1983) proposed two alternative ways to handle situations of multiple discreteness within a behaviorally-consistent utility maximizing framework. Both approaches assume a direct utility function $U(\boldsymbol{x})$ that is assumed to be quasi-concave, increasing, and continuously differentiable with respect to the consumption quantity vector $\boldsymbol{x}$.[3] Consumers maximize the utility function subject to a linear budget constraint, which is binding in that all the available budget is invested in the consumption of the goods; that is, the budget constraint has an equality sign rather than a '$\leq$' sign. This binding nature of the budget constraint is the result of assuming an increasing utility function, and also implies that at least one good will be consumed. The difference in the two alternative approaches proposed by Wales and Woodland (1983) is in how stochasticity, non-negativity of consumption, and corner solutions (*i.e.*, zero consumption of some goods) are accommodated, as briefly discussed below (see Wales and Woodland, 1983 and Phaneuf *et al.*, 2000 for additional details).

The first approach, which Wales and Woodland label as the Amemiya-Tobin approach, is an extension of the classic microeconomic approach of adding normally distributed stochastic terms to the budget-constrained utility-maximizing share equations. In this approach, the direct utility function $U(\boldsymbol{x})$ itself is assumed to be deterministic by the analyst, and stochasticity is introduced post-utility maximization. The justification for the addition of such normally distributed stochastic terms to the deterministic utility-maximizing allocations is based on the notion that consumers make errors in the utility-maximizing process, or that there are measurement errors in the collection of share data, or that there are unknown factors (from the analyst's perspective) influencing actual consumed shares. However, the addition of normally distributed error terms to the share equations in no way restricts the shares to be positive and less than 1. The contribution of Wales and Woodland was to devise a stochastic formulation, based on the earlier work of Tobin (1958) and Amemiya (1974), that (a) respects the unit simplex range constraint for the shares, (b) accommodates the restriction that the shares sum to one, and (c) allows corner solutions in which one or more alternatives are not consumed. They achieve this by assuming that the observed shares for the (K-1) of the K alternatives follow a truncated multivariate normal distribution (note that since the shares across alternatives have to sum to

---

from the "horizontal" variety seeking model considered in this paper, where the choice is considered to be among inherently imperfect substitutes at the choice occasion (see Kim *et al.*, 2002 and Bhat, 2005).

[3] The assumption of a quasi-concave utility function is simply a manifestation of requiring the indifference curves to be convex to the origin (see Deaton and Muellbauer, 1980, page 30 for a rigorous definition of quasi-concavity). The assumption of an increasing utility function implies that $U(\boldsymbol{x}^1) > U(\boldsymbol{x}^0)$ if $\boldsymbol{x}^1 > \boldsymbol{x}^0$.

one, there is a singularity generated in the K-variate covariance matrix of the K shares, which can be accommodated by dropping one alternative). However, an important limitation of the Amemiya-Tobin approach of Wales and Woodland is that it does not account for corner solutions in its underlying behavior structure. Rather, the constraint that the shares have to lie within the unit simplex is imposed by *ad hoc* statistical procedures of mapping the density outside the unit simplex to the boundary points of the unit simplex.

The second approach suggested by Wales and Woodland, which they label as the Kuhn-Tucker approach, is based on the Kuhn Tucker or KT (1951) first-order conditions for constrained random utility maximization (see Hanemann, 1978, who uses such an approach even before Wales and Woodland). Unlike the Amemiya-Tobin approach, the KT approach employs a more direct stochastic specification by assuming the utility function $U(\boldsymbol{x})$ to be random (from the analyst's perspective) over the population, and then derives the consumption vector for the random utility specification subject to the linear budget constraint by using the KT conditions for constrained optimization. Thus, the stochastic nature of the consumption vector in the KT approach is based fundamentally on the stochastic nature of the utility function. Consequently, the KT approach immediately satisfies all the restrictions of utility theory, and the stochastic KT first-order conditions provide the basis for deriving the probabilities for each possible combination of corner solutions (zero consumption) for some goods and interior solutions (strictly positive consumption) for other goods. The singularity imposed by the "adding-up" constraint is accommodated in the KT approach by employing the usual differencing approach with respect to one of the goods, so that there are only (K-1) interdependent stochastic first-order conditions.

Among the two approaches discussed above, the KT approach constitutes a more theoretically unified and behaviorally consistent framework for dealing with multiple discreteness consumption patterns. However, the KT approach did not receive much attention until relatively recently because the random utility distribution assumptions used by Wales and Woodland lead to a complicated likelihood function that entails multi-dimensional integration. Kim *et al.* (2002) addressed this issue by using the Geweke-Hajivassiliou-Keane (or GHK) simulator to evaluate the multivariate normal integral appearing in the likelihood function in the KT approach. Also, different from Wales and Woodland, Kim *et al.* used a generalized variant of the well-known translated constant elasticity of substitution (CES) direct utility function (see

Pollak and Wales, 1992; page 28) rather than the quadratic direct utility function used by Wales and Woodland. In any case, the Kim *et al.* approach, like the Wales and Woodland approach, is unnecessarily complicated because of the need to evaluate truncated multivariate normal integrals in the likelihood function. In contrast, Bhat (2005) introduced a simple and parsimonious econometric approach to handle multiple discreteness, also based on the generalized variant of the translated CES utility function but with a multiplicative log-extreme value error term. Bhat's model, labeled the multiple discrete-continuous extreme value (MDCEV) model, is analytically tractable in the probability expressions and is practical even for situations with a large number of discrete consumption alternatives. In fact, the MDCEV model represents the multinomial logit (MNL) form-equivalent for multiple discrete-continuous choice analysis and collapses exactly to the MNL in the case that each (and every) decision-maker chooses only one alternative.

Independent of the above works of Kim *et al.* and Bhat, there has been a stream of research in the environmental economics field (see Phaneuf *et al.*, 2000; von Haefen *et al.*, 2004; von Haefen, 2003a; von Haefen, 2004; von Haefen and Phaneuf, 2005; Phaneuf and Smith, 2005) that has also used the KT approach to multiple discreteness. These studies use variants of the linear expenditure system (LES) as proposed by Hanemann (1978) and the translated CES for the utility functions, and use multiplicative log-extreme value errors. However, the error specification in the utility function is different from that in Bhat's MDCEV model, resulting in a different form for the likelihood function (more on this in Section 6).

Within the context of the KT approach to handling multiple discreteness, the purpose of this research is five-fold. The first objective is to reformulate the utility specification used in earlier studies in a way that explicitly clarifies the role of each parameter in the utility specification. The second objective is to present identification considerations related to both the functional form as well as the stochastic nature of the utility specification. The third objective is to derive the MDCEV model expression for the case when there is price variation across goods and to extend the MDCEV model to accommodate generalized extreme value (GEV)-based and other correlation structures. The fourth objective is to discuss the relationship between the models of Kim *et al.* (2002), the KT formulations used in Environmental Economics, and the MDCEV formulation. The fifth objective is to illustrate the technical issues related to the properties and identification of the MDCEV model through empirical illustrations.

The rest of the paper is structured as follows. The next section formulates a functional form for the utility specification that enables the isolation of the role of different parameters in the specification. This section also identifies empirical identification considerations in estimating the parameters in the utility specification. Section 3 discusses the stochastic form of the utility specification, the resulting general structure for the probability expressions, and associated identification considerations. Section 4 derives the MDCEV structure for the new utility functional form used in the current paper, and extends this structure to more general error structure specifications. For presentation ease, Sections 2 through 4 consider the case of the absence of an outside good. In Section 5, we extend the discussions of the earlier sections to the case when an outside good is present. Section 6 compares the earlier multiple discrete-continuous models used in the literature with the one formulated in the current paper. Section 7 provides empirical illustrations to reinforce the theoretical issues discussed in earlier sections. The final section concludes the paper.

2. FUNCTIONAL FORM OF UTILITY SPECIFICATION

We consider the following functional form for utility in this paper, based on a generalized variant of the translated CES utility function:

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

where $U(\boldsymbol{x})$ 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 (1) 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 outside good, so that corner solutions (*i.e.*, zero consumptions) are allowed for all the goods k (this assumption is being made only to streamline the presentation and should not be construed as limiting in any way; the assumption is relaxed in a straightforward manner as discussed in Section 5). The possibility of corner solutions implies that the term γ_k, which is a translation parameter, should be greater than zero for all k.[4] The reader will note that there is an assumption of additive separability of preferences in the utility form of Equation (1), which immediately implies that none of the goods are *a priori* inferior and all the goods are strictly Hicksian substitutes (see Deaton and Muellbauer, 1980; page 139). Additionally, additive separability implies that the marginal utility with respect to any good is independent of the levels of all other goods.[5]

The form of the utility function in Equation (1) is different from that used in earlier studies. The reason for the specific functional form adopted here is to highlight the role of the various parameters ψ_k, γ_k and α_k, and explicitly indicate the inter-relationships between these parameters that relate to theoretical and empirical identification issues.[6] Finally, it should be noted that the utility form of Equation (1) collapses to the following linear expenditure system (LES) form when $\alpha_k \to 0 \; \forall \; k$ (see Appendix A; the LES form of the type below appears to have been first used by Hanemann, 1978).

$$U(\boldsymbol{x}) = \sum_{k=1}^{K} \gamma_k \psi_k \ln\left(\frac{x_k}{\gamma_k} + 1\right) \tag{2}$$

---

[4] As illustrated in Kim *et al.* (2002) and Bhat (2005), the presence of the translation parameters makes the indifference curves strike the consumption axes at an angle (rather than being asymptotic to the consumption axes), thus allowing corner solutions.

[5] Some other studies assume the overall utility to be derived from the characteristics embodied in the goods, rather than using the goods as separate entities in the utility function. The reader is referred to Chan (2006) for an example of such a characteristics approach to utility.

[6] As we will show later, however, the utility form we adopt is behaviorally and observationally indistinguishable from those used in Bhat (2005) and Kim *et al.* (2002) if γ_k is normalized to 1 for all k and $0 < \alpha_k < = 1$. It is also observationally indistinguishable from the utility form used in environmental economics under the condition that $\alpha_k \to 0$. Specifically, all these utility forms imply an identical set of Kuhn-Tucker first order conditions and demand. However, the various utility forms may not yield identical welfare measures. In our formulation of utility, we impose the untestable, but intuitive, condition of 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, or $U_k = 0$ if $x_k = 0$, where U_k is the sub-utility function for the k*th* good). The reader is referred to Hanemann (1984), von Haefen (2004), and Herriges *et al.* (2004) for a detailed discussion of the advantages of using the weak complementarity assumption. The use of the weak complementarity condition essentially amounts to a cardinal normalization restriction on utilities. But, as Herriges *et al.* (2004) indicate, the analyst will have to place some kind of a cardinal restriction on preferences anyway for welfare measurement, and weak complementarity is a natural choice in many circumstances. We will maintain the weak complementary cardinal normalization in the rest of this paper to simplify the algebra, though the ordinality of utilities should always be kept in mind.

2.1 Role of Parameters in Utility Specification

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:

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

It is obvious from above that ψ_k represents the baseline marginal utility, or the marginal utility at the point of zero consumption. Alternatively, the marginal rate of substitution between any two goods *k* and *l* at the point of zero consumption of both goods is $\frac{\psi_k}{\psi_l}$. This is the case regardless of the values of γ_k and α_k (unlike in earlier studies where the baseline marginal utility and the marginal rate of substitution, in general, are functions of multiple parameters). For two goods *i* and *j* with same unit prices, a higher baseline marginal utility for good *i* relative to good *j* implies that an individual will increase overall utility more by consuming good *i* rather than *j* at the point of no consumption of any goods. That is, the consumer will be more likely to consume good *i* than good *j*. Thus, a higher baseline ψ_k implies less likelihood of a corner solution for good *k*.

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 γ_2 = 1), while they are asymptotic to the y-axis at $x = -\gamma_1$.

Figure 1 also points to another role of the γ_k term as a satiation parameter. Specifically, the indifference curves get steeper in the positive orthant as the value of γ_1 increases, which implies a stronger preference (or lower satiation) for good 1 as γ_1 increases (with steeper indifference curve slopes, the consumer is willing to give up more of good 2 to obtain 1 unit of good 1). This point is particularly clear if we examine the profile of the sub-utility function for alternative k. Figure 2 plots the function for alternative k for $\alpha_k \to 0$ and ψ_k = 1, and for different values of γ_k. All of the curves have the same slope ψ_k = 1 at the origin point, because of the functional form used in this paper. 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. It is important to note that the entire range of satiation effects from immediate and full satiation (flat line) to linear satiation (constant marginal utility) can be accommodated by different values of γ_k for any given α_k value.

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 α_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 as proposed by Deaton and Muellbauer (1980) and applied in Hanemann (1984), Chiang (1991), Chintagunta (1993), and Arora *et al.* (1998), among others. Intuitively, when there is no satiation and the unit good prices are all the same, the consumer will invest all expenditure on the single good with the highest baseline (and constant) marginal utility (*i.e.*, the highest ψ_k value). This is the case of single discreteness.[7] 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 form in Equation (2), 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. Thus, for a given extensive margin of choice of good k, γ_k and α_k influence the quantity of good k consumed (or the intensive margin of choice) through their impact on satiation effects. 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, 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. This is illustrated in Figures 4a through 4d for $\psi_k = 1$

[7] If there is price variation across goods, one needs to take the derivative of the utility function with respect to expenditures (e_k) on the goods. In the case that α_k= 1 for all k, $U = \sum_k \psi_k(e_k / p_k)$, where ψ_k is the unit price of good k. Then $\partial U / \partial e_k = \psi_k / p_k$. In this situation, the consumer will invest all expenditures on the single good with the highest price-normalized marginal (and constant) utility ψ_k / p_k.

and for different satiation levels. In these figures, the subutility functions based solely on γ_k assume α_k = 0 (*i.e.*, these functions take the form of Equation (2)), while those based solely on α_k for incorporating satiation assume γ_k = 1 (note that γ_k, even if fixed, has to be positive to allow corner solutions). In all the figures, the profile based only on γ_k (the γ_k-profile) or α_k (the α_k-profile) tracks the profile based on the combination of values (the combination profile) reasonably well. For moderate satiations (Figures 4a and 4b), one of the two profiles does better than the other, based on how close the γ_k and α_k values in the combination profile are to the assumed value of α_k^* = 0 for the γ_k-profile and γ_k^{**} = 1 for the α_k-profile. For very low and very high satiations, both the α_k-profile and the γ_k-profile track the combination profile very closely. In actual application, it would behoove the analyst to estimate models based on both the α_k-profile and the γ_k-profile, and choose a specification that provides a better statistical fit.[8]

In cases where α_k values are estimated, these values need to be bounded from above at the value of 1. To enforce these conditions, α_k can be parameterized as $[1-\exp(-\delta_k)]$, with δ_k being the parameter that is estimated. Further, to allow the satiation parameters (*i.e.*, the α_k values) to vary across individuals, Bhat (2005) writes $\delta_k = \theta_k' y_k$, where y_k is a vector of individual characteristics impacting satiation for the k^{th} alternative, and θ_k is a corresponding vector of parameters. In cases where γ_k values are estimated, these values need to be greater than zero, which can be maintained by reparameterizing γ_k as $\exp(\mu_k)$. Additionally, the translation parameters can be allowed to vary across individuals by writing $\mu_k = \varphi_k' w_k$, where w_k is a vector of individual characteristics for the k^{th} alternative, and φ_k is a corresponding vector of parameters.

[8] 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.

3. STOCHASTIC FORM OF UTILITY FUNCTION

The KT 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 KT approach for multiple discreteness, a multiplicative random element is introduced to the baseline marginal utility of each good as follows:

$$\psi(z_k, \varepsilon_k) = \psi(z_k) \cdot e^{\varepsilon_k}, \tag{4}$$

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 j. 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:

$$\psi(z_k, \varepsilon_k) = \exp(\beta' z_k + \varepsilon_k). \tag{5}$$

The z_k vector in the above equation includes a constant term. The overall random utility function of Equation (1) then takes the following form:

$$U(\boldsymbol{x}) = \sum_k \frac{\gamma_k}{\alpha_k} \left[\exp(\beta' z_k + \varepsilon_k)\right] \cdot \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\} \tag{6}$$

From the analyst's perspective, the individual is maximizing random utility subject to the binding linear budget constraint that $\sum_{k=1}^{K} e_k = E$, where E is total expenditure or income (or some other appropriately defined total budget quantity), $e_k = p_k x_k$, and p_k is the unit price of good k.

