Allowing for Complementarity and Rich Substitution Patterns in Multiple Discrete-Continuous Models

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

Marisol Castro
Universidad de Chile
Transport System Division
Casilla 228-3, Santiago, Chile
Tel: (562) 29784380, Fax: (562)6894206
Email: marisolcastro@gmail.com

Abdul Rawoof Pinjari
University of South Florida
Department of Civil and Environmental Engineering
4202 E. Fowler Ave., ENC 2503, Tampa, Florida 33620
Tel: 813-974-9671, Fax: 813-974-2957
Email: apinjari@usf.edu

*corresponding author

Original version: January 2013
First revision: February 2015
Second revision: August 2015

ABSTRACT

Many consumer choice situations are characterized by the simultaneous demand for multiple alternatives that are imperfect substitutes for one another, along with a continuous quantity dimension for each chosen alternative. To model such multiple discrete-continuous choices, most multiple discrete-continuous models in the literature use an additively-separable utility function, with the assumption that the marginal utility of one good is independent of the consumption of another good. In this paper, we develop model formulations for multiple discrete-continuous choices that accommodate rich substitution structures and complementarity effects in the consumption patterns, and demonstrate an application of the model to transportation-related expenditures using data drawn from the 2002 Consumer Expenditure (CEX) Survey.

Keywords: Discrete-continuous system, multiple discreteness, Karush-Kuhn-Tucker demand systems, random utility maximization, non-additively separable utility form, transportation expenditure.

1. INTRODUCTION

Multiple discrete-continuous (MDC) choice situations are quite ubiquitous in consumer decision-making, and constitute a generalization of the more classical single discrete-continuous choice situation. Examples of MDC contexts include the participation decision of individuals in different types of activities over the course of a day and the duration in the chosen activity types (see Bhat, 2005, Chikaraishi *et al.*, 2010, and Wang and Li, 2011), household holdings of multiple vehicle body/fuel types and the annual vehicle miles of travel on each vehicle (Ahn *et al.*, 2008), and consumer purchase of multiple brands within a product category and the quantity of purchase (Kim *et al*., 2002).

To date, most MDC modeling frameworks, including Bhat's (2005, 2008) MDCEV model, have considered the case of imperfect substitutes and perfect substitutes, but not the case of complementary goods (the case of imperfect as well as perfect substitutes can be handled through a nested MDC-SDC model, as in Bhat *et al*., 2009). However, complementary goods occur quite frequently in consumer choice situations. For example, in the consumer expenditure literature, consider the case of annual household expenditures on transportation and other commodities (such as housing, clothing, and food). Also, let the transportation expenditures be disaggregated into such categories as vehicle purchase, gasoline/oil, vehicle insurance, vehicle maintenance, and public transportation. Then, there are likely to be complementarity effects in the expenditures on gasoline, vehicle insurance, and vehicle maintenance, as well as strong substitution effects between these three categories of auto-related expenditures and public transportation expenditures. If the public transportation category is further broken down by rail or bus, it is possible that these two sub-categories are perfect substitutes in that there is expenditure on only one or the other of these two alternatives. This example context then is a case of alternatives that are complementary, imperfect substitutes, as well as perfect substitutes. Similarly, in the activity-based travel modeling and time-use literature, an analyst may be interested in daily non-work, non-sleep time-use patterns in such activities as relaxing, running indoors, running outdoors, and eating. Here, relaxing and eating may be complements, while relaxing and running outdoors may be imperfect substitutes, and running indoors and running outdoors may be perfect substitutes.

The reason why most earlier MDC studies are unable to consider complementarity stems from the use of an additively separable utility function (ASUF) and the usual assumption of a quasi-concave and increasing utility function with respect to the consumption of goods (see Deaton and Muellbauer, 1980, page 139). Besides, the additive utility structure makes it difficult to incorporate even rich imperfect substitution patterns across alternatives because the marginal rate of substitution between any pair of goods is dependent only on the quantities of the two goods in the pair, and independent of the quantity of other goods (see Pollak and Wales, 1992). Thus, back to the activity-based travel modeling and time-use fields, consider an individual living alone with three recreation activity options: watching TV at home, visiting friends, and going to the movies. Let this individual currently spend all her time watching TV at home. As she spends more and more time watching TV, the traditional utility formulation does recognize that there is satiation and that the marginal utility of an additional unit of time spent watching TV decreases. However, the additive utility formulation assumes that the utility of visiting friends is unaffected by the amount of time watching TV. But as the time spent watching TV increases, it may increase the marginal utility of visiting friends. If the latter is true, it would imply a higher likelihood to participate in visiting friends and a higher time investment in visiting friends, relative to the case when this interaction between the investment in one alternative and the utility of another is completely ignored (as in the additive utility function). Of course, whether such an interaction exists, and the direction of such an interaction, may be an empirical issue. This suggests that one should consider a richer non-additive utility function and then examine its performance against a traditional utility function.

Overall, the additively separable assumption substantially reduces the ability of the utility function to accommodate rich and flexible substitution patterns, as well as to accommodate complementarity effects. At the same time, the literature on MDC models that adopt a non-additively separable (NAS) utility function is very limited, and research in this area has arisen only in the past five years or so. Song and Chintagunta (2007) and Mehta (2007) accommodated complementarity and substitution effects in an MDC utility function to model purchase quantity decisions of house cleaning products. However, both studies use an indirect utility approach instead of a direct utility approach. As clearly articulated by Bunch (2009), the direct utility approach has the advantage of being closely tied to an underlying behavioral theory, so that interpretation of parameters in the context of consumer preferences is clear and straightforward. Further, the direct utility approach provides insights into identification issues. Later, Lee and Allenby (2009) proposed a direct utility approach that incorporates a NAS utility structure. For this purpose, they grouped goods in categories assuming that goods in the same category are substitutes, while goods in different categories are complements. However, their modeling framework does not allow consumers to choose multiple goods within each category. Lee *et al.* (2010) proposed a direct utility model for measuring asymmetric complementarity. Their model formulation, however, was developed for the simple case of only two goods.

Vásquez-Lavín and Hanemann (2008) or VH extended Bhat's (2008) additively separable linear form allowing the marginal utility of each good to be dependent on the level of consumption of other goods. In this paper, we use the VH utility formulation (VHUF) as the starting point, but suppress a term used in the VHUF that can create interpretation and identification problems. The resulting utility forms remain flexible, while also being easy to estimate and expanding the range of local consistency of the utility function relative to the VHUF. We also develop several ways to introduce stochasticity in the utility specification. The stochastic forms we introduce essentially acknowledge two different sources of errors. The first source of errors arises when consumers make random "mistakes" in maximizing their utility function, and the second source of errors originates from the analyst's inability to observe all factors relevant to the consumer's utility formation. To our knowledge, this is the first time that such a distinction is being made between the two sources of errors in a NAS-MDC model. Basically, the first source assumes that the analyst knows exactly how consumers valuate goods (that is, the analyst knows the utility functions of consumers exactly), but the analyst also acknowledges that there may be a difference between the optimal consumptions as computed by the analyst based on the "exact" utilities and as actually observed to be made by the consumers. This may be because consumers do not go through a rigorous mathematical optimization process, and make random "mistakes" about (statistically speaking) what the actual consumption patterns must have been. This causes two consumers who are exactly the same, or the same consumer in exactly the same choice environment, to reveal different consumption patterns. We call this as the deterministic utility-random maximization (DU-RM) stochastic specification in the rest of this paper. The second source is the more traditional one used in the economic and transportation literature. Here the analyst introduces stochasticity directly in the utility function to acknowledge that the analyst does not know all the factors that is considered by the consumer in her/his valuation pattern for goods. However, the consumer is assumed to make a perfect optimization decision given her or his utility formation. We refer to this as the random utility-deterministic maximization (RU-DM) stochastic specification.[1] A third approach combines the random utility as well as the random maximization specifications in what may be considered the most realistic situation. We refer to this as the random utility-random maximization (RU-RM) stochastic specification. For each of the three proposed stochastic formulations, we are able to retain a relatively simple form for the model, and the structure of the Jacobian in the likelihood function is also relatively simple. The formulations are applied to the empirical context of household transportation expenditures.