3.1 Optimal Expenditure Allocations

The analyst can solve for the optimal expenditure allocations by forming the Lagrangian and applying the Kuhn-Tucker (KT) conditions.[9] The Lagrangian function for the problem is:

$$\mathscr{L} = \sum_k \frac{\gamma_k}{\alpha_k} \left[\exp(\beta' z_k + \varepsilon_k)\right] \left\{ \left(\frac{e_k}{\gamma_k p_k} + 1 \right)^{\alpha_k} - 1 \right\} - \lambda \left[\sum_{k=1}^{K} e_k - E \right], \tag{7}$$

[9] For reasons that will become clear later, we solve for the optimal expenditure allocations e_k for each good, not the consumption amounts x_k of each good. This is different from earlier studies that focus on the consumption of goods.

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

$$\left[\frac{\exp(\beta' z_k + \varepsilon_k)}{p_k}\right]\left(\frac{e_k^*}{\gamma_k p_k}+1\right)^{\alpha_k - 1} - \lambda = 0\text{, if } e_k^* > 0\text{, } k = 1, 2, \ldots, K \tag{8}$$

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

The optimal demand satisfies the conditions in Equation (8) plus the budget constraint $\sum_{k=1}^{K} e_k^* = E$. The budget constraint implies that only K-1 of the e_k^* values need to be estimated, since the quantity consumed of any one good is automatically determined from the quantity consumed of all the other goods. To accommodate this constraint, designate activity purpose 1 as a purpose to which the individual allocates some non-zero amount of consumption (note that the individual should participate in at least one of the K purposes, given that $E > 0$). For the first good, the KT condition may then be written as:

$$\lambda = \frac{\exp(\beta' z_1 + \varepsilon_1)}{p_1}\left(\frac{e_1^*}{\gamma_1 p_1}+1\right)^{\alpha_1 - 1} \tag{9}$$

Substituting for λ from above into Equation (8) for the other activity purposes (k = 2,…, K), and taking logarithms, we can rewrite the KT conditions as:

$$V_k + \varepsilon_k = V_1 + \varepsilon_1 \text{ if } e_k^* > 0 \;\; (k = 2, 3, \ldots, K)$$

$$V_k + \varepsilon_k < V_1 + \varepsilon_1 \text{ if } e_k^* = 0 \;\; (k = 2, 3, \ldots, K)\text{, where} \tag{10}$$

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

Also, note that, in Equation (10), a constant cannot be identified in the $\beta' z_k$ term for one of the K alternatives (because only the difference in the V_k from V_1 matters). Similarly, individual-specific variables are introduced in the V_k's for (K-1) alternatives, with the remaining alternative serving as the base.[10]

[10] These identification conditions are similar to those in the standard discrete choice model, though the origin of the conditions is different between standard discrete choice models and the multiple discrete-continuous models. In

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 allocates expenditure to the first M of the K goods is:

$$P(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0) = |J| \int_{\varepsilon_1=-\infty}^{+\infty} \int_{\varepsilon_{M+1}=-\infty}^{V_1-V_{M+1}+\varepsilon_1} \int_{\varepsilon_{M+2}=-\infty}^{V_1-V_{M+2}+\varepsilon_1} \cdots \int_{\varepsilon_{K-1}=-\infty}^{V_1-V_{K-1}+\varepsilon_1} \int_{\varepsilon_K=-\infty}^{V_1-V_K+\varepsilon_1}$$
$$f(\varepsilon_1, V_1 - V_2 + \varepsilon_1, V_1 - V_3 + \varepsilon_1, ..., V_1 - V_M + \varepsilon_1, \varepsilon_{M+1}, \varepsilon_{M+2}, ..., \varepsilon_{K-1}, \varepsilon_K) \tag{11}$$
$$d\varepsilon_K d\varepsilon_{K-1} ... d\varepsilon_{M+2} d\varepsilon_{M+1} d\varepsilon_1 \,,$$

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

$$J_{ih} = \frac{\partial[V_1 - V_{i+1} + \varepsilon_1]}{\partial e_{h+1}^*} = \frac{\partial[V_1 - V_{i+1}]}{\partial e_{h+1}^*}; \; i, h = 1, 2, \ldots, M-1. \tag{12}$$

The probability expression in Equation (11) is a (K-M+1)-dimensional integral. The expression for the probability of all goods being consumed is one-dimensional, while the expression for the probability of only the first good being consumed is K-dimensional. The dimensionality of the integral can be reduced by one by noticing that the KT 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 $g(\tilde{\varepsilon}_{21}, \tilde{\varepsilon}_{31}, ..., \tilde{\varepsilon}_{K1})$. Then, Equation (11) may be written in the equivalent (K-M)-integral form shown below:

$$P(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0) = |J| \int_{\tilde{\varepsilon}_{M+1,1}=-\infty}^{V_1-V_{M+1}} \int_{\tilde{\varepsilon}_{M+2,1}=-\infty}^{V_1-V_{M+2}} \cdots \int_{\tilde{\varepsilon}_{K-1,1}=-\infty}^{V_1-V_{K-1}} \int_{\tilde{\varepsilon}_{K,1}=-\infty}^{V_1-V_K}$$
$$g(V_1 - V_2, V_1 - V_3, ..., V_1 - V_M, \tilde{\varepsilon}_{M+1,1}, \tilde{\varepsilon}_{M+2,1}, ..., \tilde{\varepsilon}_{K,1}) d\tilde{\varepsilon}_{K,1} d\tilde{\varepsilon}_{K-1,1} ... d\tilde{\varepsilon}_{M+1,1} \tag{13}$$

The equation above indicates that the probability expression for the observed optimal expenditure pattern of goods is completely characterized by the (K-1) error terms in difference form. Thus, all that is estimable is the (K-1)x(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 $g(.)$

---

standard discrete choice models, individuals choose the alternative with highest utility, so that all that matters is relative utility. In multiple discrete-continuous models, the origin of these conditions is the adding up (or budget) constraint associated with the quantity of consumption of each good that leads to the KT first order conditions of Equation (10).

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. To see this, consider the case of independent and identically distributed error terms for the ε_k terms, which leads to a (*K*-1)x(*K*-1) covariance matrix for $\tilde{\varepsilon}_{k1}$ (k = 2,3,…,K) with diagonal elements equal to twice the value of scale parameter of the ε_k terms and off-diagonal elements equal to the scale parameter of the ε_k terms. Let the unit prices of all goods be the same (see Bhat, 2005; Bhat and Sen, 2006; Bhat *et al.*, 2006 for examples where the weights or prices on the goods in the budget constraint are equal). Consider the utility function in Equation (6) and another utility function as given below:

$$\tilde{U} = \sum_k \frac{\gamma_k}{\alpha_k^*}\left[\exp\{\sigma \times (\beta' z_k + \varepsilon_k)\}\right] \cdot \left\{\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k^*} - 1\right\} \tag{14}$$

The scale of the error terms in the utility function in the above expression is σ times the scale of the error terms in Equation (6). Let $\alpha_k^* = \sigma(\alpha_k - 1) + 1$, where α_k is the satiation parameter in the original Equation (6).[11] The KT conditions for optimal expenditure for this modified utility function can be shown to be:

$$V_k^* + \sigma\varepsilon_k = V_1^* + \sigma\varepsilon_1 \text{ if } e_k^* > 0 \ \ (k = 2, 3, \ldots, K)$$

$$V_k^* + \sigma\varepsilon_k < V_1^* + \sigma\varepsilon_1 \text{ if } e_k^* = 0 \ \ (k = 2, 3, \ldots, K), \text{ where} \tag{15}$$

[11] Note that α_k^* is less than or equal to 1 by definition, because α_k is less than or equal to 1 and the scale σ should be non-negative.

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

If the unit prices are not all the same (*i.e.*, the unit prices of at least two of the K goods are different), the KT conditions above are different from the KT conditions in Equation (10). That is, the utility function in Equation (14) is unique and different from the utility function in Equation (6), which implies that the scale σ is identified. However, if the unit prices are all the same ($p_k = p \forall k$), it is straightforward to note that the KT conditions above collapse exactly to the KT conditions in Equation (10). In this case, the utility function in Equation (14) cannot be uniquely identified from the utility function in Equation (6), which implies that the scale σ is not identified theoretically. For convenience, the analyst can set the scale to 1.

In the case that the analyst uses a heteroscedastic specification with no variation in unit prices across alternatives, the scale of one of the alternatives has to be set to unity (similar to the case of the heteroscedastic extreme value or HEV model of Bhat, 1995). With a general error structure and no variation in unit prices, the identification considerations associated with a standard discrete choice model with correlated errors apply (see Train, 2003; Chapter 2).

4. SPECIFIC MODEL STRUCTURES

4.1 The MDCEV Model Structure

Following Bhat (2005), we specify an extreme value distribution for ε_k and assume that ε_k is independent of z_k $(k = 1, 2, \ldots, K)$. 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:

$$V_k = \beta' z_k + (\alpha_k - 1)\ln\left(\frac{e_k^*}{p_k} + 1\right) - \ln p_k \;\; (k = 1, 2, 3, ..., K), \text{ when the } \alpha \text{ - profile is used, and}$$
$$V_k = \beta' z_k - \ln\left(\frac{e_k^*}{\gamma_k p_k} + 1\right) - \ln p_k \; (k = 1, 2, 3, ..., K), \text{ when the } \gamma \text{ - profile is used.} \qquad (16)$$

As discussed earlier, it is generally not possible to estimate the V_k form in Equation (10), because the α_k terms and γ_k terms serve a similar satiation role.

From Equation (11), the probability that the individual allocates expenditure to the first M of the K goods ($M \geq 1$) is:

$$P\left(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0\right)$$

$$= |J| \int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty} \left\{ \left(\prod_{i=2}^{M} \frac{1}{\sigma} \lambda\left[\frac{V_1 - V_i + \varepsilon_1}{\sigma}\right] \right) \right\} \times \left\{ \prod_{s=M+1}^{K} \Lambda\left[\frac{V_1 - V_s + \varepsilon_1}{\sigma}\right] \right\} \frac{1}{\sigma} \lambda\left(\frac{\varepsilon_1}{\sigma}\right) d\varepsilon_1 , \tag{17}$$

where λ is the standard extreme value density function and Λ is the standard extreme value cumulative distribution function. The expression in Equation (17) simplifies to a remarkably simple and elegant closed-form expression. Bhat derived the form of the Jacobian for the case of equal unit prices across goods, which however can be extended in a simple fashion to accommodate the more general case of different unit prices. The resulting form for the determinant of the Jacobian has a compact structure given by:

$$|J| = \left(\prod_{i=1}^{M} c_i \right) \left(\sum_{i=1}^{M} \frac{1}{c_i} \right), \text{ where } c_i = \left(\frac{1-\alpha_i}{e_i^* + \gamma_i p_i} \right). \tag{18}$$ [12]

The integration in Equation (17) also collapses to a closed form expression (see Appendix B), providing the following overall expression:[13]

$$P\left(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0\right)$$

$$= \frac{1}{\sigma^{M-1}} \left[\prod_{i=1}^{M} c_i \right] \left[\sum_{i=1}^{M} \frac{1}{c_i} \right] \left[\frac{\prod_{i=1}^{M} e^{V_i/\sigma}}{\left(\sum_{k=1}^{K} e^{V_k/\sigma} \right)^M} \right] (M-1)! \tag{19}$$

In the case when $M = 1$ (*i.e.*, only one alternative is chosen), there are no satiation effects ($\alpha_k = 1$ 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 (19) collapses to the standard MNL model. Thus, the MDCEV model is a multiple discrete-continuous extension of the standard MNL model.[14]

[12] It is important to note that this compact Jacobian form is independent of the assumptions regarding the density and correlation structure of the error terms.

[13] One can also derive the expression below from the difference form of Equation (13), using the properties of the multivariate logistic distribution (see Appendix C).

[14] Note that when $\alpha_k = 1$ for all k, $V_k = \beta' z_k - \ln p_k$. Even if $M = 1$, when Equation (19) collapses to the MNL form, the scale σ is estimable as long as the utility takes the functional form $V_k = \beta' z_k - \ln p_k$ and there is price variation across goods. This is because the scale is the inverse of the coefficient on the $\ln p_k$ term (see Hanemann, 1984).

The expression for the probability of the consumption pattern of the goods (rather than the expenditure pattern) can be derived to be:

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

$$= \frac{1}{p_1} \cdot \frac{1}{\sigma^{M-1}} \left[\prod_{i=1}^{M} f_i \right] \left[\sum_{i=1}^{M} \frac{p_i}{f_i} \right] \left[\frac{\prod_{i=1}^{M} e^{V_i / \sigma}}{\left(\sum_{k=1}^{K} e^{V_k / \sigma} \right)^M} \right] (M-1)!, \tag{20}$$

where V_k is as defined earlier (see Equation 16) and $f_i = \left(\frac{1-\alpha_i}{x_i^* + \gamma_i} \right)$. The expression in Equation (20) is, however, not independent 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). 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, but it is still awkward to have different probability values for the same consumption pattern. This is particularly the case because different log-likelihood values at convergence will be obtained for different designations of the first good. Thus, the preferred approach is to use the probability expression for expenditure allocations, which will provide the same probability for a given expenditure pattern regardless of the good labeled as the first good. However, in the case that the first good is an outside numeraire good that is always consumed (see Section 5), then $p_1 = 1$ and one can use the consumption pattern probability expression or the expenditure allocation probability expression.

4.2 The Multiple Discrete-Continuous Generalized Extreme-Value (MDCGEV) Model Structure

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 alternatives using a generalized extreme value (GEV) error structure. The remarkable advantage of the GEV structure is that it continues to result in closed-form probability expressions for any and all expenditure patterns. However, the derivation is tedious, and the expressions get unwieldy. In this paper, we provide the expressions for a specific nested logit structure with 4 alternatives, two alternatives (labeled 1 and 2) in nest A and the other two alternatives (labeled 3 and 4) in nest B (the derivation is available on request from the author). The cumulative distribution function for the error terms in the utility expressions take the following form:

$$\Lambda(\varepsilon_1 < s_1, \varepsilon_2 < s_2, \varepsilon_3 < s_3, \varepsilon_4 < s_4) = \exp\left[-\left(e^{-s_1/\sigma\theta_A} + e^{-s_2/\sigma\theta_A}\right)^{\theta_A} - \left(e^{-s_3/\sigma\theta_B} + e^{-s_4/\sigma\theta_B}\right)^{\theta_B}\right] \quad (21)$$

Define the following:

$$A_{12} = \left(e^{V_1/\sigma\theta_A} + e^{V_2/\sigma\theta_A}\right),$$

$$B_{34} = \left(e^{V_3/\sigma\theta_B} + e^{V_4/\sigma\theta_B}\right),$$

$$H = A_{12}^{\theta_A} + B_{34}^{\theta_B}.$$

Then,

$$P(x_1^*,0,0,0) = \frac{A_{12}^{\theta_A-1} \cdot e^{V_1/\sigma\theta_A}}{H}$$

$$P(x_1^*,x_2^*,0,0) = \frac{1}{\sigma}|J|\frac{e^{V_1/\sigma\theta_A} \cdot e^{V_2/\sigma\theta_A} \cdot A_{12}^{\theta_A-2}}{H}\left(\frac{A_{12}^{\theta_A}}{H} + \frac{1-\theta_A}{\theta_A}\right)$$

$$P(x_1^*,0,x_3^*,0) = \frac{1}{\sigma}|J|\frac{e^{V_1/\sigma\theta_A} \cdot e^{V_3/\sigma\theta_B} \cdot A_{12}^{\theta_A-1}B_{34}^{\theta_B-1}}{H^2}$$

$$P(x_1^*,x_2^*,x_3^*,0) = \frac{1}{\sigma^2}|J|\frac{e^{V_1/\sigma\theta_A} \cdot e^{V_2/\sigma\theta_A} \cdot e^{V_3/\sigma\theta_B} \cdot A_{12}^{\theta_A-2}B_{34}^{\theta_B-1}}{H^2}\left(\frac{2A_{12}^{\theta_A}}{H} + \frac{1-\theta_A}{\theta_A}\right)$$

$$P(x_1^*,x_2^*,x_3^*,x_4^*) = \frac{1}{\sigma^3}|J|\frac{e^{V_1/\sigma\theta_A} \cdot e^{V_2/\sigma\theta_A} \cdot e^{V_3/\sigma\theta_B} \cdot e^{V_4/\sigma\theta_B} \cdot A_{12}^{\theta_A-2}B_{34}^{\theta_B-1}}{H^2}$$
$$\left\{\frac{6A_{12}^{\theta_A}B_{34}^{\theta_B}}{H^2} + \frac{2(1-\theta_A)}{\theta_A}\cdot\frac{B_{34}^{\theta_B}}{H} + \frac{2(1-\theta_B)}{\theta_B}\cdot\frac{A_{12}^{\theta_A}}{H} + \left(\frac{(1-\theta_A)}{\theta_A}\right)\left(\frac{(1-\theta_B)}{\theta_B}\right)\right\}$$

The probabilities for all other expenditure patterns for the 4 goods can be obtained by interchanging the labels 1, 2, 3, and 4.[15]

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

[15] In all the expressions corresponding to the nested structure above, σ is identified only when there is price variation across alternatives (see Section 3.2).