The rest of this paper is structured as follows. The next section formulates a functional form for the non-additive 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 alternative stochastic forms of the utility specification and the resulting general structures for the probability expressions. Section 4 provides an empirical demonstration of the model proposed in this paper. The final section concludes the paper.

2. FUNCTIONAL FORM OF UTILITY SPECIFICATION

The starting point for our utility functional form is Bhat (2008), who proposes a linear Box-Cox version of the constant elasticity of substitution (CES) direct utility function for MDC models:

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

where $U(\boldsymbol{x})$ is strictly quasi-concave, non-decreasing in all arguments and strictly increasing in at least one, and a continuously differentiable function with respect to the consumption quantity (K×1)-vector $\boldsymbol{x}$ ($x_k \geq 0$ for all k). The ψ_k, γ_k and α_k parameters are associated with good k. The reader will note that there is an assumption of additive separability of preferences in the utility form of Equation (1) (Bhat, 2008 discusses the many reasons for the use of the Box-Cox form of Equation (1) for MDC models; for ease in presentation, we will refer to Equation (1) as Bhat's Additively Separable Utility function or the B-ASUF). The B-ASUF is a valid utility function if $\psi_k > 0$, $\gamma_k > 0$, and $\alpha_k \leq 1$ for all k. For presentation ease, we assume temporarily that there is no "essential good" (that is, we present the case of "non-essential goods only"), so

[1] Of course, many other interpretations may be provided for these two sources of error.

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; in fact, as we will show later, the econometrics become much easier when there is an essential good for which there is always positive consumption).[2] We also assume for now that the utility function is deterministic to focus on functional form issues (important modeling issues arise when we introduce stochasticity, which we discuss in Section 3). The possibility of corner solutions implies that the term γ_k in the B-ASUF, which is a translation parameter, should be greater than zero for all *k*.[3]

Vásquez-Lavín and Hanemann or VH (2008) extended the ASUF and presented a quadratic version of it that relaxes the additively separable form, as below:

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

where $\psi_k > 0$, $\gamma_k > 0$, and $\alpha_k \le 1$ for all *k* (we will refer to Equation (2) as the VH-UF) The new interaction parameters θ_{km} allow quadratic effects as well as allow the marginal utility of good *k* to be dependent on the level of consumption of other goods. Positive interaction parameters accommodate complementarity effects, while negative interaction parameters accommodate substitution effects. Of course, if $\theta_{km} = 0$ for all *k* and *m*, the utility function collapses to the B-ASUF. Also, following Bhat (2008), it is very difficult to disentangle the γ_k and α_k effects separately (as in the ASUF). Thus, for identification purposes, we either have to constrain α_k to zero for all goods (technically, assume $\alpha_k \to 0 \ \forall k$) and estimate the γ_k parameters (*i.e.*, the γ-profile utility form), or constrain γ_k to 1 for all goods and estimate the α_k parameters (*i.e.*, the α-profile utility form).[4]

---

[2] In general, utility specifications are structured so that, if present, there is a single essential good that is characterized as an "outside" good and not of primary interest to the analyst, but is included simply to make the total consumption on the goods of interest endogenous. The non-essential goods are typically characterized as the "inside" goods. For a detailed discussion of the notions of "inside " and "outside" goods, the reader is referred to von Haefen (2010) and Bhat *et al*. (2013).

[3] 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.

[4] The γ -profile equivalent of VH-UF is: $U(\boldsymbol{x}) = \sum_{k=1}^{K} \gamma_k \psi_k \ln\left(\frac{x_k}{\gamma_k}+1\right) + \frac{1}{2}\sum_{k=1}^{K}\sum_{m=1}^{K}\theta_{km}\gamma_k\gamma_m \ln\left(\frac{x_k}{\gamma_k}+1\right)\ln\left(\frac{x_m}{\gamma_m}+1\right)$.

The VH-UF, like the basic translog utility function and quadratic utility function, is a flexible functional form that has enough parameters to provide a second-order approximation to any true unknown direct twice-differentiable utility functional form at a local point (see Pollack and Wales, 1992, page 53, 60, and Sauer *et al*., 2006). It also is a non-additive functional form. Of course, because of the budget constraint, as in the additive, translog and quadratic utility forms, one must place a normalization on the ψ_k values (see Wales and Woodland, 1983, Pollack and Wales, 1992, page 57, and Holt and Goodwin, 2009). Many earlier studies impose the identification condition that $\sum_k \psi_k = 1$, though we will impose a different normalization during our estimations. Also, to adhere to the utility maximization principle that we use as the decision rule, one must impose symmetry of the interaction parameters; that is $\theta_{km} = \theta_{mk} \ \forall\, m, k$ (see Jorgenson and Lau, 1975, and Holt and Goodwin, 2009). This guarantees the symmetry of the second derivatives (or Hessian) matrix of the utility function with respect to the consumption quantities. Additionally, there is another positivity condition that is needed to ensure that the utility function is increasing, as we discuss in Section 2.1.1. Finally, to ensure correct curvature (that is, the quasi-concavity of the utility function with respect to quantities, or the negative semi-definiteness of the Hessian matrix) at the consumption points represented in the empirical sample, one can reparameterize the interaction parameter matrix as $\boldsymbol{\theta} = -\boldsymbol{L}\boldsymbol{L}'$, where $\boldsymbol{L}$ is a lower triangular Cholesky matrix (see Ryan and Wales, 1998 and Holt and Goodwin, 2009). However, we do not impose any homogeneity restrictions related to expenditure shares being invariant to expenditure level; that is, we do not impose the restriction that $\sum_m \theta_{km} = 0$ for each good k.

As indicated earlier, the VH-UF form can provide a local approximation to any direct twice-differentiable utility function. The restrictions imposed above also help obtain local consistency of the utility function. In the next section, we clarify the role of parameters, present empirical identification considerations, and recommend a flexible form that is easier to estimate and expands the range of local consistency of the utility function relative to the VH-UF.