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, $\eta \sim N(0, \Omega)$.[16]

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:

$$P\left(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0\right)$$

$$= \int_{\eta} \frac{1}{\sigma^{M-1}} \left[\prod_{i=1}^{M} c_i\right] \left[\sum_{i=1}^{M} \frac{1}{c_i}\right] \left[\frac{\prod_{i=1}^{M} e^{(V_i+\eta_i)/\sigma}}{\left(\sum_{k=1}^{K} e^{(V_k+\eta_k)/\sigma}\right)^M}\right] (M-1)!\, dF(\eta). \tag{22}$$

where F is the multivariate cumulative normal distribution (see Bhat, 2005; Bhat and Sen, 2006; and Bhat *et al.*, 2006).

The model in Equation (22) 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).

4.3.1 Heteroscedastic structure within the MMDCEV framework

Consider the case where there is price variation across the alternatives, and the overall errors ε_k are heteroscedastic, but not correlated. Assuming a 4-alternative case for ease in presentation, the heteroscedastic structure may be specified in the form of the following covariance matrix for $\varepsilon = (\varepsilon_{k1}, \varepsilon_{k2}, \varepsilon_{k3}, \varepsilon_{k4})$:

---

[16] 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 (see Bhat, 2007).

$$Cov(\varepsilon) = \frac{\pi^2\sigma^2}{6}\begin{bmatrix} 1 & 0 & 0 & 0 \\ & 1 & 0 & 0 \\ & & 1 & 0 \\ & & & 1 \end{bmatrix} + \begin{bmatrix} \omega_1^2 & 0 & 0 & 0 \\ & \omega_2^2 & 0 & 0 \\ & & \omega_3^2 & 0 \\ & & & \omega_4^2 \end{bmatrix}, \tag{23}$$

where the first component on the right side corresponds to the IID covariance matrix of $\zeta = (\zeta_1, \zeta_2, \zeta_3, \zeta_4)$ and the second component is the heteroscedastic covariance matrix of η. The covariance of error differences with respect to the first alternative is:

$$Cov(\tilde{\varepsilon}_1 = \tilde{\varepsilon}_{21}, \tilde{\varepsilon}_{31}, \tilde{\varepsilon}_{41}) = \frac{\pi^2\sigma^2}{6}\begin{bmatrix} 2 & 1 & 1 \\ & 2 & 1 \\ & & 2 \end{bmatrix} + \begin{bmatrix} \omega_1^2 + \omega_2^2 & \omega_1^2 & \omega_1^2 \\ & \omega_1^2 + \omega_3^2 & \omega_1^2 \\ & & \omega_1^2 + \omega_4^2 \end{bmatrix} \tag{24}$$

An inspection of the matrix above shows only four independent equations (the rank condition), implying that at most four parameters are estimable[17]. There are two ways to proceed with a normalization, as discussed below.

The first approach is to normalize σ and estimate the heteroscedastic covariance matrix of η (*i.e.*, ω_1, ω_2, ω_3, and ω_4). Assume that σ is normalized to $\tilde{\sigma}$, and let the corresponding values of ω_k be $\tilde{\omega}_k$ (k = 1, 2, 3, 4). Then, the following equalities should hold, based on Equation (24), for any normalization of σ to $\tilde{\sigma}$ ($q = \pi^2 / 6$ below):

$$\begin{aligned} &\omega_1^2 + q\sigma^2 = \tilde{\omega}_1^2 + q\tilde{\sigma}^2 \\ &\omega_1^2 + \omega_k^2 + 2q\sigma^2 = \tilde{\omega}_1^2 + \tilde{\omega}_k^2 + 2q\tilde{\sigma}^2 \quad (k = 2, 3, 4) \end{aligned} \tag{25}$$

The above equalities can be rewritten as:

$$\tilde{\omega}_k^2 = \omega_k^2 + q\sigma^2 - q\tilde{\sigma}^2 \quad (k = 1, 2, 3, 4) \tag{26}$$

The normalized variance terms $\tilde{\omega}_k^2$ must be greater than or equal to zero, which implies that the following conditions should hold:

$$\omega_k^2 + q\sigma^2 \geq q\tilde{\sigma}^2 \quad (k = 1, 2, 3, 4) \tag{27}$$

Intuitively, the above condition implies that the normalization on $\tilde{\sigma}$ must be set low enough so that the overall "true" variance of each error term $(= \omega_k^2 + q\sigma^2)$ is larger than $q\tilde{\sigma}^2$. For example, setting σ to 1 would be inappropriate if the "true" variance of one or more alternatives is less

[17] Strictly speaking, one can estimate all the five parameters (σ, ω_1, ω_2, ω_3, and ω_4) because of the difference in the extreme value distributions used for ζ_k and the normal distributions used for η_k (see Walker, 2002). However, the model will be near singular, and it is important to place the order/rank constraint.

than $\pi^2/6$. Since the "true" variance is unknown, the best the analyst can do is to normalize σ to progressively smaller values and statistically examine the results.

The second approach is to normalize one of the ω_k terms instead of the σ term. In this case, from Equation (24), we can write:

$$q\tilde{\sigma}^2 = \omega_1^2 + q\sigma^2 - \tilde{\omega}_1^2 = \frac{1}{2}\left[\omega_1^2 + \omega_k^2 + 2q\sigma^2 - \tilde{\omega}_1^2 - \tilde{\omega}_k^2\right];\ k = 2, 3, 4. \tag{28}$$

After some manipulations, the above equation may be rewritten as:

$$\tilde{\omega}_k^2 = \omega_k^2 + \tilde{\omega}_1^2 - \omega_1^2;\ k = 2, 3, 4. \tag{29}$$

Next, imposing the condition that the normalized terms $\tilde{\omega}_k^2$ must be greater than or equal to zero implies the following:

$$\tilde{\omega}_1^2 \geq \omega_1^2 - \omega_k^2 \ \ (k = 2, 3, 4). \tag{30}$$

The above condition is automatically satisfied as long as the first alternative is the minimum variance alternative. An associated convenient normalization is $\tilde{\omega}_1^2 = 0$, since the resulting model nests the MDCEV model. The minimum variance alternative can be determined by estimating an unidentified model with all the k ω_k terms, and identifying the alternative with the minimum variance (see Walker *et al.*, 2004, for an equivalent procedure for a heteroscedastic specification within the mixed multinomial logit model).

The above discussion assumes there is price variation across goods. In the case of no price variation, the scale σ is not identifiable. In this case, the easiest procedure is to normalize σ to 1 and the ω_k^2 value for the minimum variance alternative *k* to zero.

4.3.2 The general error covariance structure within the MMDCEV framework

Appropriate identification normalizations will have to placed on σ and the covariance matrix of η when the analyst is estimating an error-components structure to allow correlation in unobserved factors influencing the baseline utility of alternatives, since only a (*K*-1)x(*K*-1) covariance of error differences is identified. This can be accomplished by imposing a structure based on *a priori* beliefs or intuitive considerations. However, the analyst must ensure that the elements of the assumed restricted covariance structure can be recovered from the (*K*-1)x(*K*-1) covariance of error differences that is actually estimable.

In the most general error covariance structure, and when there is price variation, one way to achieve identification is the following: (1) Normalize the scale parameter σ to be a small value such that the variance of the minimum variance alternative exceeds $\pi^2\sigma^2/6$ (since this variance is not known, the analyst will have to experiment with alternative fixed σ values), (2) Normalize ω_k for the minimum variance alternative *k* to zero, and (3) Normalize all correlations of this minimum variance alternative with other alternatives to zero. Together, these normalizations leave only $K(K-1)/2$ parameters to be estimated, and are adequate for identification. In the case of no price variation, an additional restriction will have to be imposed. One approach would be to set $\sum_{k=2}^{K}\omega_k^2 = 1$ to set the scale in the covariance matrix of η.

5. THE MODEL 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. Also, for identification, let $\psi(x_1,\varepsilon_1) = e^{\varepsilon_1}$. Then, the utility functional form needs to be modified as follows:

$$U(\boldsymbol{x}) = \frac{1}{\alpha_1}\exp(\varepsilon_1)x_1^{\alpha_1} + \sum_{k=2}^{K}\frac{\gamma_k}{\alpha_k}\exp(\beta' z_k + \varepsilon_k)\left\{\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right\} \tag{31}$$

Note that there is no translation parameter γ_1 for the first good, because the first good is always consumed. As in the "no-outside good" case, the analyst will generally not be able to estimate both α_k and γ_k for the inside goods 2, 3, …, *K*. The analyst can estimate one of the following three utility forms:

$$U(\boldsymbol{x}) = \frac{1}{\alpha_1}\exp(\varepsilon_1)x_1^{\alpha_1} + \sum_{k=2}^{K}\frac{1}{\alpha_k}\exp(\beta' z_k + \varepsilon_k)\left\{(x_k+1)^{\alpha_k} - 1\right\}$$

$$U(\boldsymbol{x}) = \frac{1}{\alpha_1}\exp(\varepsilon_1)x_1^{\alpha_1} + \sum_{k=2}^{K}\gamma_k\exp(\beta' z_k + \varepsilon_k)\ln\left(\frac{x_k}{\gamma_k}+1\right) \tag{32}$$

$$U(\boldsymbol{x}) = \frac{1}{\alpha}\exp(\varepsilon_1)x_1^{\alpha} + \sum_{k=2}^{K}\frac{\gamma_k}{\alpha}\exp(\beta' z_k + \varepsilon_k)\left\{\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha} - 1\right\}$$

The last functional form above is estimable now because the constant α parameter is obtaining a "pinning effect" from the satiation parameter for the outside good. The analyst can estimate all the three possible functional forms and select the one that fits the data best based on statistical and intuitive considerations. The identification considerations discussed for the "no-outside good" case carries over to the "with outside good" case. The probability expression for the expenditure allocation on the various goods (with the first good being the outside good) is identical to Equation (19), while the probability expression for consumption of the goods (with the first good being the outside good) is

$$P\left(x_1^*, x_2^*, x_3^*, ..., x_M^*, 0, 0, ..., 0\right)$$
$$= \frac{1}{\sigma^{M-1}}\left[\prod_{i=1}^{M} f_i\right]\left[\sum_{i=1}^{M} \frac{p_i}{f_i}\right]\left[\frac{\prod_{i=1}^{M} e^{V_i/\sigma}}{\left(\sum_{k=1}^{K} e^{V_k/\sigma}\right)^M}\right](M-1)!, \tag{33}$$

where $f_i = \left(\dfrac{1-\alpha_i}{x_i^* + \gamma_i}\right)$.[18]

The expressions for V in Equation (19) and Equation (33) are as follows for each of the three utility forms in Equation (32):

First form - $V_k = \beta' z_k + (\alpha_k - 1)\ln(x_k^* + 1) - \ln p_k \ (k \geq 2);\ V_1 = (\alpha_1 - 1)\ln(x_1^*)$

Second form - $V_k = \beta' z_k - \ln(\dfrac{x_k^*}{\gamma_k} + 1) - \ln p_k \ (k \geq 2);\ V_1 = (\alpha_1 - 1)\ln(x_1^*)$ (34)

Third form - $V_k = \beta' z_k + (\alpha - 1)\ln(\dfrac{x_k^*}{\gamma_k} + 1) - \ln p_k \ (k \geq 2);\ V_1 = (\alpha - 1)\ln(x_1^*)$

6. COMPARISON WITH EARLIER MULTIPLE DISCRETE-CONTINUOUS MODELS

In this section, we discuss how the model developed in this paper differs from the model of Kim *et al.* (2002), those in Environmental Economics, and the earlier models by Bhat and colleagues. The discussion is in the context of the basic structure with identically and independently distributed error terms across alternatives.

[18] The Gauss code, documentation, and test data sets for estimating the MDCEV model (with and without an outside good) are available at: http://www.caee.utexas.edu/prof/bhat/MDCEV.html

6.1 Kim *et al.*'s Model

Kim *et al.* (2002) use the following translated constant elasticity of substitution (CES) direct utility form:

$$U = \sum_{k=1}^{K} \psi_k^1 (x_k + \gamma_k^1)^{\alpha_k^1}, \tag{35}$$

where the superscript '1' is to distinguish the parameters in this functional form from those in the form of Equation (1). In this section, for ease in comparison, we will consider the case when there is no outside good. In the utility form above, $\psi_k^1 > 0$, $\gamma_k^1 > 0$, and $0 \le \alpha_k^1 \le 1$. To empirically identify the utility form, Kim *et al.* impose the restriction that $\gamma_k^1 = 1$ for all *k* (*i.e.*, they estimate the α -profile). The reader will note that Kim *et al.*'s form does not incorporate the weak complementarity assumption. However, this can be easily remedied by revising the utility form above to an empirically indistinguishable alternative form provided below:

$$U = \sum_{k=1}^{K} \psi_k^1 \left\{ (x_k + \gamma_k^1)^{\alpha_k^1} - (\gamma_k^1)^{\alpha^1} \right\}, \tag{36}$$

The KT conditions and optimal consumptions for both Equations (35) and (36) are identical. But the latter form assigns zero utility to good *k* when it is not consumed, while still allowing corner solutions. However, in either of the two forms, the interpretation of ψ_k^1 is not straightforward. Specifically, the baseline marginal utility of a good (or marginal utility when no quantity of the good is consumed) is $\psi_k^1 \alpha_k^1$, which depends on both ψ_k^1 and α_k^1 (for γ_k^1 fixed to 1 for all *k*). Of course, the satiation rate for good *k* with respect to the baseline marginal utility is still determined by α_k^1, as in the α -profile based utility form adopted in this paper (*i.e.*, Equation (1) with γ_k fixed to 1 for all *k*). In fact, the estimation results and optimal consumptions from using Equation (1) and Equation (35) (with all γ_k 's fixed to 1) will be identical, as long as $\alpha_k \ge 0$ for all *k* in Equation (1). The only cosmetic difference will be a shift in the constant terms between the ψ_k and ψ_k^1 terms. Specifically, if $\psi_k = \exp(\tau_k + \beta' z_k + \varepsilon_k)$ and $\psi_k^1 = \exp(\tau_k^1 + \beta^{1\prime} z_k + \varepsilon_k^1)$, where the constant term is removed out from the $\beta' z_k$ term, the following relationship will hold between the two models as long as the same error distributions are assumed (and assuming that the first alternative is considered as the base):

$$\tau_k = \tau_k^1 + \ln\left(\frac{\alpha_k^1}{\alpha_1^1}\right) \forall\, k$$

$$\alpha_k = \alpha_k^1 \;\forall\, k \tag{37}$$

$$\beta = \beta^1 \,.$$

An important technical nuance is in order here. The form of Equation (35) is restrictive compared to the form of Equation (1) adopted in this paper. Specifically, Equation (1) covers the overall baseline marginal utility satiation space more completely than Equation (35) because $\alpha_k \le 1$ in Equation (1), while $0 \le \alpha_k^1 \le 1$ in Equation (35). But the analyst, at times, may have to impose the constraint $\alpha_k \ge 0$ in the functional form of Equation (1) to provide stability in estimation, especially when a scale parameter is being estimated with limited price variation.