The α-profile equivalent of VH-UF is:

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

2.1. Role of Parameters in Non-Additively Separable Utility Specification

2.1.1. Role of ψ_k

The marginal utility of consumption with respect to good *k* can be written from the VH-UF as:

$$\frac{\partial U(\boldsymbol{x})}{\partial x_k} = \left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k-1}\left\{\psi_k + \sum_{m=1}^{K}\theta_{km}\frac{\gamma_m}{\alpha_m}\left[\left(\frac{x_m}{\gamma_m}+1\right)^{\alpha_m}-1\right]\right\}. \quad (3)$$

The difference between the above expression and the corresponding one in the B-ASUF is the presence of the second term in parenthesis, which includes the consumptions of other goods. Thus, the formulation is not additively separable, but one in which the marginal utility of a good is dependent on the consumption amounts of other goods. The marginal utility at zero consumption of good *k* (that is, the baseline utility of good *k*) collapses to:

$$\left.\frac{\partial U(\boldsymbol{x})}{\partial x_k}\right|_{x_k=0} = \tilde{\pi}_k = \psi_k + \sum_{m\neq k}\theta_{km}\frac{\gamma_m}{\alpha_m}\left[\left(\frac{x_m}{\gamma_m}+1\right)^{\alpha_m}-1\right]. \quad (4)$$

From above, it is clear that ψ_k is no longer the baseline (marginal) utility at the point at which good *k* has not been consumed (as it is in the B-ASUF). Rather, in the VH-UF, it is the baseline (marginal) utility of good *k* at the point at which no good has been "consumed"; that is, when $x_m = 0 \;\; \forall m$ (no consumption decision has yet been made). This also indicates that, if prices of all goods are the same, then the good with the highest value of ψ_k will definitely see some positive consumption.[5]

Another important point to note from Equation (3) is that for the utility function to be increasing in $x_k \;\; (k = 1,2,...,K)$, the following condition should be satisfied for all possible values of the consumption vector $\boldsymbol{x}$:

$$\psi_k + \sum_{m=1}^{K}\theta_{km}\frac{\gamma_m}{\alpha_m}\left[\left(\frac{x_m}{\gamma_m}+1\right)^{\alpha_m}-1\right] > 0 \text{ for all } k. \quad (5)$$

This is in addition to the condition in the B-ASUF where $\psi_k > 0 \; \forall \; k$.

[5] If there is price variation across goods, the good with the highest price-normalized marginal utility ψ_k / p_k will definitely see some positive consumption (see Bhat, 2008 and Pinjari and Bhat, 2011 for discussions). Also, the reader will note that the idea that consumers start from a clean slate and then work their way to the optimal "basket" of consumption based on marginal utilities is consistent with the utility maximization principle, though it need not represent the way many consumers actually reach their optimal consumption point.

2.1.2. Role of γ_k

As in the B-ASUF, the γ_k parameter allows for corner solutions. In particular, the γ_k terms 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*. In addition to allowing corner solutions, the γ_k terms also serve as satiation parameters. In general, the higher the value of γ_k, the less is the satiation effect in the consumption of x_k. However, unlike in the B-ASUF, γ_k affects satiation for good *k* in two ways. The first effect is through the first term on the right side of the VH-UF, and the second is through the second term on the right side of the VH-UF that generates quadratic effects. The overall effect depends on the sign and magnitude of the parameter θ_{kk} in the second term. If this term is negative, and particularly for high values of γ_k, we can get an inappropriate parabolic shape for the contribution of alternative *k* to overall utility within the range of x_k. In particular, beyond a certain point of consumption of alternative *k*, there is negative marginal utility. This is because of the violation of the condition in Equation (5). An illustration is provided in Figure 1, which plots the utility contribution of alternative *k* for $\psi_k = 1$, $\alpha_k \to 0$, $\theta_{kk} = -0.02$, $\theta_{km} = 0 \;\; \forall m \neq k$, and different values of γ_k ($\gamma_k = 1$, 10, and 30). As can be observed, for the γ_k value of 30, we get a profile that peaks at about 110 units, and violates the requirement that the utility function be strictly increasing (this is also shown in Vásquez-Lavín and Hanemann, 2008). On the other hand, if θ_{kk} is positive and quite high in magnitude, it is possible that, for high γ_k values, there is in fact an increase in the marginal utility effect at low values of x_k (essentially a violation of the strictly quasi-concave assumption of the utility function). This is because the left side of Equation (5) becomes an increasing function of x_k at low x_k values. Figure 2 illustrates such a case for $\psi_k = 1$, $\alpha_k \to 0$, $\theta_{kk} = +0.2$, $\theta_{km} = 0 \;\; \forall m \neq k$, and different values of γ_k ($\gamma_k = 1$, 10, and 30). For $\gamma_k = 10$, one can discern the increasing marginal utility until about 6.5 units after which the shape becomes one of decreasing marginal utility. The increasing marginal utility at low values is particularly pronounced for $\gamma_k = 30$, which continues until a value of 40 units before starting to decrease in marginal utility. We will return to these issues in Section 2.2.

The translation parameters γ_m of other goods also have an impact on the utility contribution of good *k*, through the influence on the baseline (marginal) utility of good *k* (see Equation (4)). Specifically, for a given value of x_m, the baseline (marginal) utility for good *k* increases as γ_m increases for positive θ_{km} values and decreases as γ_m increases for negative θ_{km} values.

2.1.3. Role of α_k

The express role of α_k is to reduce the marginal utility with increasing consumption of good *k*; that is, it represents a satiation parameter. However, as in the case of the γ_k effect on consumption of good *k*, there are two effects of the α_k parameter – one through the first term on the right side of the VH-UF and the second through the quadratic effect caused by the combination of the first and second terms on the right side of the VH-UF. The overall α_k effect depends on the sign and magnitude of the parameter θ_{kk} in the second term. If this term is negative, and particularly for values of α_k close to 1, we can get a "nonsensical" parabolic shape for the utility contribution of alternative *k* within the usual possible range of x_k. An illustration is provided in Figure 3, which plots the utility contribution of alternative *k* for $\gamma_k = 1$, $\psi_k = 1$, $\theta_{kk} = -0.03$, $\theta_{km} = 0 \ \forall m \neq k$, and different values of α_k. As can be observed, at the α_k value of 0.6, we get a profile that peaks at about 150 units and decreases thereafter, violating the requirement that the utility function be strictly increasing. On the other hand, if θ_{kk} is positive and quite high in magnitude, it is possible that, for high α_k values, there is in fact an increase in the marginal utility effect at some low values of x_k. Figure 4 illustrates such a case for $\gamma_k = 1$, $\psi_k = 1$, $\theta_{kk} = +0.2$, $\theta_{km} = 0 \ \forall m \neq k$, and different values of α_k. The non-conforming utility profile is obvious for the α_k value of 0.8.

The α_m parameters for other goods also impact the baseline (marginal) utility of good *k* (see Equation (4)). For a given value of x_m, the baseline (marginal) utility for good *k* decreases as α_m falls down from 1 for positive θ_{km} values and increases as α_m falls down from 1 for negative θ_{km} values.