Another important difference between Kim *et al.*'s model and the model here is the distribution of the error terms. Kim *et al.* assume that the error terms are independent and identically distributed normal. They then use the differencing form of Equation (13) to develop the probabilities, with the *g*(.) function being a multivariate normal density. This form requires an appropriate decomposition of the density function for the continuous and discrete components, and multivariate normal integration. The approach is not practical for most realistic applications. As Bhat (2005) noted, if one considers the error terms to be IID gumbel instead of IID normal, the model structure collapses to the closed-form MDCEV form used here.

6.2 Models in Environmental Economics

The studies in Environmental Economics, unlike Kim *et al.*, use the utility function corresponding to the γ_k-profile. These studies always consider the presence of an outside good, and so we will consider the case when there is an outside good in this section. This outside good may be arbitrarily designated as the first good in our notational framework with $p_1 = 1$ (as in Section 5).

The utility function in the Environmental Economics studies takes the LES form (see von Haefen and Phaneuf (2005), von Haefen (2003b), Phaneuf *et al.* (2000), Phaneuf and Herriges (2000), and Herriges *et al.* (2004):

$$U = \sum_{k=1}^{K} \psi_k^2 \ln(x_k + \gamma_k^2) \;. \tag{38}$$

The superscript '2' above is to distinguish the parameters from those in Equation (1), and should not be confused with the square power function. In the function above, the utility accrued from zero consumption of a good is positive since $\psi_k^2 > 0$ and $\gamma_k^2 > 0$. However, this can be accommodated by re-writing the utility form in the empirically indistinguishable alternative form shown below[19]:

$$U = \sum_{k=1}^{K} \psi_k^2 \ln(\frac{x_k}{\gamma_k^2} + 1) \ , \tag{39}$$

In particular, it can be readily seen that the KT first order conditions and the optimal consumptions are identical for the utility forms in Equations (38) and (39), as also observed by Herriges *et al.* (2004). However, in both forms, the interpretation of ψ_k^2 is not straightforward, since the baseline marginal utility is ψ_k^2 / γ_k^2. But, for a given baseline marginal utility, γ_k^2 serves as a satiation parameter (in addition to allowing corner solutions). In fact, the estimation results and optimal consumptions from using Equation (2) and Equation (39) will be identical, except for a shift in the constant terms between the ψ_k and ψ_k^2 terms. Specifically, if $\psi_k = \exp(\tau_k + \beta' z_k + \varepsilon_k)$ and $\psi_k^2 = \exp(\tau_k^2 + \beta^{2}{}' z_k + \varepsilon_k^2)$, the following relationship will hold between the two models as long as the same error distributions are assumed (and assuming that the first alternative is considered as the base):

$$\begin{aligned} &\tau_k = \tau_k^2 + \ln\left(\frac{\gamma_1^2}{\gamma_k^2}\right) \forall\, k \\ &\gamma_k = \gamma_k^2 \ \forall\, k \\ &\beta = \beta^2 \end{aligned} \tag{40}$$

An important point to note about the Environmental Economics studies is that they consider the utility of the outside good as being deterministic (*i.e.*, ε_1 = 0), and then consider the error terms of the utilities of the inside goods to be independent and (typically) identically extreme value distributed. To see this, consider Equation (17), which is the appropriate probability expression for the expenditure pattern if there are independent and identically

[19] von Haefen *et al.* (2004) and von Haefen and Phaneuf (2003) also recognize this weak complementarity problem in the functional form of Equation (38), where the quality attributes of good *k* contribute to utility even if the good is not consumed. They address it by interacting the quality attributes with x_k, rather than including the quality attributes in ψ_k.

distributed error terms in all utilities (with the first alternative being the outside good). The equivalent probability expression for the consumption pattern is:

$$P\left(x_1^*, x_2^*, x_3^*, ..., x_M^*, 0, 0, ..., 0\right)$$
$$= |J|_x \int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty} \left\{ \left(\prod_{i=2}^{M} \frac{1}{\sigma} \lambda\left[\frac{V_1 - V_i + \varepsilon_1}{\sigma}\right] \right) \right\} \times \left\{ \prod_{s=M+1}^{K} \Lambda\left[\frac{V_1 - V_s + \varepsilon_1}{\sigma}\right] \right\} \frac{1}{\sigma} \lambda\left(\frac{\varepsilon_1}{\sigma}\right) d\varepsilon_1 , \tag{41}$$

where $|J|_x = \left(\prod_{i=1}^{M} f_i \right) \left(\sum_{i=1}^{M} \frac{p_i}{f_i} \right)$, $p_1 = 1$, and $f_i = \left(\frac{1-\alpha_i}{x_i^* + \gamma_i} \right)$. The values of α_i and γ_i in f_i will depend on the utility function form used in the presence of an outside good. For the first functional form in Equation (32), $\gamma_1 = 0$ and $\gamma_i = 1$ for all $i \neq 1$. For the second functional form in Equation (32), $\gamma_1 = 0$ and $\alpha_i = 0$ for all $i \neq 1$. For the final functional form, $\gamma_1 = 0$ and all α values are equal across alternatives.

The expression in Equation (41) collapses to the closed form expression provided in Equation (33), yielding the MDCEV model. But now assume $\varepsilon_1 = 0$ in Equation (41) as in the Environmental Economics studies. The integral in Equation (41) then drops out, and the equation becomes:

$$P\left(x_1^*, x_2^*, x_3^*, ..., x_M^*, 0, 0, ..., 0\right)$$
$$= |J|_x \left\{ \left(\prod_{i=2}^{M} \frac{1}{\sigma} \lambda\left[\frac{V_1 - V_i}{\sigma}\right] \right) \right\} \times \left\{ \prod_{s=M+1}^{K} \Lambda\left[\frac{V_1 - V_s}{\sigma}\right] \right\}. \tag{42}$$

Substituting $\lambda(w) = e^{-e^{-w}} \cdot e^{-w}$ and $\Lambda(w) = e^{-e^{-w}}$, the expression may be written as:

$$P\left(x_1^*, x_2^*, x_3^*, ..., x_M^*, 0, 0, ..., 0\right)$$
$$= |J|_x \left\{ \prod_{i=2}^{M} \frac{\exp(-g_i/\sigma)}{\sigma} \right\} \left\{ \exp\left(-\sum_{k=2}^{K} \exp(-g_k/\sigma) \right) \right\}, \tag{43}$$

where $g_k = V_1 - V_k$. The form above is the same as the likelihood function in Equation (10) of Von Haefen and Phaneuf (2003). Thus, the models in Environmental Economics are obtained by assuming away stochasticity in the utility of the outside good. Basically, the Environmental Economics studies recognize the singularity imposed by the budget constraint by directly assuming $\varepsilon_1 = 0$, so that there are only (K-1) error terms in the (K-1) KT conditions (see Equation 10). The MDCEV model, on the other hand, recognizes the singularity imposed by the budget constraint by considering all utilities as random, and then explicitly acknowledging the singularity among the K error terms in the (K-1) KT conditions (see also Kim *et al.*, 2002 and Wales and Woodland, 1983, who use the latter approach). The latter approach is conceptually consistent in considering the utilities of all alternatives as being random (strict random utility maximization), while the former approach assumes that the analyst knows all consumer-related and market-related factors going into the valuation of the outside good, but not for the inside goods (partial random utility maximization). Further, in the Environmental Economics approach, if instead of the outside good's utility, the utility of some other inside good is considered deterministic to accommodate the singularity, we obtain different probability expressions and probability values for the same consumption pattern. Specifically, if the error term of alternative l is fixed to zero where $l \leq M$, the probability expression for the consumption pattern corresponding to Equation (43) is:

$$P\left(x_1^*, x_2^*, x_3^*, ..., x_M^*, 0, 0, ..., 0\right)$$
$$= \frac{1}{p_l} |J|_x \{\exp(+g_l/\sigma)\}^{M-1} \left\{ \prod_{\substack{i=1 \\ i \neq l}}^{M} \frac{\exp(-g_i/\sigma)}{\sigma} \right\} \left\{ \exp\left(-\sum_{\substack{k=1 \\ k \neq l}}^{K} \exp(-g_k/\sigma) \right) \right\}^{\exp(+g_l/\sigma)}, \tag{44}$$

where p_l is the price of the l^{th} good. On the other hand, if the error term of alternative l is fixed to zero where $l > M$, the probability expression for the consumption pattern is:

$$P\left(x_1^*, x_2^*, x_3^*, ..., x_M^*, 0, 0, ..., 0\right)$$
$$= |J|_x \left\{ \prod_{i=2}^{M} \frac{\exp(-g_i/\sigma)}{\sigma} \right\} \left\{ \sum_{\substack{k=1 \\ k \neq l}}^{K} \exp(-g_k/\sigma) \right\}^{-M} \tag{45}$$
$$\left\{(M-1)! - \exp(d_l)(-1)^{M+1}\left(d_l^{M-1} - (M-1)d_l^{M-2} + (M-1)(M-2)d_l^{M-3} + ...(-1)^{M-1}(M-1)!\right)\right\},$$

where $d_l = -\left\{ \sum_{\substack{k=1 \\ k \neq l}}^{K} \exp(-g_k/\sigma) \right\} \exp(+g_l/\sigma)$.[20]

As one can observe from Equation (43), (44), and (45), the probability expressions are quite different based on which alternative's error term is fixed to zero. To see this even more clearly, consider the specific case of 3 goods with the first good being the outside good. Then, the

[20] The derivations of Equations (44) and (45) are tedious, and are available from the author.

probability expression for only the outside good being consumed if the outside good's utility is fixed is (from Equation 43 with $M = 1$):[21]

$$P^1(x_1^*,0,0) = \exp\left[-\sum_{k=2}^{3} \exp(-g_k/\sigma)\right], \tag{46}$$

where $g_k = V_1 - V_k$ as earlier. The corresponding expression if the second good's utility is fixed is (from Equation 45 with $M = 1$):

$$P^2(x_1^*,0,0) = \{1+\exp(-g_3/\sigma)\}^{-1}\{1-\exp[-\{1+\exp(-g_3/\sigma)\}\{\exp(+g_2/\sigma)\}]\}. \tag{47}$$

The superscript '2' above is to distinguish from the expression in Equation (46), and should not be confused with the square power function. Clearly, Equations (46) and (47) are different. This is diagrammatically shown in Figure 5, which plots $P^1(x_1^*,0,0)$ and $P^2(x_1^*,0,0)$ for different values of $g_2 = g_3 = g$ and $\sigma = 1$. As can be seen, $P^2(x_1^*,0,0) > P^1(x_1^*,0,0)$. Thus, one gets different probability profiles depending on which alternative's utility is considered fixed, and there is no obvious reason to fix the outside good's utility rather than an inside good's utility. In fact, this issue of considering which alternative's utility to fix becomes particularly apparent when there is more than one outside good defined in a certain empirical context, or when there is no outside good (*i.e.*, the choice of only the inside goods is modeled). On the other hand, the probability in our approach that includes error terms in all alternatives is:

$$P^3(x_1^*,0,0) = \frac{\exp(V_1/\sigma)}{\sum_{k=1}^{3} \exp(V_k/\sigma)} = \frac{1}{1+\exp(-g_2/\sigma)+\exp(-g_3/\sigma)}, \tag{48}$$

which is, of course, the simple multinomial logit model (the corresponding Environmental Economics-based formula of Equation (46) does not collapse to the multinomial logit). The profile of Equation (48) is also drawn in Figure 1, and it lies between the two earlier profiles.

Another point to note is that the Environmental Economics studies (and the Kim *et al.* study) do not use the compact and simple structure of the Jacobian we have derived. The Jacobian expression $|J|_x$ under Equation (41) is equivalent to the tedious expansion formulas for the Jacobian used in the Environmental Economics studies. As a result of these tedious Jacobian expressions, the Environmental Economics studies employ a numerical gradient

[21] Note that the Jacobian term drops out in the expressions below because the probabilities are being computed for the case where only the outside good is consumed.

method in the likelihood estimation, which is less precise than the simple form of the analytic gradient that can be obtained by writing the Jacobian as in our approach. Further, while the numerical gradients may provide accurate estimates at acceptable speeds for the simple model form with IID error terms, the approach is extremely slow (by a large order of magnitude) compared to the analytic gradient approach for the case when random coefficients or richer error correlation patterns are introduced using a mixing approach.

6.3 Bhat's Earlier Models

The models of Bhat (2005), Bhat and Sen (2006), and Bhat *et al.* (2006) use the translated CES direct utility function form as in Kim *et al.* (2002), rather than the more easy-to-interpret and general utility form used in the current paper. The studies assume independent and identically distributed gumbel error kernel terms that leads to the MDCEV form and its mixed variants, as in the current paper. The studies assume unit "prices" for all the alternatives, and so do not have to deal with the many issues arising from price variation as discussed in this paper. These earlier efforts also do not address identification considerations, nor do they shed light on the role played by the parameters in the utility function.

7. EMPIRICAL ILLUSTRATIONS

7.1 Absence of Outside Good Case

In this section, we supplement and demonstrate the scaling and identification issues associated with the case of no price variation and price variation (as discussed in earlier sections) using empirical examples for the case of the absence of an outside good. The section also shows the equivalency in parameters when using the consumption pattern formulation and the expenditure pattern formulation for the case of price variation across goods.

The data set used for these illustrations is the same as that used by Bhat and Sen (2006), and is drawn from the 2000 San Francisco Bay Area Travel Survey (BATS). The sample includes 3500 households with non-zero vehicle ownership. The vehicles owned by each household are categorized into one of five vehicle types based on their make, model and year. The five vehicle types are (1) Passenger car, (2) Sports Utility Vehicle (SUV), (3) Pickup truck, (4) Minivan, and (5) Van. In this paper, we estimate different MDCEV specifications with these five alternatives. Households may hold a combination of these vehicle types and use different vehicle types in different ways, leading to the discrete-continuous choice of vehicle holding mix and vehicle miles of travel by vehicle type.

In all the results to be discussed, we used the most basic specification with constants in the baseline marginal utility and satiation parameters. This is because the motivation here is to discuss scaling and identification issues, not variable specification considerations. We present the baseline marginal utility parameters in their parameterized form because this allows us to show the equivalence between models in a straightforward manner. That is, we present the β_k parameters, where $\psi_k = \exp(\beta_k)$ for all *k*.

7.1.1 Case of no price variation

Consider the situation where the right side of the "budget" constraint is simply the total annual miles of travel across all vehicle types. This is the case of prices not appearing in the budget constraint (or equivalently, the cost of traveling a mile by each vehicle being unity).

Table 1 shows six estimations, the first four being based on the α-profile and the last two being based on the γ-profile (as expected, we were unable to estimate a model with both the α parameters and the γ parameters for each vehicle type). The first two estimations fix the scale parameter and estimate all the α satiation parameters. It can be observed from the log-likelihood values at convergence (see last row of the table) that whether one fixes the scale to the value of '1' or '2' does not matter. As discussed in Section 3.2, the β parameters in Model 2 are scaled by a factor of 2 and the α parameters in model 2 are related to those in Model 1 by the relationship $\alpha_{\text{model 2}} = 2*(\alpha_{\text{model 1}} - 1) + 1$. That is, the error term scale is unidentified when one estimates all the α parameters.

The third and fourth models apply a different normalization where one of the α parameters is fixed to the value of '0'. In these cases, one can estimate the error term scale. However, these models are no different from Models 1 and 2. Specifically, the β parameters in Models 3 and 4 are equal to the estimated scale parameters in these models times the β parameters in Model 1. The α parameters in Models 3 and 4 are related to that in Model 1 by $\alpha_{\text{model d}} = \sigma_d * (\alpha_{\text{model 1}} - 1) + 1$ ($d = 3,4$).