2.2. Empirical Identification Issues Associated with Utility Form

The total number of parameters in the flexible utility functional form of Equation (2) rises rapidly with the number of alternatives, especially in the θ_{km} terms ($k = 1,2,...,K;\ m = 1,2,...,K$). There are also empirical identification issues that arise with the utility form. In addition to the consideration that only one set of the γ_k and α_k parameters are empirically identifiable, there is an additional empirical identification issue. This is because the θ_{kk} parameters in the quadratic utility functional form essentially also serve as "satiation" parameters by providing appropriate curvature to the utility function. However, empirically speaking, it is difficult to disentangle the θ_{kk} effects from the γ_k effects (for the γ - profile) and from the α_k effects (for the α - profile) as long as the θ_{kk} effects do not become that negative as to bring on a parabolic shape at even low to moderate consumption levels (this latter case would anyway be inappropriate to represent the utility function). In fact, a utility profile based on a combination of θ_{kk} and γ_k values for the γ - profile case can be closely approximated by a utility function based solely on γ_k values with $\theta_{kk} = 0$. Similarly, a utility profile based on a combination of θ_{kk} and α_k values for the α - profile case can be closely approximated by a utility function based solely on α_k values with $\theta_{kk} = 0$. While a rigorous mathematical proof is not provided here, this may be illustrated as in Figure 5 for the γ - profile, with $\psi_k = 1$ and $\theta_{km} = 0\ \forall m \neq k$. The figure shows that alternative k's contribution to utility based on a certain combination of γ_k and θ_{kk} values can be closely replicated by other combination values of γ_k and θ_{kk}. In particular, the utility profiles corresponding to combinations of γ_k and θ_{kk} values can be replicated very closely by a profile that corresponds to $\gamma_k = 1$ and some specific θ_{kk} value, or by a profile that corresponds to $\theta_{kk} = 0$ and some specific γ_k value. Thus, in Figure 5, the utility profiles corresponding to $\gamma_k = 7.45$ and $\theta_{kk} = 0$, and $\gamma_k = 1$ and $\theta_{kk} = 1.3$, are able to closely replicate all the other utility profiles. A similar situation may be observed from Figure 6 for the α - profile, where the utility profiles of different combinations of θ_{kk} and α_k values can be approximated closely by the profile corresponding to $\alpha_k = 0.442$ and $\theta_{kk} = 0$, and $\alpha_k = 0$ and $\theta_{kk} = 1.38$.

The discussion above suggests that, while one may be able to theoretically estimate the VH-UF by retaining α_k, γ_k, and θ_{kk}, this will be empirically difficult to estimate. Thus, without much loss of empirical generality, one should be able to normalize $\gamma_k = 1$ (and estimate θ_{kk}) or set $\theta_{kk} = 0$ (and estimate γ_k) for each good *k* in the γ-profile case (the first normalization represents Christiansen et al.'s direct basic translog utility function, as should be obvious from imposing the normalization in the γ-profile utility function in footnote 4). In the α-profile case, one can normalize $\alpha_k = 0$ (and estimate θ_{kk}) or set $\theta_{kk} = 0$ (and estimate α_k) for each good *k* in the utility function (the first normalization represents Wales and Woodland's quadratic utility function, as can be observed from imposing the normalization in the α-profile utility function in footnote 4). Unlike the implicit normalizations adopted by the earlier studies, we propose, apparently for the first time in the literature, to employ the second normalization in each of the cases just discussed. That is, we set $\theta_{kk} = 0$ for each good, since this immediately removes the possibility of a parabolic shape for the utility contribution of good *k*. At the same time, along with the symmetry and local quasi-concavity restrictions discussed just before Section 2.1, we reduce the range of consumption bundles for which the utility function may not be consistent. The result is also improved clarity in the interpretation of the γ_k and α_k parameters, which now have the same interpretation as satiation parameters corresponding to good *k* as in the B-ASUF. Besides, the baseline marginal utility of good *k* now remains unchanged with the consumption of good *k*, which is intuitive. Indeed, the alternative of including θ_{kk} leads to the rather strange notion that the marginal valuation at the point of no consumption of a good varies with the consumption level of the good. Overall, the resulting modified form of the VH-UF with $\theta_{kk} = 0$ is much more general than the B-ASUF, while improving parameter interpretation compared to the VH-UF, allowing a more intuitive marginal utility structure, and removing the possibility of parabolic shapes and increasing marginal utilities. However, the restriction that the baseline marginal utility of all goods should be positive for all consumption bundles ($\tilde{\pi}_k > 0$, $k = 1,2,...,K$) still needs to be maintained. The only way this condition will hold globally is if

$\theta_{km} \geq 0$ for all *k* and *m* (see Equation (4)). The condition $\theta_{km} > 0$ implies that the goods *k* and *m* are complements (since the consumption of good *m* would increase the baseline marginal utility of good *k* and therefore consumption of good *k*). But we would also like to allow rich substitution patterns in the utilities of goods by allowing $\theta_{km} < 0$ for some pairs of goods (this refers to substitution patterns beyond that implied by the presence of a budget constraint). As we discuss later, our methodology accommodates this, while also recognizing the constraint $\tilde{\pi}_k > 0$ ($k = 1,2,...,K$) during estimation and ensuring that it holds in the range of consumptions observed in the data.

To summarize, we propose the following general and modified formulation for the non-additively separable utility form (or NASUF for short) when there are no essential goods:

$$U(\boldsymbol{x}) = \sum_{k=1}^{K} \psi_k \frac{\gamma_k}{\alpha_k}\left[\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right] + \frac{1}{2}\sum_{k=1}^{K}\sum_{m \neq k} \theta_{km}\left(\frac{\gamma_k}{\alpha_k}\left[\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right]\right)\left(\frac{\gamma_m}{\alpha_m}\left[\left(\frac{x_m}{\gamma_m}+1\right)^{\alpha_m} - 1\right]\right) \quad (6)$$

Note that this is different from the VH-UF (Equation (2)) in the absence of the θ_{kk} terms. Further, as discussed earlier, the analyst will need to estimate the γ-profile or the α-profile. The γ-profile of our NASUF takes the following form:

$$U(\boldsymbol{x}) = \sum_{k=1}^{K} \psi_k \gamma_k \ln\left(\frac{x_k}{\gamma_k}+1\right) + \frac{1}{2}\sum_{k=1}^{K}\sum_{m \neq k} \theta_{km}\, \gamma_k \gamma_m \ln\left(\frac{x_k}{\gamma_k}+1\right) \ln\left(\frac{x_m}{\gamma_m}+1\right), \quad (7)$$

and the α-profile takes the following form:

$$U(\boldsymbol{x}) = \sum_{k=1}^{K} \frac{\psi_k}{\alpha_k}\left[(x_k+1)^{\alpha_k} - 1\right] + \frac{1}{2}\sum_{k=1}^{K}\sum_{m \neq k} \frac{\theta_{km}}{\alpha_k \alpha_m}\left[(x_k+1)^{\alpha_k} - 1\right]\left[(x_m+1)^{\alpha_m} - 1\right]. \quad (8)$$

In the case that a γ-profile is estimated, the γ_k values need to be greater than zero, which can be maintained by reparameterizing γ_k as $\exp(\kappa_k)$. Additionally, the translation parameters can be allowed to vary across individuals by writing $\kappa_k = \tilde{\boldsymbol{\kappa}}_k' \boldsymbol{w}_k$, where $\boldsymbol{w}_k$ is a vector of individual characteristics for the k^{th} alternative, and $\tilde{\boldsymbol{\kappa}}_k'$ is a corresponding vector of parameters. In the case when a α-profile is estimated, the α_k 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 $\delta_k = \tilde{\boldsymbol{\delta}}_k' \boldsymbol{y}_k$, where $\boldsymbol{y}_k$ is a vector of individual characteristics impacting satiation for the k^{th} alternative, and $\tilde{\boldsymbol{\delta}}_k$ is a corresponding vector of parameters. In actual application, it would behoove the analyst to estimate models based on both the estimable profiles above, and choose the one that provides a better statistical fit. In the rest of this paper, we will use the general form (that is, the NASUF of Equation (6)) for the "no-essential good" case for ease in presentation.

Thus far, the discussion has assumed that there is no essential good. If an essential (outside) good is present, label this essential good as the first good which now has a unit price of one (*i.e.*, $p_1 = 1)$. This good, being an essential good, has no interaction term effects with the inside goods; *i.e.*, $\theta_{1m} = 0 \; \forall m \, (m \neq 1)$. Our NASUF utility functional form for this essential goods case takes the following structure:

$$U(\boldsymbol{x}) = \frac{1}{\alpha_1}(x_1 + \gamma_1)^{\alpha_1}\psi_1 + \sum_{k=2}^{K}\psi_k \frac{\gamma_k}{\alpha_k}\left[\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right] + \frac{1}{2}\sum_{k=2}^{K}\sum_{\substack{m\neq k \\ m\neq 1}}\theta_{km}\left(\frac{\gamma_k}{\alpha_k}\left[\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right]\right)\left(\frac{\gamma_m}{\alpha_m}\left[\left(\frac{x_m}{\gamma_m}+1\right)^{\alpha_m} - 1\right]\right) \qquad (9)$$

In the above formula, we need $\gamma_1 \leq 0$, while $\gamma_k > 0$ for $k > 1$. Also, we need $x_1 + \gamma_1 > 0$. The magnitude of γ_1 may be interpreted as the required lower bound (or a "subsistence value") for consumption of the essential good. As in the "non-essential goods only" case, the analyst will have to use either an α-profile or a γ-profile, though we will use the general form above for ease in presentation. For identification, we impose the condition that $\psi_1 = 1$.

3. THE ECONOMETRIC MODEL

We first consider the "no-essential" good setting, because the econometrics is more cumbersome in this case. When an essential (outside) good is also present, the econometrics simplify considerably.

3.1. Optimal Consumption Allocations

The consumer maximizes utility $U(\boldsymbol{x})$ as provided by Equation (6) subject to the budget constraint that $\sum_{k=1}^{K} p_k x_k = E$, where p_k is the unit price of good k and E is total expenditure across all goods. The analyst can solve for the optimal consumption allocations by forming the Lagrangian and applying the KKT conditions. The Lagrangian function for the problem is:

$$L = U(\boldsymbol{x}) - \lambda \left[\sum_{k=1}^{K} x_k p_k - E \right], \tag{10}$$

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

$$\frac{\partial U(\boldsymbol{x})}{\partial x_k^*} - \lambda p_k = 0, \text{ if } x_k^* > 0, \ k = 1,2,...,K \tag{11}$$

$$\frac{\partial U(\boldsymbol{x})}{\partial x_k^*} - \lambda p_k < 0, \text{ if } x_k^* = 0, \ k = 1,2,...,K .$$

The precise form of the KKT conditions depends on how stochasticity is introduced in the model, and determines the model structure (note that the discussions in Section 2 were based on the assumption of a deterministic utility function).

3.2. Introducing Stochasticity in the Additively Separable (AS) Case

To complete the econometric model, the analyst needs to introduce stochasticity. This is an important component of the model formulation. In the B-ASUF (and in other restricted versions of this formulation), stochasticity is introduced using the following random specification:

$$U(\boldsymbol{x}) = \sum_k \frac{\gamma_k}{\alpha_k} \left[\psi(\boldsymbol{z}_k) \exp(\varepsilon_k) \right] \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\}, \tag{12}$$

where $\boldsymbol{z}_k$ is a set of attributes characterizing alternative k and the decision maker, and the ε_k terms are independent and identically distributed (IID) across alternatives. ε_k captures idiosyncratic (unobserved) characteristics that impact the baseline utility for good k (the above stochastic utility form is equivalent to assuming a stochastic baseline (marginal) utility function given by $\psi(\boldsymbol{z}_k)\exp(\varepsilon_k)$). The exponential form for the introduction of the random term guarantees the positivity of the baseline marginal utility as long as $\psi(\boldsymbol{z}_k) > 0$. To ensure this latter condition, $\psi(\boldsymbol{z}_k)$ is further parameterized as $\exp(\boldsymbol{\beta}'\boldsymbol{z}_k)$, where $\boldsymbol{\beta}$ is a vector of parameters. The KKT conditions corresponding to the random utility functional form of Equation (12) are thus stochastic and take the following form:

$$\psi(\boldsymbol{z}_k)\exp(\varepsilon_k)\left(\frac{x_k^*}{\gamma_k}+1\right)^{\alpha_k-1} - \lambda p_k = 0, \text{ if } x_k^* > 0,\ k = 1,2,...,K \tag{13}$$

$$\psi(\boldsymbol{z}_k)\exp(\varepsilon_k)\left(\frac{x_k^*}{\gamma_k}+1\right)^{\alpha_k-1} - \lambda p_k < 0, \text{ if } x_k^* = 0,\ k = 1,2,...,K\ .$$

According to this approach, any stochasticity in the KKT conditions originates from the analyst's inability to observe all factors relevant to the consumer's utility formation. Individuals are assumed to know all relevant factors impacting choice, and make an error-free maximization of overall utility (subject to the budget constraint) to determine their consumption patterns (this is the random utility-deterministic maximization or RU-DM decision postulate, as already discussed).

Note, however, that the stochastic KKT conditions above of the AS model could as well have been obtained using a deterministic utility specification (rather than a random utility specification) as follows (we will occasionally write $\psi(\boldsymbol{z}_k)$ as ψ_k):

$$U(\boldsymbol{x}) = \sum_k \frac{\gamma_k}{\alpha_k}\psi_k \left\{\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right\}. \tag{14}$$

The KKT conditions corresponding to the above form are also deterministic (the conditions are identical to Equation (13), without the presence of the term $\exp(\varepsilon_k)$). But stochasticity can then be introduced explicitly in the KKT conditions in a multiplicative exponential form to once again obtain Equation (13). According to this view, not only is the consumer aware of all factors relevant to utility formation, but the analyst observes all of these factors too. However, consumers are assumed to make random mistakes ("errors") in maximizing utility (subject to the budget constraint), which gets manifested in the form of stochasticity in the KKT conditions (this is the deterministic utility-random maximization or DU-RM decision postulate; though they do not characterize this perspective as the DU-RM postulate, Wales and Woodland explicitly identify this alternative perspective for KKT models – see footnote 5 in their paper, page 268).[6] While the DU-RM postulate is seldom used for KKT models in the econometric literature, it certainly is a plausible one that should not be summarily dismissed. It also allows the usual computations of compensating variation for welfare analysis (a common reason for modeling consumer preferences) as does the RU-DM postulate.

In the AS case, both the DU-RM and RU-DM decision postulates lead to exactly the same model (further, when the error terms ε_k are assumed to be extreme value and independently and identically distributed (IID) across alternatives, the resulting model collapses to the surprisingly simple MDCEV model after using a logarithm transformation on the KKT conditions of Equation (13), as illustrated by Bhat, 2008). Since the two postulates are empirically indistinguishable, one can use either postulate to motivate the model. However, this ceases to be the case when moving from the AS utility form to the non-additively separable (NAS) utility functional form. In the next two sections, we discuss the DU-RM and RU-DM formulations, and show how a formulation that combines these two formulations in a random utility-random maximization (RU-RM) decision postulate is particularly convenient and general for the NAS case.