The fifth and sixth models correspond to the γ profile with a fixed α. The fifth model constrains α to 0 and the scale σ to 1. The log-likelihood at convergence for this model indicates that the γ profile provides a better fit in this empirical case than does the α profile. The sixth model sets the scale value to 2 rather than 1. This sixth model provides the same results as the fifth model because the α parameter has been set to $\alpha_{\text{model}\,6} = \sigma_{\text{model}\,6} * (\alpha_{\text{model}\,5} - 1) + 1 = -1$. Thus, in estimating the γ-profile, the scale can be set to any value as long as the α parameter is fixed appropriately. The normalization used in Model 5 is most convenient in this case.

7.1.2 Case of price variation

The data assembled by Bhat and Sen does not include a composite per mile usage price measure for each vehicle type because of the wide variation in the fixed costs of vehicle purchase and use costs per mile within each broad vehicle type category. For the analysis here, we used synthetic per mile price data by generating uniform random numbers with a mean of 1,2,2,3, and 4 for sedans, sports utility vehicles (SUV), pick-up trucks, minivans, and vans, respectively. The intent here is simply to demonstrate the issues discussed earlier in the context of price variation, rather than to develop an actual model for vehicle type choice and use.

Table 2 shows the results of six model estimations. The first two estimations are α-profiles based on the consumption probability expression of Equation (20). In both these estimations, we had to restrict the α parameters to be between 0 and 1 for convergence considerations. The only difference between these first two estimations is that the good labeled as the first good is different (note that any good that is consumed may be designated as the first good). As can be observed from the results, there is effectively no difference in the parameter estimates at convergence (the minor differences are a result of the optimization convergence process). However, the log-likelihood values at convergence are different, because the probability assigned to the same consumption pattern varies based on the vehicle type designated as the first good.

The third model is also an α-profile specification (with the α's constrained to between 0 and 1), but based on the expenditure probability expression of Equation (19). The results of this third model are identical to those in the first two models in terms of parameter estimates.

However, this model, by structure, is independent of the alternative that is designated as the first good.

The remaining models in Table 2 are all based on the expenditure probability expression. The fourth model, also an α-profile specification, restricts the scale parameter to one. A comparison of this model with the third model in terms of a nested likelihood ratio test yields a test statistic of 603.74, which rejects the null hypothesis of a unit scale parameter at any reasonable level of statistical significance. This is a clear indication that the scale parameter is estimable when there is price variation.

The last two models in Table 2 correspond to the γ-profile, with the sixth model restricting the scale to one. A comparison of these last two models with one another using a nested likelihood ratio test again clearly rejects the null hypothesis that the scale value is immaterial. We also estimated another model that constrains the α value to $(1/\sigma_{\text{model5}})*(\alpha_{\text{model5}}-1)+1=0.418$. If the scale did not matter, this model should provide identical results to Model 5, as in the case of no price variation. The log-likelihood at convergence of this model was –9046.3, which is substantially worse than the convergence value for Model 5. Taken together, the results clearly indicate that the scale is identifiable when there is price variation.

7.2 Presence of Outside Good

The data set used for the case with an outside good is the same as that used in Bhat *et al.* (2006), and is also drawn from the 2000 San Francisco Bay Area Travel Survey (BATS). The sample comprises the time use characteristics (participation and duration of participation) of 2000 adult individuals in ten different activity purposes over a weekend day: (1) maintenance activities (in-home meals, in-home and out-of-home personal household chores and personal care, in-home and out-of-home personal business, out-of-home maintenance shopping, and out-of-home medical appointments), (2) in-home relaxation, (3) in-home recreation (hobbies, TV, *etc.*), (4) non-work internet use, (5) social activities (in-home and out-of-home), (6) out-of-home meals, (7) out-of-home non-maintenance shopping, (8) out-of-home volunteer activities (including civic and religious activities), (9) out-of-home recreation (hobbies, exercise, *etc.*), and (10) pure recreation (travel episodes that began and ended at home without any stops in-between, such as walking or bicycling around the neighborhood). The reader is referred to Bhat *et al.* for further details of the activity typology and definitions. The first activity purpose, maintenance activities, is an "outside" good in which every individual participates.

To demonstrate the issues discussed earlier in the paper in the context of price variation, we synthetically generated unit prices (*i.e.*, cost per minute of participation in each type of activity) for each of the 10 activity types. The unit price for the maintenance activity category (the outside good) is set to unity, while those for the other activity types are generated using draws from a uniform distribution with preset mean values. We only discuss the results for the case of price variation here for brevity (and do not include the case of no price variation).

Table 3 shows the estimation results of four different specifications. The first specification is based on an α-profile corresponding to the expenditure formulation of Equation (19), while the second is based an α-profile corresponding to the consumption formulation of Equation (33). Both these specifications use the first V_k form in Equation (34) and provide the same parameter estimates, but with different log-likelihood values. The α values had to be bounded between 0 and 1 for convergence. As can be observed, the satiation parameter for the first "good" is zero (however, note that we did not explicitly restrict this parameter to zero). The third specification is the same as Model 2, but constrains the scale parameter to 1. A nested likelihood ratio test between Models 2 and 3 clearly indicates the statistically superior fit of Model 2, showing that the scale is identifiable.

The remaining two model specifications in Table 2 are also based on the consumption formulation. The fourth model corresponds to the γ-profile for the internal goods (that is, the second utility form of Equation (32)). Interestingly, the α parameter that is estimable for the outside good was estimated to be zero (we had to bound this α value to be between 0 and 1 for obtaining convergence). This is equivalent to a log-form for both the outside and inside goods in the second utility form in Equation (32). The log-likelihood value for this γ profile is much superior to the α-profile specification of the second model. The fifth specification estimates a common α parameter across the outside and inside alternatives (third utility form in Equation (32)). This specification can be compared to the fourth specification, because the α parameter for the outside good in the fourth model turned out to be zero (in the fourth model, the α parameters for the inside goods are constrained to be zero). The common α parameter in the fifth model is statistically significantly different from zero, indicating that this model is statistically superior to the fourth model. Overall, this fifth specification turns out to be the preferred one in the current empirical context.

8. CONCLUSIONS

Classical discrete and discrete-continuous models deal with situations where only one alternative is chosen from a set of mutually exclusive alternatives. Such models assume that the alternatives are perfectly substitutable for each other. On the other hand, many consumer choice situations are characterized by the simultaneous demand for multiple alternatives that are imperfect substitutes for one another.

A simple and parsimonious econometric approach to handle multiple discreteness was formulated by Bhat (2005) based on a specific satiation-based formulation within the broader Kuhn-Tucker (KT) multiple discrete-continuous economic model of consumer demand. Bhat's model, labeled the MDCEV model, is analytically tractable in the probability expressions and is practical even for situations with a large number of discrete consumption alternatives.

This paper examines several issues associated with extant KT multiple discrete-continuous models. Specifically, the paper proposes a new utility function form that enables clarity in the role of each parameter in the utility specification. The paper also presents identification considerations associated with the utility specification, extends the MDCEV model to the case of price variation across goods and to general error covariance structures, discusses the relationship between earlier KT-based multiple discrete-continuous models, and illustrates the many technical nuances and identification considerations of the multiple discrete-continuous model structure through empirical examples.

Overall, the paper contributes toward the modeling of multiple discrete-continuous choice situations, a field of research that is at an exciting and challenging stage. There have been important contributions to the area from marketing, transportation, and environmental economics, especially within the past five years. At the same time, several challenges lie ahead, including (1) Accommodating more than one constraint in the utility maximization problem (for example, recognizing both time and money constraints in activity type choice and duration models; see Anas, 2006 for a recent theoretical effort to accommodate such multiple constraints), (2) Incorporating latent consideration sets in a theoretically appropriate way within the MDCEV structure, and (3) Using more flexible utility structures that can handle both complementarity as well as substitution among goods, and that do not impose the constraints of additive separability.

ACKNOWLEDGEMENTS

The author would like to thank Ken Vaughn and Chuck Purvis of the Metropolitan Transportation Commissions (MTC) in Oakland for providing help with data related issues. The author appreciates the efforts of Sudeshna Sen on data assembly and preparation of the figures and tables in the paper. Thanks also to Abdul Pinjari and Naveen Eluru for helping with the development of Appendix C and the formatting/documentation of the code accompanying the paper (available at the website indicated in the paper). The author is grateful to Lisa Macias for her help in typesetting and formatting this document. Three anonymous referees provided valuable comments on an earlier version of this paper. The research in this paper was partially supported by an International Visiting Research Fellowship and faculty grant from the University of Sydney.

REFERENCES

Amemiya, T., 1974. Multivariate regression and simultaneous equation models when the dependent variables are truncated normal. *Econometrica*, 42, 999-1012.

Anas, A., 2006. A unified theory of consumption and travel. Presented at the Conference Honoring Kenneth A. Small, Department of Economics, University of California, Irvine, February 3-4, 2006.

Arora, N., Allenby, G.M., and Ginter, J.L., 1998. A hierarchical Bayes model of primary and secondary demand. *Marketing Science*, 17, 29-44.

Baltas, G., 2004. A model for multiple brand choice. *European Journal of Operational Research*, 154, 144-149.

Bhat, C.R., 1995. A heteroscedastic extreme value model of intercity travel mode choice. *Transportation Research Part B*, 29(6), 471-483.

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., 2007. Econometric choice formulations: alternative model structures, estimation techniques, and emerging directions. In *Moving Through Nets: The Physical and Social Dimensions of Travel - Selected papers from the 10th International Conference on Travel Behaviour Research*, K. W. Axhausen (Ed.), Elsevier, pp. 45-80.

Bhat, C.R., and Sen, S., 2006. Household vehicle type holdings and usage: an application of the multiple discrete-continuous extreme value (MDCEV) model. *Transportation Research Part B*, 40(1), 35-53.

Bhat, C.R., and Srinivasan, S., 2005. A multidimensional mixed ordered-response model for analyzing weekend activity participation. *Transportation Research Part B*, 39(3), 255-278.

Bhat, C.R., Srinivasan, S., and Sen, S., 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.

Chan, T.Y., 2006. Estimating a continuous hedonic-choice model with an application to demand for soft drinks. *RAND Journal of Economics*, 37(2), 466-482, Summer.

Chiang, J., 1991. A simultaneous approach to whether to buy, what to buy, and how much to buy. *Marketing Science*, 4, 297-314.

Chintagunta, P.K., 1993. Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Science*, 12, 194-208.

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

Dube, J.P., 2004. Multiple discreteness and product differentiation: Demand for carbonated soft drinks. *Marketing Science*, 23(1), 66-81.

Edwards, Y.D., and Allenby, G.M., 2003. Multivariate analysis of multiple response data. *Journal of Marketing Research*, 40, 321-334.

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. The discrete/continuous model of consumer demand. *Econometrica*, 52, 541-561.

Hendel, I., 1999. Estimating multiple-discrete choice models: An application to computerization returns. *Review of Economic Studies*, 66, 423-446.

Herriges, J.A., Kling, C.L., and Phaneuf, D.J., 2004. What's the use? Welfare estimates from revealed preference models when weak complementarity does not hold. *Journal of Environmental Economics and Management*, 47, 55-70.

Johnson, N.L., and Kotz, S., 1976. *Distributions in Statistics: Continuous Multivariate Distributions*. John Wiley and Sons, New York.

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

Kuhn, H.W. and Tucker, A.W., 1951. Nonlinear programming. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, J. Neyman (Ed.), University of California Press, Berkeley, CA, 481-492.

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

Manchanda, P., Ansari, A., and Gupta, S., 1999. The "shopping basket": A model for multicategory purchase incidence decisions. *Marketing Science*, 18, 95-114.

Phaneuf, D.J., and Herriges, J.A., 2000. Choice set definition issues in a Kuhn-Tucker model of recreation demand. *Marine Resource Economics*, 14, 343-355.

Phaneuf, D.J., and Smith, V.K., 2005. Recreation demand models. In *Handbook of Environmental Economics*, Volume 2, K-G. Mäler and J.R. Vincent (Eds.), North Holland.

Phaneuf, D.J., Kling, C.L., and Herriges, J.A., 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.

Pollak, R., and Wales, T., 1992. *Demand System Specification and Estimation*. Oxford University Press, New York.

Tobin, J., 1958. Estimation of relationships for limited dependent variables. *Econometrica*, 26, 24-36.

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

von Haefen, R.H., 2003a. Latent consideration sets and continuous demand system models. Working paper, Department of Agricultural & Resource Economics, University of Arizona.

von Haefen, R.H., 2003b. Incorporating observed choice into the construction of welfare measures from random utility models. *Journal of Environmental Economics & Management*, 45(2), 145-165.

von Haefen, R.H., 2004. Empirical strategies for incorporating weak complementarity into continuous demand system models. Working paper, Department of Agricultural and Resource Economics, University of Arizona.

von Haefen, R.H., and Phaneuf, D.J., 2003. Estimating preferences for outdoor recreation: a comparison of continuous and count data demand system frameworks. *Journal of Environmental Economics and Management*, 45(3), 612-630.

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

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

Wales, T.J., and Woodland, A.D., 1983. Estimation of consumer demand systems with binding non-negativity constraints. *Journal of Econometrics*, 21(3), 263-85.

Walker, J., 2002. Mixed logit (or logit kernel) model: Dispelling misconceptions of identification. *Transportation Research Record*, 1805, 86-98.

Walker, J., Ben-Akiva, M., and Bolduc, D., 2004. Identification of the logit kernel (or mixed logit) model. Presented at the 10th International Conference on Travel Behavior Research, Lucerne, Switzerland, August.

Walsh, J.W., 1995. Flexibility in consumer purchasing for uncertain future tastes. *Marketing Science*, 14, 148-165.

Appendix A: Form of Utility Function as $\alpha_k \to 0$ for all Goods k

From Equation (1),

$$U(\boldsymbol{x}) = \sum_{k=1}^{K} \gamma_k \psi_k \left[\frac{\left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\}}{\alpha_k} \right] \qquad \text{(A.1)}$$

Consider the expression in parenthesis and write it in the $\frac{0}{0}$ form shown below when $\alpha_k \to 0$:

$$\lim_{\alpha_k \to 0} \left[\frac{\left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\}}{\alpha_k} \right] = \lim_{\alpha_k \to 0} \left[\frac{e^{\alpha_k \ln\left(\frac{x_k}{\gamma_k} + 1 \right)} - 1}{\alpha_k} \right] \qquad \text{(A.2)}$$

Using L'Hospital's rule, we can write the above expression as:

$$\lim_{\alpha_k \to 0} \left[e^{\alpha_k \ln\left(\frac{x_k}{\gamma_k} + 1 \right)} \cdot \ln\left(\frac{x_k}{\gamma_k} + 1 \right) \right] = \ln\left[\frac{x_k}{\gamma_k} + 1 \right]. \qquad \text{(A.3)}$$

Thus, $U(\boldsymbol{x})$ collapses to Equation (2) when $\alpha_k \to 0$ for all *k*.

Appendix B: Derivation of the Structure of the Multiple Discrete-Continuous Extreme Value Model with Error Scale Parameter

From Equation (17) of the text:

$$P\left(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0\right) \tag{B.1}$$

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

$$= |J| \frac{1}{\sigma^{M-1}} \left[\prod_{i=2}^{M} e^{-\left(\frac{V_1 - V_i}{\sigma}\right)} \right] \left[\int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty} \left(e^{-\frac{\varepsilon_1}{\sigma}} \right)^{M-1} \cdot e^{-\sum_{k=1}^{K} \left[e^{-\left(\frac{V_1 - V_k + \varepsilon_1}{\sigma}\right)} \right]} \cdot \frac{1}{\sigma} e^{-\frac{\varepsilon_1}{\sigma}} d\varepsilon_1 \right]$$

Now, consider the last term within the integral in the expression above, and let $t = e^{-\frac{\varepsilon_1}{\sigma}}$. Then $dt = -e^{-\frac{\varepsilon_1}{\sigma}} \cdot \frac{1}{\sigma} d\varepsilon_1$, and we can write the integral as:

$$- \int_{t=+\infty}^{0} t^{M-1} \cdot e^{-\left[t \cdot \sum_{k=1}^{K} e^{-\left(\frac{V_1 - V_k}{\sigma}\right)} \right]} \cdot dt = \frac{(M-1)!}{\left[\sum_{k=1}^{K} e^{-\left(\frac{V_1 - V_k}{\sigma}\right)} \right]^M} \tag{B.2}$$

Putting this back in (B.1), we get

$$P\left(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0\right) \tag{B.3}$$

$$= \frac{1}{\sigma^{M-1}} |J| \left[\prod_{i=2}^{M} e^{-\left(\frac{V_1 - V_i}{\sigma}\right)} \right] \left[\frac{1}{\left(\sum_{k=1}^{K} e^{-\left(\frac{V_1 - V_i}{\sigma}\right)} \right)^M} \right] (M-1)!$$

$$= \frac{1}{\sigma^{M-1}} \left[\prod_{i=1}^{M} c_i \right] \left[\sum_{i=1}^{M} \frac{1}{c_i} \right] \left[\frac{\prod_{i=1}^{M} e^{V_i/\sigma}}{\left(\sum_{k=1}^{K} e^{V_k/\sigma} \right)^M} \right] (M-1)!, \text{ where } c_i = \left(\frac{1-\alpha_i}{e_i^* + \gamma_i p_i} \right).$$

Appendix C: Derivation of the Structure of the Multiple Discrete Continuous Extreme Value (MDCEV) Model from a Differenced Error Structure

From Equation 14 of the text,

$$P(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0) = |J| \int_{\tilde{\varepsilon}_{M+1,1}=-\infty}^{V_1-V_{M+1}} \int_{\tilde{\varepsilon}_{M+2,1}=-\infty}^{V_1-V_{M+2}} \cdots \int_{\tilde{\varepsilon}_{K-1,1}=-\infty}^{V_1-V_{K-1}} \int_{\tilde{\varepsilon}_{K,1}=-\infty}^{V_1-V_K}$$
$$g(V_1-V_2, V_1-V_3, ..., V_1-V_M, \tilde{\varepsilon}_{M+1,1}, \tilde{\varepsilon}_{M+2,1}, ..., \tilde{\varepsilon}_{K,1}) d\tilde{\varepsilon}_{K,1} d\tilde{\varepsilon}_{K-1,1} ... d\tilde{\varepsilon}_{M+1,1} \quad \text{(C.1)}$$

The reader will note that $g(.)$ is a K-1 multivariate logistic distribution with a variance-covariance matrix whose diagonal elements are $\frac{\pi^2\sigma^2}{3}$, and off-diagonal elements are $\underset{l \neq j}{cov}(\tilde{\varepsilon}_{l1}, \tilde{\varepsilon}_{j1}) = \underset{l \neq j}{cov}(\varepsilon_l - \varepsilon_1, \varepsilon_j - \varepsilon_1) = \text{var}(\varepsilon_1) = \frac{\pi^2\sigma^2}{6}$. The probability density function corresponding to $g(.)$ is given by (see page 293, Johnson and Kotz, 1976):

$$g(V_1-V_2, V_1-V_3, ..., V_1-V_M, \tilde{\varepsilon}_{M+1,1}, \tilde{\varepsilon}_{M+2,1}, ..., \tilde{\varepsilon}_{K,1})$$
$$= \frac{(K-1)!}{\sigma^{K-1}} \left(1 + \sum_{j=2}^{M} e^{\frac{-(V_1-V_j)}{\sigma}} + \sum_{j=M+1}^{K} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right)^{-K} e^{-\left[\sum_{j=2}^{M}\left(\frac{V_1-V_j}{\sigma}\right) + \sum_{j=M+1}^{K}\frac{\tilde{\varepsilon}_{j,1}}{\sigma}\right]} \quad \text{(C.2)}$$
$$= \frac{(K-1)!}{\sigma^{K-1}} \left(\sum_{j=1}^{M} e^{\frac{-(V_1-V_j)}{\sigma}} + \sum_{j=M+1}^{K} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right)^{-K} \prod_{j=1}^{M} e^{\frac{-(V_1-V_j)}{\sigma}} \prod_{j=M+1}^{K} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}$$

The probability expression in Equation C.1 can be simplified by evaluating the (K-M)-dimensional integral, one integral at a time. Specifically, rewrite the probability expression in Equation C.1 as:

$$P(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0) =$$
$$|J| \frac{(K-1)!}{\sigma^{K-1}} \prod_{j=1}^{M} e^{\frac{-(V_1-V_j)}{\sigma}} \int_{\tilde{\varepsilon}_{M+1,1}=-\infty}^{V_1-V_{M+1}} \int_{\tilde{\varepsilon}_{M+2,1}=-\infty}^{V_1-V_{M+2}} \cdots \int_{\tilde{\varepsilon}_{K-1,1}=-\infty}^{V_1-V_{K-1}} I_{\tilde{\varepsilon}_{K,1}} d\tilde{\varepsilon}_{K-1,1} ... d\tilde{\varepsilon}_{M+1,1} \quad \text{(C.3)}$$

where, $I_{\tilde{\varepsilon}_{K,1}} = \int_{\tilde{\varepsilon}_{K,1}=-\infty}^{V_1-V_K} \left(\left[\sum_{j=1}^{M} e^{\frac{-(V_1-V_j)}{\sigma}}\right] + \left[\sum_{j=M+1}^{K-1} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right] + e^{\frac{-\tilde{\varepsilon}_{K,1}}{\sigma}}\right)^{-K} \left(\prod_{j=M+1}^{K-1} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right) e^{\frac{-\tilde{\varepsilon}_{K,1}}{\sigma}} d\tilde{\varepsilon}_{K,1}.$

Next, to evaluate $I_{\tilde{\varepsilon}_{K,1}}$, let $s=\left[\sum_{j=1}^{M} e^{\frac{-(V_1-V_j)}{\sigma}}\right]+\left[\sum_{j=M+1}^{K-1} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right]+e^{\frac{-\tilde{\varepsilon}_{K,1}}{\sigma}}$. Then $ds=\frac{-1}{\sigma}e^{\frac{-\tilde{\varepsilon}_{K,1}}{\sigma}}d\tilde{\varepsilon}_{K,1}$, and the integral $I_{\tilde{\varepsilon}_{K,1}}$ can be written as:

$$
\begin{aligned}
I_{\tilde{\varepsilon}_{K,1}} &= \frac{\sigma}{K-1}\left(\left[\sum_{j=1}^{M} e^{\frac{-(V_1-V_j)}{\sigma}}\right]+\left[\sum_{j=M+1}^{K-1} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right]+e^{\frac{-\tilde{\varepsilon}_{K,1}}{\sigma}}\right)^{-K+1}\left(\prod_{j=M+1}^{K-1} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right)\Bigg|_{\tilde{\varepsilon}_{K,1}=-\infty}^{\tilde{\varepsilon}_{K,1}=V_1-V_K} \\
&= \frac{\sigma}{K-1}\left(\left[\sum_{j=1}^{M} e^{\frac{-(V_1-V_j)}{\sigma}}\right]+\left[\sum_{j=M+1}^{K-1} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right]+e^{\frac{-(V_1-V_K)}{\sigma}}\right)^{-K+1}\left(\prod_{j=M+1}^{K-1} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right).
\end{aligned}
\tag{C.4}
$$

Using the above expression for $I_{\tilde{\varepsilon}_{K,1}}$, the probability expression in Equation C.3 can be rewritten as:

$$
\begin{aligned}
&P(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0) = \\
&|J|\frac{(K-2)!}{\sigma^{K-2}}\prod_{j=1}^{M} e^{\frac{-(V_1-V_j)}{\sigma}}\int_{\tilde{\varepsilon}_{M+1,1}=-\infty}^{V_1-V_{M+1}}\int_{\tilde{\varepsilon}_{M+2,1}=-\infty}^{V_1-V_{M+2}}\cdots\int_{\tilde{\varepsilon}_{K-2,1}=-\infty}^{V_1-V_{K-2}} I_{\tilde{\varepsilon}_{K-1,1}}\,d\tilde{\varepsilon}_{K-2,1}...d\tilde{\varepsilon}_{M+1,1}
\end{aligned}
\tag{C.5}
$$

$$
\begin{aligned}
\text{where } I_{\tilde{\varepsilon}_{K-1,1}} &= \int_{\tilde{\varepsilon}_{K-1,1}=-\infty}^{V_1-V_{K-1}}\left(\left[\sum_{j=1}^{M} e^{\frac{-(V_1-V_j)}{\sigma}}\right]+\left[\sum_{j=M+1}^{K-2} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right]+e^{\frac{-\tilde{\varepsilon}_{K-1,1}}{\sigma}}\right)^{-K+1}\left(\prod_{j=M+1}^{K-2} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right)e^{\frac{-\tilde{\varepsilon}_{K-1,1}}{\sigma}}d\tilde{\varepsilon}_{K-1,1} \\
&= \frac{\sigma}{K-2}\left(\left[\sum_{j=1}^{M} e^{\frac{-(V_1-V_j)}{\sigma}}\right]+\left[\sum_{j=M+1}^{K-2} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right]+e^{\frac{-(V_1-V_K)}{\sigma}}\,e^{\frac{-(V_1-V_{K-1})}{\sigma}}\right)^{-K+2}\left(\prod_{j=M+1}^{K-2} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right).
\end{aligned}
$$

In a similar fashion, the probability expression in Equation C.3 can be rewritten in a general form as:

$$
\begin{aligned}
&P(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0) = \\
&|J|\frac{(K-i-1)!}{\sigma^{K-i-1}}\prod_{j=1}^{M} e^{\frac{-(V_1-V_j)}{\sigma}}\int_{\tilde{\varepsilon}_{M+1,1}=-\infty}^{V_1-V_{M+1}}\int_{\tilde{\varepsilon}_{M+2,1}=-\infty}^{V_1-V_{M+2}}\cdots\int_{\tilde{\varepsilon}_{K-(i+1),1}=-\infty}^{V_1-V_{K-(i+1)}} I_{\tilde{\varepsilon}_{K-i,1}}\,d\tilde{\varepsilon}_{K-i,1}...d\tilde{\varepsilon}_{M+1,1}
\end{aligned}
\tag{C.6}
$$

where, for $i=0,1,\ldots,K-M-1$,

$$
I_{\tilde{\varepsilon}_{K-i,1}} = \frac{\sigma}{K-(i+1)}\left(\left[\sum_{j=1}^{M} e^{\frac{-(V_1-V_j)}{\sigma}}\right]+\left[\sum_{j=M+1}^{K-i} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right]+\sum_{l=0}^{i} e^{\frac{-(V_1-V_{K-l})}{\sigma}}\right)^{-K+i+1}\left(\prod_{j=M+1}^{K-(i+1)} e^{\frac{-\tilde{\varepsilon}_{j,1}}{\sigma}}\right)
$$

Of course, the entire integration is completed when $i = K - M - 1$. At this juncture, the probability expression in Equation C.3 simplifies to the MDCEV probability expression:

$$P(e_1^*, e_2^*, e_3^*, ..., e_M^*, 0, 0, ..., 0) = |J| \frac{(M-1)!}{\sigma^{M-1}} \frac{\prod_{j=1}^{M} e^{\frac{-(V_1 - V_j)}{\sigma}}}{\left(\sum_{j=1}^{K} e^{\frac{-(V_1 - V_j)}{\sigma}} \right)^M} = |J| \frac{(M-1)!}{\sigma^{M-1}} \frac{\prod_{j=1}^{M} e^{V_j / \sigma}}{\left(\sum_{j=1}^{K} e^{V_j / \sigma} \right)^M}. \tag{C.7}$$