3.2.1 The DU-RM non-additively separable (NAS) utility formulation and model

Consider the NASUF of Equation (6). For this deterministic utility form, the corresponding deterministic KKT conditions are:

$$\tilde{\pi}_k \left(\frac{x_k^*}{\gamma_k} + 1 \right)^{\alpha_k - 1} - \lambda p_k = 0, \text{ if } x_k^* > 0, \; k = 1,2,...,K \tag{15}$$

[6] In the DU-RM viewpoint, the introduction of stochasticity within the KKT conditions has nothing to do with the random utility structure in Equation (12), and can be technically done in many different ways as long as the values of the random elements are admissible. In this regard, the consideration of stochasticity in the KKT conditions by including $\exp(\varepsilon_k)$ in the leading term on the left side of Equation (13) (with ε_k spanning the real line) is helpful, because the KKT conditions can be rewritten in an equivalent logarithmic form with ε_k appearing linearly in the transformed conditions. On the other hand, for example, including ε_k directly in the leading term on the left side of Equation (13) would result in a logarithm version of the KKT conditions that would involve $\ln(\varepsilon_k)$, which, if ε_k spanned the real line, would result in inadmissible maximization errors. Thus, even in the subsequent NAS forms to be discussed later, we will maintain the exponential form of introduction of maximization errors in the KKT conditions.

$$\tilde{\pi}_k \left(\frac{x_k^*}{\gamma_k} + 1 \right)^{\alpha_k - 1} - \lambda p_k < 0, \text{ if } x_k^* = 0, \; k = 1,2,...,K,$$

where $\tilde{\pi}_k$ is the baseline marginal utility as provided in Equation (4). Stochasticity may be introduced explicitly in the KKT conditions in the usual multiplicative exponential form as follows (again, other ways to introduce stochasticity in the KKT conditions):

$$\tilde{\pi}_k \exp(\varepsilon_k) \left(\frac{x_k^*}{\gamma_k} + 1 \right)^{\alpha_k - 1} - \lambda p_k = 0, \text{ if } x_k^* > 0, \; k = 1,2,...,K \tag{16}$$

$$\tilde{\pi}_k \exp(\varepsilon_k) \left(\frac{x_k^*}{\gamma_k} + 1 \right)^{\alpha_k - 1} - \lambda p_k < 0, \text{ if } x_k^* = 0, \; k = 1,2,...,K.$$

Note that, unlike in the AS case, one cannot develop an analytic random utility specification that corresponds to the KKT stochastic conditions in the equation above. This is an important issue, because a random utility formulation should typically start from the introduction of stochasticity, conceptualized in a specific way.

The optimal demand satisfies the conditions in Equation (16) plus the budget constraint. The structure is now exactly the same as the model of Bhat (2005, 2008). Specifically, consider an extreme value distribution for ε_k and assume that ε_k is independent of ψ_k, γ_k, and α_k ($k = 1,2,...,K$). The ε_k terms are also assumed to be IID across alternatives with a scale parameter of σ (σ can be normalized to one if there is no variation in unit prices across goods; see Bhat, 2008 for a detailed discussion of identification issues). In this case, the probability expression collapses to the following MDCEV closed-form:

$$P\left(x_1^*, x_2^*, x_3^*, ..., x_M^*, 0, 0, ..., 0\right) = |J_M| \frac{1}{\sigma^{M-1}} \left[\frac{\prod_{i=1}^{M} e^{V_i/\sigma}}{\left(\sum_{k=1}^{K} e^{V_k/\sigma} \right)^M} \right] (M-1)!, \tag{17}$$

where $V_k = \ln(\tilde{\pi}_k) - \ln p_k + (\alpha_k - 1)\ln\left(\frac{x_k^*}{\gamma_k} + 1 \right)$ ($k = 1,2,...,K$), and the elements of the Jacobian J_M are given by:

$$J_{ih} = \frac{\partial[V_1 - V_{i+1} + \varepsilon_1]}{\partial x_{h+1}^*} = \frac{\partial[V_1 - V_{i+1}]}{\partial x_{h+1}^*}, \; i = 1,2,...,M-1; \; h = 1,2,...,M-1, \tag{18}$$

where the first alternative is an alternative to which the consumer allocates some non-zero budget amount (note that the consumer should allocate budget to at least one alternative, given that the total expenditure across all alternatives is a positive quantity). To write these Jacobian elements, define $z_{ih}=1$ if $i=h$, and $z_{ih}=0$ if $i\neq h.$ Also, define the following:

$$\omega_k=\frac{1}{p_k}\left(\frac{x_k^*}{\gamma_k}+1\right)^{\alpha_k-1}\text{ for }k=1,2,...,K. \tag{19}$$

Then, the elements of the Jacobian can be derived to be:[7]

$$J_{ih}=\omega_{h+1}p_{h+1}\left[\frac{\theta_{1,h+1}}{\tilde{\pi}_1}-\left(1-z_{ih}\right)\frac{\theta_{i+1,h+1}}{\tilde{\pi}_{i+1}}\right]+\omega_1 p_1^2\frac{\theta_{1,i+1}}{\tilde{\pi}_{i+1}}+p_{h+1}\left[L_1p_1+z_{ih}L_{i+1}\right], \tag{20}$$

where $L_k=\dfrac{1-\alpha_k}{p_k(x_k^*+\gamma_k)}$. Unfortunately, there is no concise form for the determinant of the Jacobian for $M>1$ (unlike the case of the additively separable case, where Bhat derived a simple form for any value of *M*). When $M=1$ (*i.e.*, only one alternative is chosen) for all individuals, there are no satiation effects ($\alpha_k=1$ for all *k*), $\theta_{km}=0\ \forall k,m\ (k\neq m)$ 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 (17) collapses to the standard MNL model.

In estimating the DU-RM model, as discussed in Section 2.1.1, we should ensure $\tilde{\pi}_k>0$ for each good *k*. This is recognized in the logarithmic transformation of $\tilde{\pi}_k$ appearing in V_k. At the same time, we also require that $\psi_k>0$, which is ensured (as in the AS case) by writing $\psi_k=\exp(\boldsymbol{\beta}'\boldsymbol{z}_k)$. Also, since only differences in the V_k from V_1 matters in the KKT conditions, a constant cannot be identified in the term for one of the *K* alternatives. Similarly, individual-specific variables are introduced in the V_k's for (*K*-1) alternatives, with the remaining alternative serving as the base (these are the equivalent of the identification condition of $\sum_k \psi_k=1$ usually imposed in flexible function forms). The parameters in the DU-RM NAS-based model may be estimated in a straightforward way using the maximum likelihood inference approach. However, it is difficult to motivate generalized extreme value error structures and variable-specific random

[7] The derivation is rather straightforward, but requires some cumbersome differentiation. Interested readers may obtain the derivation from the authors.

coefficients in the context of the DU-RM formulation. These extensions, however, are quite natural in the context of the RU-DM decision postulate, which we discuss in the next section.