LIST OF FIGURES

LIST OF TABLES

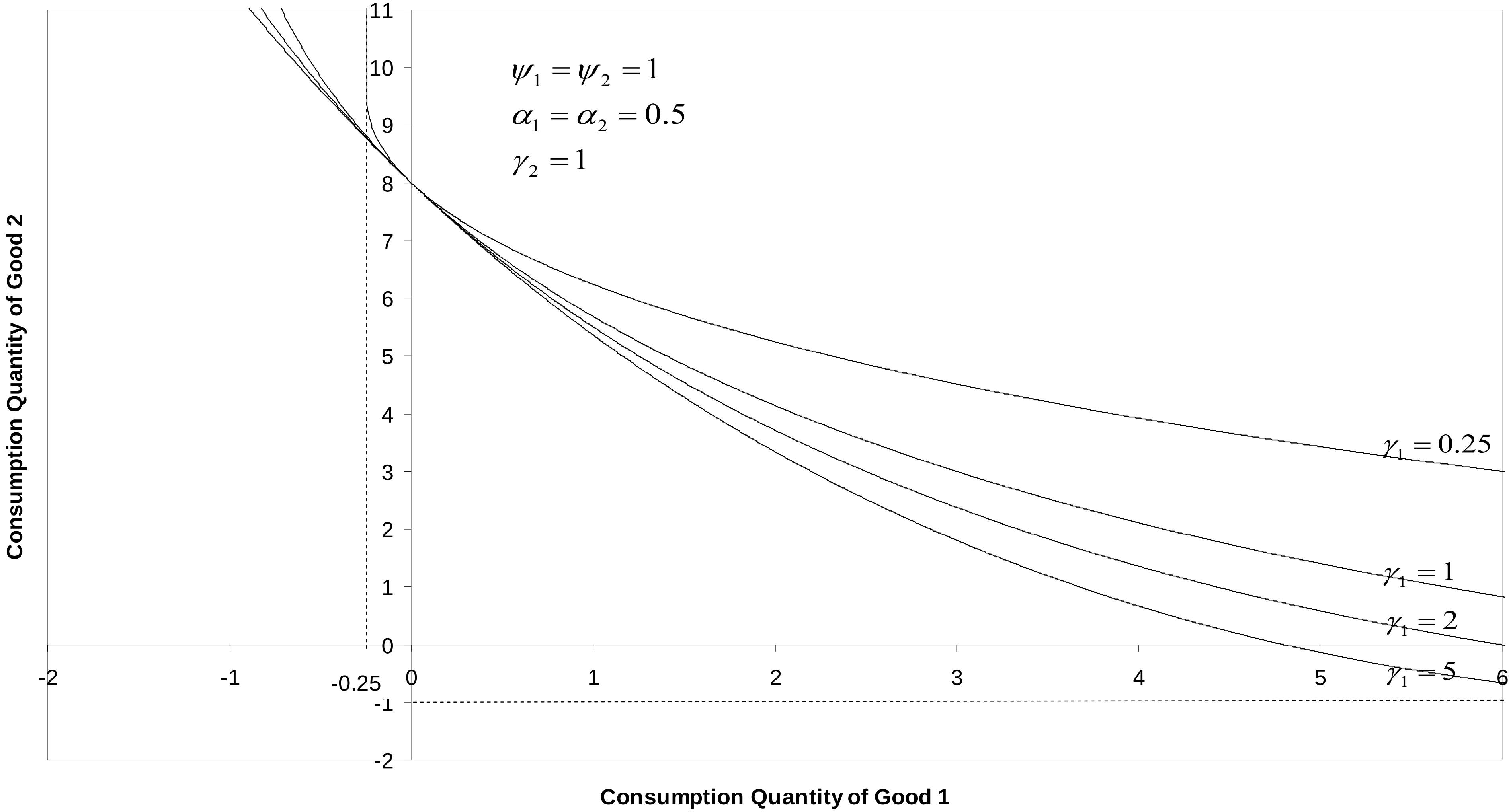

Figure 1. Indifference Curves Corresponding to Different Values of γ_1

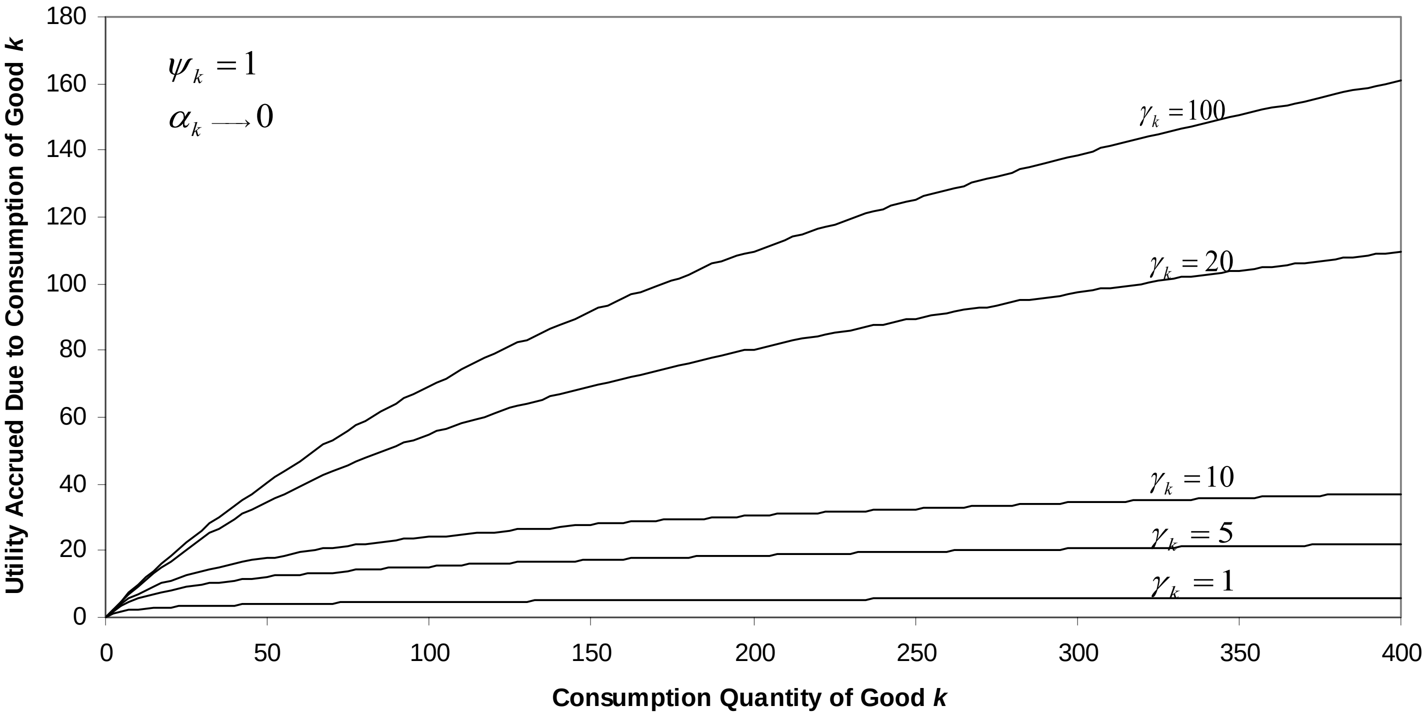

Figure 2. Effect of γ_k Value on Good *k*'s Subutility Function Profile

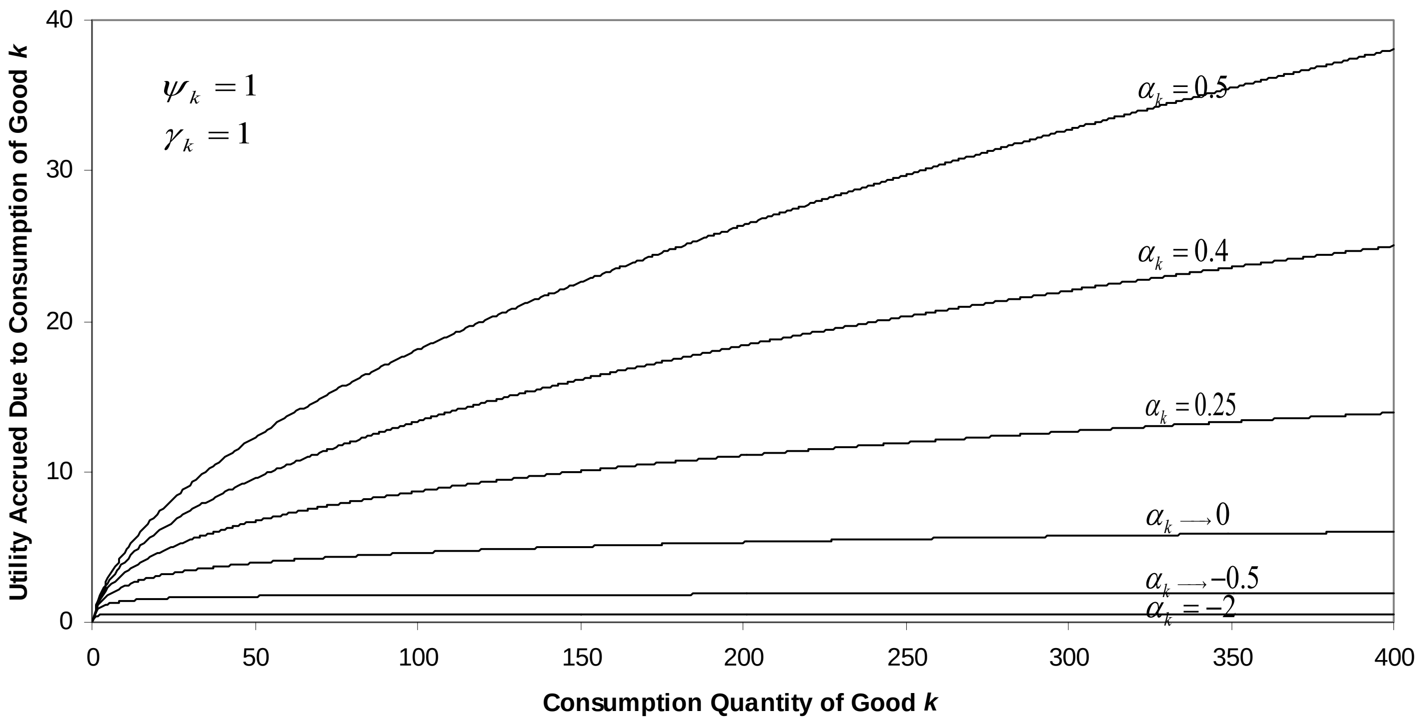

Figure 3. Effect of α_k Value on Good *k*'s Subutility Function Profile

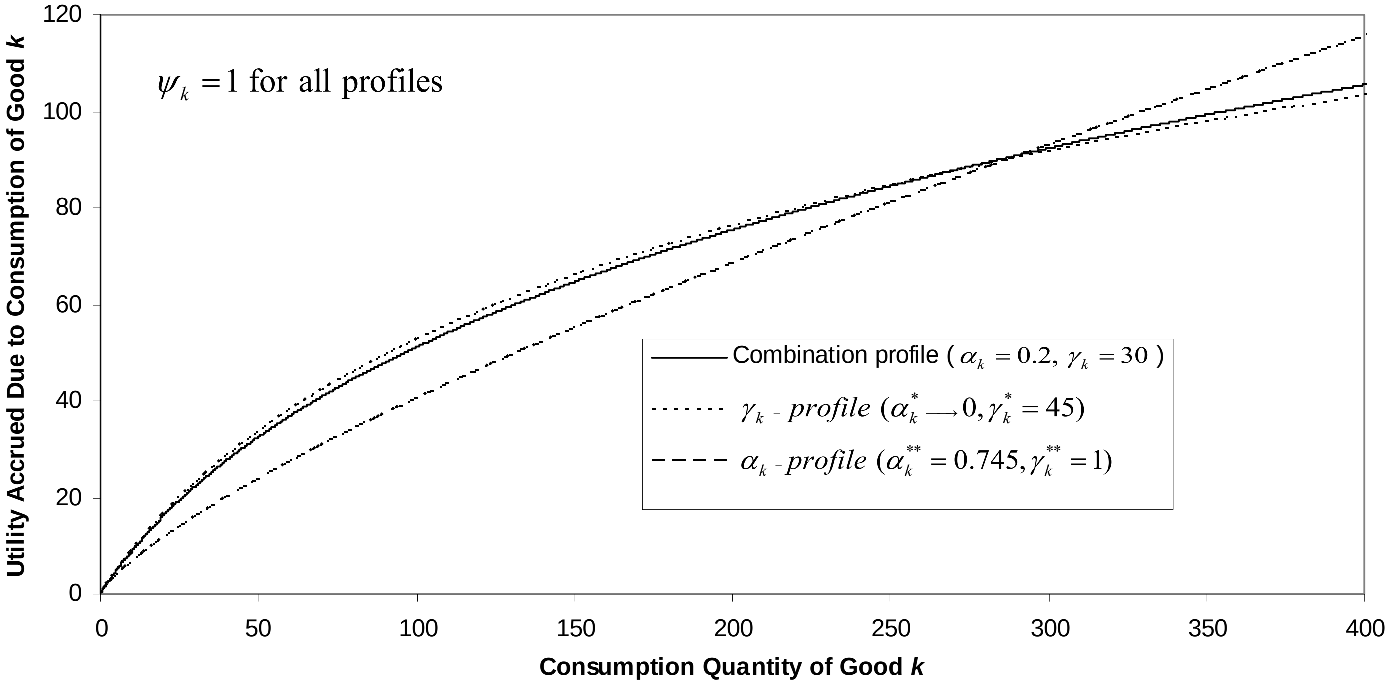

Figure 4a. Alternative Profiles for Moderate Satiation Effects with Low α_k Value and High γ_k Value

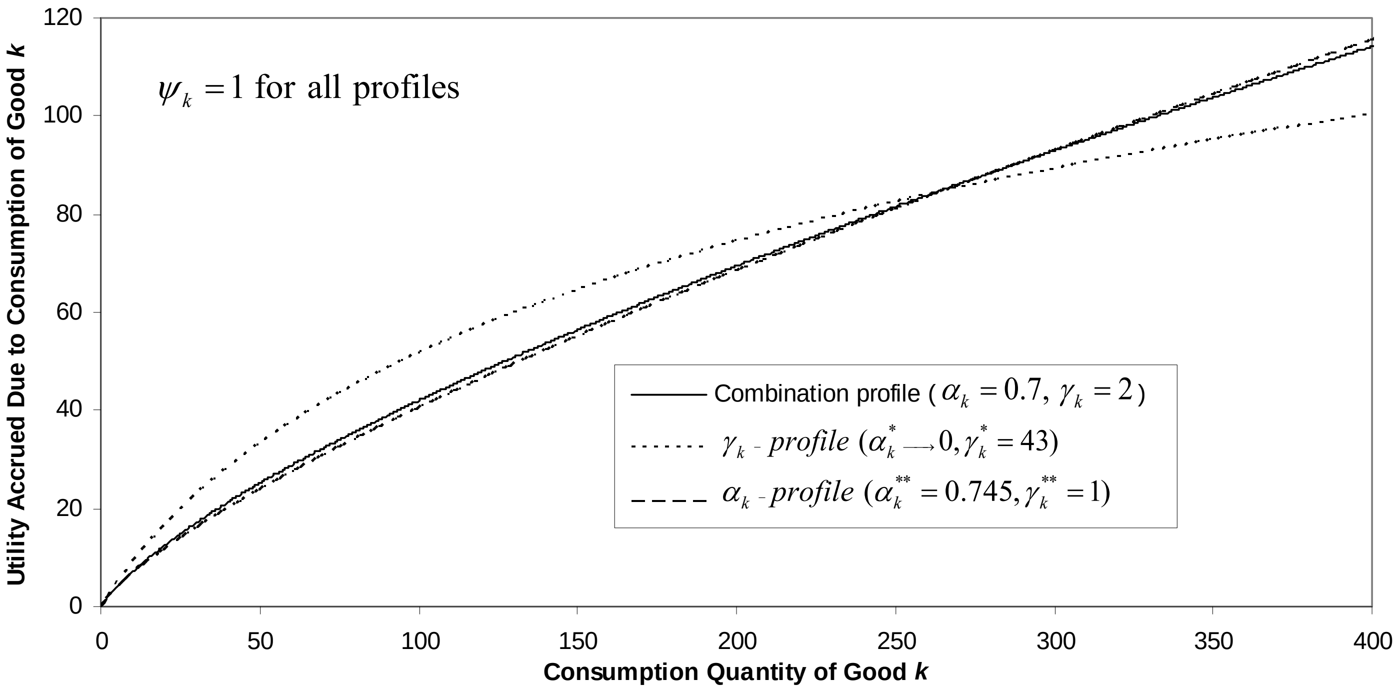

Figure 4b. Alternative Profiles for Moderate Satiation Effects with High α_k Value and Low γ_k Value

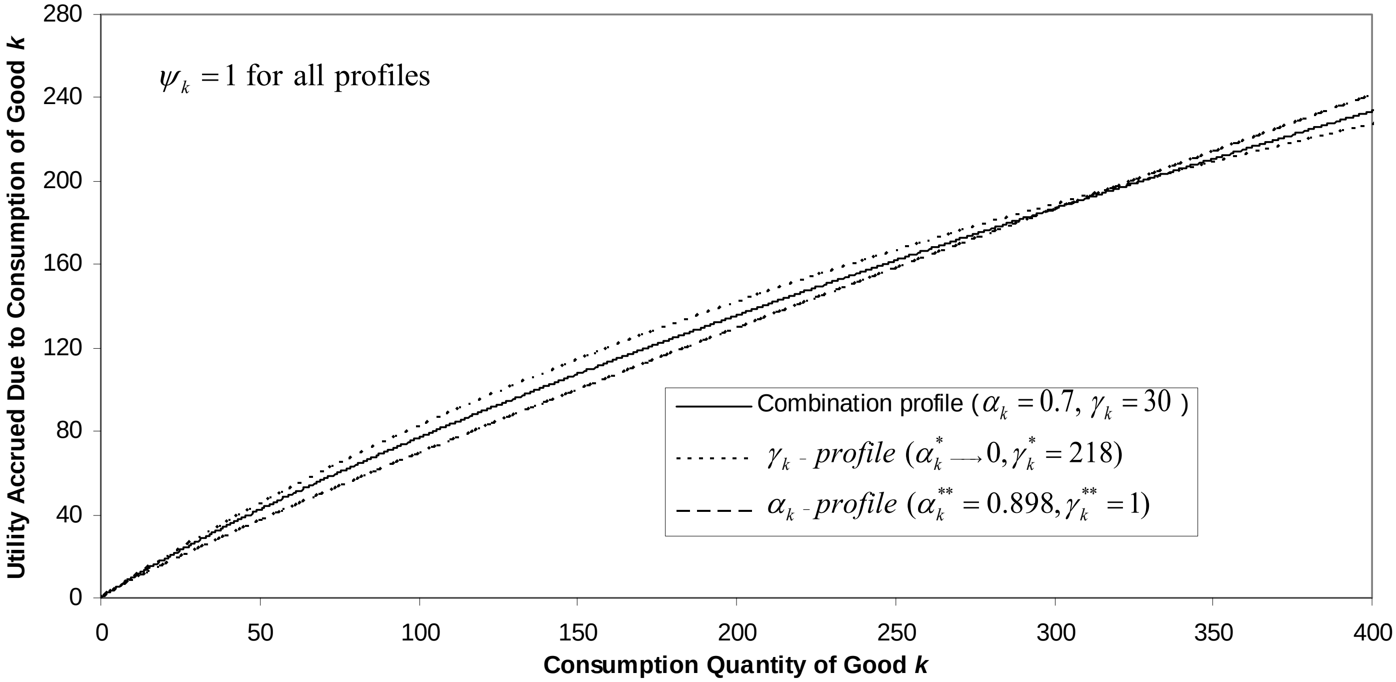

Figure 4c. Alternative Profiles for Low Satiation Effects with High α_k Value and High γ_k Value