For the DU-RM formulation with an essential outside good, the econometrics simplify considerably. One can go through the same procedure as earlier by writing the KKT conditions and introducing stochasticity corresponding to the deterministic utility expression in Equation (9) instead of Equation (6). For the outside good (say, the first alternative), we have the following: $\boldsymbol{\beta}'\boldsymbol{z}_1 = 0,\ \psi_1 = 1,$ and $p_1 = 1.$ The final expression for probability in this outside good case is the same as in Equation (17) with the following modifications to the V_k terms:

$$V_k = \ln(\tilde{\pi}_k) - \ln p_k + (\alpha_k - 1)\ln\left(\frac{x_k^*}{\gamma_k} + 1\right)\ (k > 2);\ V_1 = (\alpha_1 - 1)\ln(x_1^* + \gamma_1)\,. \tag{21}$$

The Jacobian elements in this case simplify relative to Equation (20), with $\theta_{1m} = 0\ \forall m\,(k \neq 1)$. The elements now are given as follows:

$$J_{ih} = \omega_{h+1}\left[-(1 - z_{ih})\frac{\theta_{i+1,h+1}}{\tilde{\pi}_{i+1}}\right] + p_{h+1}\left[L_1 + z_{ih}L_{i+1}\right]. \tag{22}$$

3.2.2. The RU-DM non-additively separable (NAS) utility formulation and model

Consider the following stochastic NASUF for the no-essential good case:

$$U(\boldsymbol{x}) = \sum_{k=1}^{K}\psi_k \exp(\xi_k)\frac{\gamma_k}{\alpha_k}\left[\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right] + \frac{1}{2}\sum_{k=1}^{K}\sum_{m \neq k}\theta_{km}\left(\frac{\gamma_k}{\alpha_k}\left[\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right]\right)\left(\frac{\gamma_m}{\alpha_m}\left[\left(\frac{x_m}{\gamma_m}+1\right)^{\alpha_m} - 1\right]\right) \tag{23}$$

where ξ_k is an IID (across alternatives) random error term with a scale parameter of σ (σ can be normalized to one if there is no variation in the unit prices across alternatives). ξ_k captures idiosyncratic (unobserved) characteristics that impact the baseline (marginal) utility of good *k* at the point at which no expenditure outlays have yet been made on any alternative.[8] The KKT conditions then are:

---

[8] Vásquez-Lavín and Hanemann instead write the utility function as:
$U(\boldsymbol{x}) = \sum_{k=1}^{K}(\psi_k + \xi_k)\frac{\gamma_k}{\alpha_k}\left[\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right] + \frac{1}{2}\sum_{k=1}^{K}\sum_{m \neq k}\theta_{km}\left(\frac{\gamma_k}{\alpha_k}\left[\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right]\right)\left(\frac{\gamma_m}{\alpha_m}\left[\left(\frac{x_m}{\gamma_m}+1\right)^{\alpha_m} - 1\right]\right)$, and use an error distribution for ξ_k that spans the entire real line (specifically, a normal or an extreme value distribution). The problem with this structure is that it allows negative values for the baseline utility $(\psi_k + \xi_k)$ at every consumption point, which is theoretically inappropriate since this term has to be positive for $U(\boldsymbol{x})$ to be a valid utility function. It

$$\eta_k \left(\frac{x_k^*}{\gamma_k} + 1 \right)^{\alpha_k - 1} - \lambda p_k = 0 \text{, if } x_k^* > 0 \text{, } k = 1,2,...,K \tag{24}$$

$$\eta_k \left(\frac{x_k^*}{\gamma_k} + 1 \right)^{\alpha_k - 1} - \lambda p_k < 0 \text{, if } x_k^* = 0 \text{, } k = 1,2,...,K \text{,}$$

where $\eta_k = \psi_k \exp(\xi_k) + W_k$ and $W_k = \sum_{m \neq k} \theta_{km} \frac{\gamma_m}{\alpha_m} \left[\left(\frac{x_m}{\gamma_m} + 1 \right)^{\alpha_m} - 1 \right]$.

Define ω_k as in Equation (19). Let $R_k = \eta_1 \cdot \frac{\omega_1}{\omega_k} - W_k$ and $\psi_k = \exp(\boldsymbol{\beta}' \boldsymbol{z}_k)$, and let the first alternative be the one to which the consumer allocates some non-zero budget amount. Then, the KKT conditions may be simplified as follows:

$$\exp(\xi_k) = \frac{R_k \mid \xi_1}{\exp(\boldsymbol{\beta}' \boldsymbol{z}_k)} \text{, if } x_k^* > 0 \text{, } k = 2,3,...,K \tag{25}$$

$$\exp(\xi_k) < \frac{R_k \mid \xi_1}{\exp(\boldsymbol{\beta}' \boldsymbol{z}_k)} \text{, if } x_k^* = 0 \text{, } k = 2,3,...,K \text{.}$$

Next, let $\zeta_k = \exp(\xi_k)$, and assume that $g(.)$ and $G(.)$ are the standardized versions of the probability density function and standard cumulative distribution function characterizing ζ_k. Then, the probability that the individual allocates expenditure to the first M of the K goods may be derived to be:

$$P\left(x_1^*, x_2^*, x_3^*, ..., x_M^*, 0, 0, ..., 0 \right)$$
$$= \int_{\zeta_1 = 0}^{\zeta_1 = +\infty} | J_M \mid \zeta_1 | \left\{ \left(\prod_{i=2}^{M} \frac{1}{\sigma} g\left[\frac{1}{\sigma} \frac{R_k \mid \zeta_1}{\exp(\boldsymbol{\beta}' \boldsymbol{z}_k)} \right] \right) \right\} \times \left\{ \prod_{s=M+1}^{K} G\left[\frac{1}{\sigma} \frac{R_k \mid \zeta_1}{\exp(\boldsymbol{\beta}' \boldsymbol{z}_k)} \right] \right\} f(\zeta_1) d\zeta_1 \text{,} \tag{26}$$

where $f(.)$ refers to the density function characterizing ζ_1, and $J_M \mid \xi_1$ is the Jacobian whose elements are given by ($i, h = 1,2,..., M - 1$):

is not at all clear to this author why they use this approach for the "no essentials good case", but resort to the use of the exponential of ξ_k (as we do in Equation 23) for the case of the presence of an essential good. As we show, the derivation of the probability structure is feasible and tractable even in the case of the use of the more theoretically appropriate Equation (23). Further, we are able to obtain a conditionally closed-form expression for the Jacobian, which allows us to estimate the model very quickly (VH suggested that "there is no simple solution for the Jacobian that can be generalized to any consumption pattern", which is not the case).

$$J_{ih} = \frac{\partial}{\partial x^*_{h+1}}\left[\frac{R_{i+1} \mid \zeta_1}{\exp(\boldsymbol{\beta}'\boldsymbol{z}_{i+1})}\right] = \frac{1}{\exp(\boldsymbol{\beta}'\boldsymbol{z}_{i+1})}\left\{\frac{\omega_1}{\omega_{i+1}}\left[(\eta_1 \mid \zeta_1)(p_1^2 L_1 + p_{h+1}L_{i+1}z_{ih}) + p_{h+1}\theta_{1,h+1}\omega_{h+1}\right] \right.$$
$$\left. + p_{h+1}\left[p_1\theta_{1,h+1}\omega_1 - \theta_{i+1,h+1}\omega_{h+1}(1-z_{ih})\right] \right\}. \qquad (27)$$

In the above expression, $z_{ih} = 1$ if $i = h$, and $z_{ih} = 0$ if $i \neq h$ and $L_k = \dfrac{1-\alpha_k}{p_k(x^*_k + \gamma_k)}$ ($k = 1,2,...,K$).

The probability expression in Equation (26) is a simple one-dimensional integral, which can be computed using quadrature techniques. Note that the distribution for ξ_k can be any univariate distribution, though the normal distribution may be convenient if there are also random normal coefficients in the $\boldsymbol{\beta}$ vector to capture unobserved individual heterogeneity (then the one-dimensional normal integral becomes simply a part of a multi-dimensional normal integration that can be evaluated using familiar simulation techniques). Such a random-coefficients specification allows a flexible covariance structure between the elements of the $\boldsymbol{\beta}$ vector, and can also include covariances among the baseline utilities of alternatives (as in a mixed multinomial logit structure). The model may be estimated using traditional maximum likelihood techniques, as for the DU-RM formulation. Note, however, that the marginal utility of 1any good at any point of consumption should be positive (for increasing utility functions). This condition is met by setting $\eta_k > 0$ (see Equation (24)) for each good *k*.