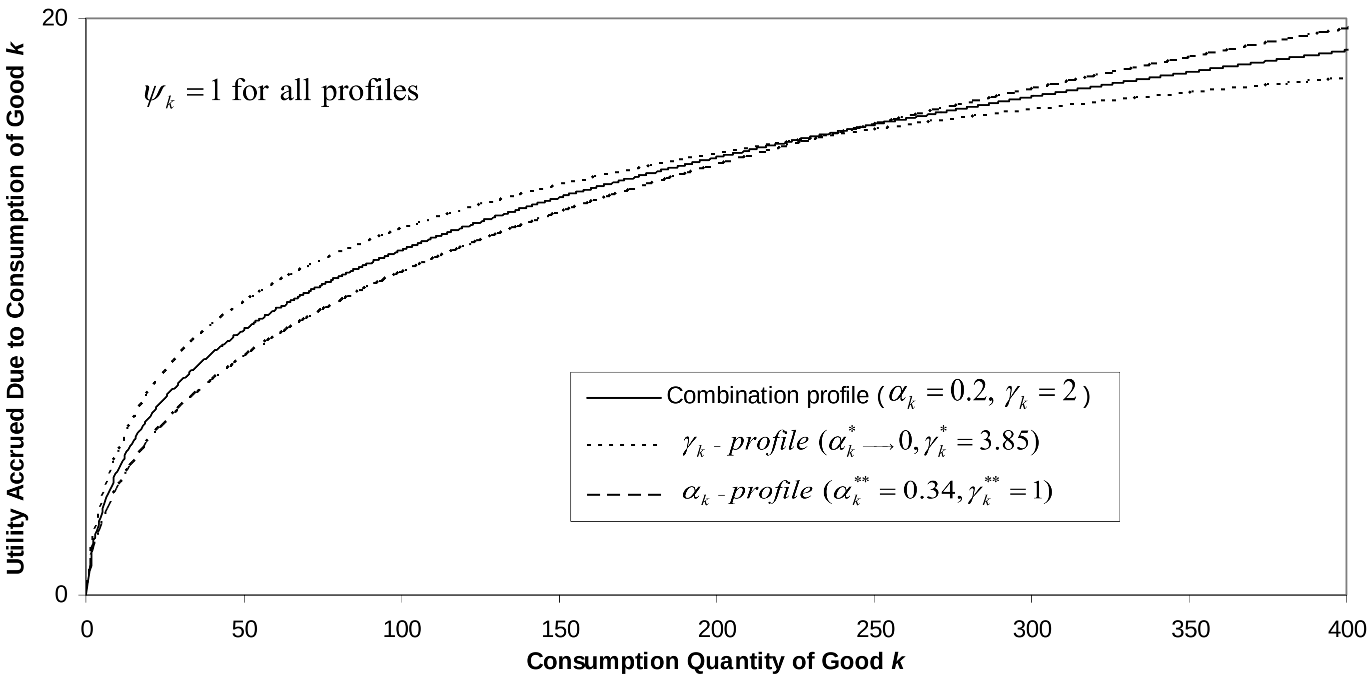

Figure 4d. Alternative Profiles for High Satiation Effects with Low α_k Value and Low γ_k Value

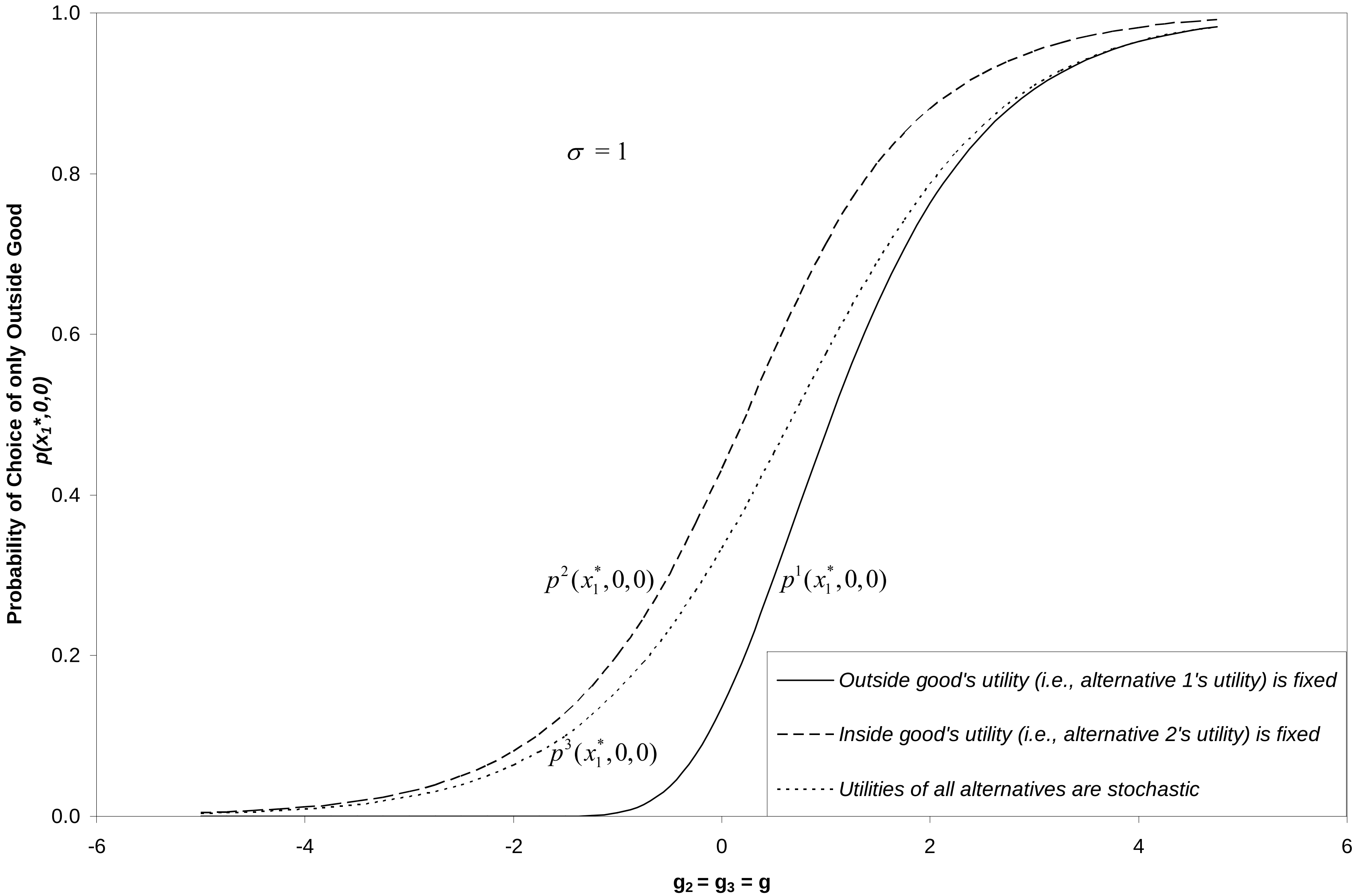

Figure 5. Probability Plots for Choice of Only the Outside Good in a Three-Good Setting

Table 1. Specifications for the "No Outside Good" Case with No Price Variables

Parameters	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Baseline marginal utility (β_k)												
Passenger car	-	-	-	-	-	-	-	-	-	-	-	-
SUV	-2.204	-38.81	-4.409	-38.81	-5.800	-29.74	-10.817	-12.05	-2.003	-37.96	-4.006	-37.96
Pickup-Truck	-1.956	-35.82	-3.912	-35.82	-5.146	-30.62	-9.597	-11.43	-1.766	-35.12	-3.533	-35.12
Minivan	-2.748	-42.08	-5.497	-42.08	-7.230	-27.57	-13.484	-12.20	-2.549	-41.17	-5.098	-41.17
Van	-4.587	-34.28	-9.175	-34.28	-12.069	-22.11	-22.508	-12.06	-4.391	-33.19	-8.782	-33.19
Satiation Parameters												
$\alpha_{Passenger\ Car}$	0.619	38.28	0.239	7.40	0.000	(fixed)	-0.865	-4.91	0.000	(fixed)	-1.000	(fixed)
α_{SUV}	0.886	60.16	0.773	26.24	0.701	16.19	0.443	5.10	0.000	(fixed)	-1.000	(fixed)
$\alpha_{Pick-up}$	0.796	51.36	0.592	19.10	0.463	9.16	0.000	(fixed)	0.000	(fixed)	-1.000	(fixed)
$\alpha_{Minivan}$	0.881	47.18	0.762	20.40	0.687	12.93	0.416	4.00	0.000	(fixed)	-1.000	(fixed)
α_{Van}	0.810	16.85	0.620	6.45	0.500	3.88	0.069	0.28	0.000	(fixed)	-1.000	(fixed)
$\gamma_{Passenger\ Car}$	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	11.470	14.79	11.469	14.79
γ_{SUV}	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	31.107	7.75	31.127	7.75
$\gamma_{Pick-up}$	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	17.880	10.56	17.877	10.56
$\gamma_{Minivan}$	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	29.155	6.60	29.159	6.60
γ_{Van}	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	19.672	3.19	19.591	3.21
Scale parameter	1.000	(fixed)	2.000	(fixed)	2.630	23.47	4.906	13.14	1.000	(fixed)	2.000	(fixed)
Log-Likelihood value at convergence	-9648.48		-9648.48		-9648.48		-9648.48		-9218.89		-9218.89	

Table 2. Specifications for the "No Outside Good" Case with Price Variables

Parameters	Model 1 (Consumption-based)		Model 2 (Consumption-based)		Model 3 (Expenditure-based)		Model 4 (Expenditure-based)		Model 5 (Expenditure-based)		Model 6 (Expenditure-based)	
	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat	Estimate	t-stat
Baseline marginal utility (β_k)												
Passenger car	-	-	-	-	-	-	-	-	-	-	-	-
SUV	-5.088	-15.44	-5.087	-15.35	-5.086	-15.40	-1.814	-32.02	-2.925	-19.95	-1.584	-29.89
Pickup-Truck	-4.465	-15.08	-4.464	-14.99	-4.464	-15.05	-1.564	-28.66	-2.522	-19.05	-1.344	-26.57
Minivan	-6.124	-14.89	-6.123	-14.81	-6.122	-14.86	-1.996	-30.52	-3.488	-18.90	1.764	-28.30
Van	-10.502	-14.50	-10.501	-14.43	-10.499	-14.47	-3.529	-26.33	-6.333	-17.99	-3.302	-24.91
Satiation Parameters												
$\alpha_{Passenger\ Car}$	0.000[22]	0.00	0.000[23]	0.00	0.000[24]	0.00	0.543	35.20	0.000	(fixed)	0.000	(fixed)
α_{SUV}	0.746	20.48	0.746	20.48	0.746	20.49	0.906	71.55	0.000	(fixed)	0.000	(fixed)
$\alpha_{Pick-up}$	0.560	12.56	0.560	12.55	0.560	12.56	0.837	59.74	0.000	(fixed)	0.000	(fixed)
$\alpha_{Minivan}$	0.868	22.47	0.868	22.47	0.868	22.47	0.952	66.51	0.000	(fixed)	0.000	(fixed)
α_{Van}	0.774	7.87	0.776	7.94	0.776	7.95	0.913	24.49	0.000	(fixed)	0.000	(fixed)
$\gamma_{Passenger\ Car}$	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	4.197	11.79	8.513	17.84
γ_{SUV}	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	20.971	6.69	40.921	7.09
$\gamma_{Pick-up}$	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	11.628	8.12	24.360	9.35
$\gamma_{Minivan}$	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	32.795	4.23	68.228	3.93
γ_{Van}	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	28.589	1.68	361.995	1.56
Scale parameter	2.565	17.93	2.565	17.85	2.564	17.90	1.000	(fixed)	1.718	26.76	1.000	(fixed)
Log-Likelihood value at convergence	-9947.07		-8986.11		-9047.08		-9348.95		-8803.51		-8939.42	

[22] The value of 0.000 was the estimated value at convergence (the standard error was 0.0595)
[23] T he value of 0.000 was the estimated value at convergence (the standard error was 0.0600)
[24] The value of 0.000 was the estimated value at convergence (the standard error was 0.0597)

Table 3. Specifications for Case with Outside Good and with Price Variables

Parameters	Model 1 (Expenditure-based)		Model 2 (Consumption-based)		Model 3 (Consumption-based)		Model 4 (Consumption-based)		Model 5 (Consumption-based)	
	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
Baseline marginal utility (β_k)										
Maintenance activity	-	-	-	-	-	-	-	-	-	-
In-home relaxation	-8.562	-31.53	-8.562	-32.21	-6.795	-52.54	-8.369	-36.75	-7.664	-29.44
In-home recreation	-8.053	-29.11	-8.053	-29.73	-6.216	-47.95	-7.785	-33.75	-7.074	-26.92
Non-work internet use	-10.546	-26.06	-10.546	-26.47	-7.514	-43.93	-9.875	-29.38	-9.025	-25.66
Social	-8.307	-29.16	-8.307	-29.77	-6.371	-48.65	-8.038	-33.74	-7.317	-27.23
Out-of-home meals	-6.768	-26.68	-6.768	-27.27	-5.238	-41.12	-6.698	-31.32	-6.019	-24.15
Out-of-home maintenance	-6.464	-25.12	-6.464	-25.67	-4.876	-38.31	-6.350	-29.33	-5.667	-22.60
Out-of-home volunteer	-9.995	-31.41	-9.995	-32.03	-7.692	-55.68	-9.590	-36.17	-8.826	-30.31
Out-of-home recreation	-7.446	-27.40	-7.446	-27.99	-5.674	-43.93	-7.214	-31.79	-6.511	-25.06
Pure recreation	-10.943	-30.66	-10.943	-31.21	-8.278	-55.02	-10.406	-35.08	-9.600	-30.21
Satiation Parameters										
α_1	0.000[25]	0.00	0.000[26]	0.00	0.186	9.05	0.000[27]	0.00	0.108	2.83
α_2	0.739	42.43	0.739	42.53	0.841	90.43	0.000	(fixed)	0.108	2.83
α_3	0.846	56.79	0.846	56.83	0.911	107.05	0.000	(fixed)	0.108	2.83
α_4	0.773	18.92	0.773	18.93	0.862	34.83	0.000	(fixed)	0.108	2.83
α_5	0.749	38.57	0.749	38.63	0.847	78.22	0.000	(fixed)	0.108	2.83
α_6	0.622	29.21	0.622	29.30	0.768	71.92	0.000	(fixed)	0.108	2.83
α_7	0.618	27.02	0.618	27.09	0.767	65.53	0.000	(fixed)	0.108	2.83
α_8	0.708	27.88	0.708	27.92	0.823	56.65	0.000	(fixed)	0.108	2.83
α_9	0.809	48.83	0.809	48.88	0.889	95.66	0.000	(fixed)	0.108	2.83
α_{10}	0.619	15.80	0.619	15.82	0.765	33.25	0.000	(fixed)	0.108	2.83

[25] The value of 0.000 was the estimated value at convergence (the standard error was 0.0372)
[26] The value of 0.000 was the estimated value at convergence (the standard error was 0.0363)
[27] The value of 0.000 was the estimated value at convergence (the standard error was 0.0311)

Table 3. Specifications for Case with Outside Good and with Price Variables (Continued)

Parameters	Model 1 (Expenditure-based)		Model 2 (Consumption-based)		Model 3 (Consumption-based)		Model 4 (Consumption-based)		Model 5 (Consumption-based)	
	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
Satiation Parameters										
γ_1	0.000	(fixed)	0.000	(fixed)	0.000	(fixed)	0.000	(fixed)	0.000	(fixed)
γ_2	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	118.015	10.77	125.194	10.77
γ_3	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	209.869	9.40	219.289	9.55
γ_4	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	92.063	3.99	97.161	4.05
γ_5	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	97.962	9.23	104.049	9.27
γ_6	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	39.642	12.01	42.158	11.95
γ_7	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	29.663	11.26	31.679	11.20
γ_8	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	86.222	8.15	91.102	8.21
γ_9	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	131.518	9.97	137.766	10.11
γ_{10}	1.000	(fixed)	1.000	(fixed)	1.000	(fixed)	35.945	6.17	38.278	6.22
Scale parameter	1.608	30.18	1.608	30.50	1.000	(fixed)	1.402	35.02	1.329	37.22
Log-Likelihood value at convergence	-30054.40		-30013.80		-30229.00		-28220.00		-28216.40	