When an essential good is present, the econometrics again simplify considerably. For the essential good (say, the first alternative), we have the following: $W_1 = 0$, $\boldsymbol{\beta}'\boldsymbol{z}_1 = 0$, $\psi_1 = 1$, $p_1 = 1$, and $\eta_1 = \zeta_1 = \psi_1 \exp(\xi_1)$. The stochasticity is introduced here similar to VH (2008). The random utility function in this case originates from Equation (9) and takes the following form (again, this is not the same as the VH-UF, because the VH-UF retains the θ_{kk} ($k = 2,3,...,K$) terms, while our NASUF removes these terms for the reasons mentioned earlier):

$$U(\boldsymbol{x}) = \frac{1}{\alpha_1}(x_1 + \gamma_1)^{\alpha_1}\ \eta_1$$
$$\sum_{k=2}^{K}\psi_k \exp(\xi_k)\frac{\gamma_k}{\alpha_k}\left[\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right] + \frac{1}{2}\sum_{k=2}^{K}\sum_{m\neq k}\theta_{km}\left(\frac{\gamma_k}{\alpha_k}\left[\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right]\right)\left(\frac{\gamma_m}{\alpha_m}\left[\left(\frac{x_m}{\gamma_m}+1\right)^{\alpha_m} - 1\right]\right). \qquad (28)$$

The probability expression takes the same form as in Equation (26) with the following modifications to the ω_k terms:

$$\omega_k = \frac{1}{p_k}\left(\frac{x_k^*}{\gamma_k}+1\right)^{\alpha_k-1} \text{ for } k=2,...,K; \;\; \omega_1 = (x_1^*+\gamma_1)^{\alpha_1-1}. \tag{29}$$

The Jacobian elements again can be computed in closed form and are as follows ($i,h=1,2,...,M-1$):

$$J_{ih} = \frac{1}{\exp(\boldsymbol{\beta}'\boldsymbol{z}_{i+1})}\left\{\frac{\omega_1}{\omega_{i+1}}\left[(\eta_1 \mid \zeta_1)(L_1+p_{h+1}L_{i+1}z_{ih})\right]-p_{h+1}\theta_{i+1,h+1}\omega_{h+1}(1-z_{ih})\right\}. \tag{30}$$

3.2.3. The RU-RM non-additively separable (NAS) utility formulation and model

Consider the random utility function of Equation (23) for the case with no essential good. The KKT conditions are given by Equation (24), but we now add stochasticity originating from consumer mistakes in the optimizing process. The KKT conditions take the form shown below:

$$\eta_k \exp(\varepsilon_k)\left(\frac{x_k^*}{\gamma_k}+1\right)^{\alpha_k-1} - \lambda p_k = 0, \text{ if } x_k^*>0\,,\; k=1,2,...,K \tag{31}$$

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

where η_k is as defined earlier in Equation (24) (and has the error term ξ_k embedded within), and the ε_k terms are independent and identically (across alternatives) extreme value distributed. Recall that the ξ_k terms represent stochasticity due to the analyst's inability to capture consumer preferences, while the ε_k terms represent stochasticity due to consumer errors in utility maximization. Let $Var(\varepsilon_k)+Var(\xi_k)=(\pi^2\sigma^2)/6 \;(k=1,2,...K)$.[9] In the RU-RM formulation, we assume that the ξ_k terms are normally distributed. This is particularly convenient when one wants to accommodate a flexible error covariance structure through a multivariate normal-distributed coefficient vector $\boldsymbol{\beta}$ and/or account for covariance in utilities across alternatives through the appropriate random multivariate specification for the ξ_k terms. To develop the probability function for consumptions, let $Var(\varepsilon_k)=\mu^2(\pi^2\sigma^2)/6$ and

[9] As earlier, we will impose the normalization that $\sigma^2=1$ if there is no price variation across the alternatives.

$Var(\xi_k) = (1-\mu^2)(\pi^2\sigma^2)/6 \;\; (k = 1,2,...,K)$, where μ is a parameter to be estimated $(0 \le \mu \le 1)$. Then, if $\mu \to 0$, and when there is no covariance among the ξ_k terms across alternatives, the RU-RM formulation approaches the RU-DM formulation of Section 3.2.2 in which the scale parameter σ is innocuously rescaled to $(\pi/\sqrt{6})\,\sigma$, so that the variance of the error terms ξ_k in the RU-DM formulation is comparable to the variance of the corresponding terms in the RU-RM formulation. However, as $\mu \to 1$, the RU-RM formulation approaches the DU-RM formulation. Thus, the parameter μ determines the extent of the mix of the RU-DM and DU-RM decision postulates leading up to the observed behavior of consumers. One can impose the constraint that $0 \le \mu \le 1$ through the use of a logistic transform $\mu = 1/(1+\exp(-\mu^*))$ and estimate the parameter μ^*.

The probability expression for consumptions in the RU-RM model formulation takes the following mixed MDCEV form:

$$P\left(x_1^*, x_2^*, x_3^*, ..., x_M^*, 0, 0, ..., 0\right) = \int_{\xi=-\infty}^{\infty} \left[|J_M \mid \xi| \frac{1}{(\mu\sigma)^{M-1}} \left[\frac{\prod_{i=1}^{M} e^{[V_i/(\mu\sigma)]|\xi_i}}{\left(\sum_{k=1}^{K} e^{[V_k/(\mu\sigma)]|\xi_k}\right)^M} \right] (M-1)! \right] dF(\xi), \quad (32)$$

where $V_k = \ln(\eta_k) - \ln p_k + (\alpha_k - 1)\ln\left(\frac{x_k^*}{\gamma_k} + 1\right)$, $\eta_k = \psi_k \exp(\xi_k) + W_k$, W_k is defined as earlier, and F is the multivariate normal distribution of the random element vector $\boldsymbol{\xi} = (\xi_1, \xi_2, ..., \xi_K)$ (each of whose elements has a variance of $(1-\mu^2)(\pi^2\sigma^2)/6$). The elements of the Jacobian are given by:

$$J_{ih} = \omega_{h+1} p_{h+1} \left[\frac{\theta_{1,h+1}}{(\eta_1 \mid \xi_1)} - \left(1 - z_{ih}\right) \frac{\theta_{i+1,h+1}}{(\eta_{i+1} \mid \xi_{i+1})} \right] + \omega_1 p_1^2 \frac{\theta_{1,i+1}}{(\eta_{i+1} \mid \xi_{i+1})} + p_{h+1}\left[L_1 + z_{ih} L_{i+1}\right]. \quad (33)$$

When there is an essential outside good, the probability expression remains the same as in Equation (33), but with $V_k = \ln(\eta_k) - \ln p_k + (\alpha_k - 1)\ln\left(\frac{x_k^*}{\gamma_k} + 1\right) \;\; (k > 2)$, $V_1 = (\alpha_1 - 1)\ln(x_1^* + \gamma_1)$, $\theta_{1m} = 0 \; \forall m \; (m \ne 1)$, $W_1 = 0$, $\boldsymbol{\beta}'\boldsymbol{z}_1 = 0$, $\psi_1 = 1$, $p_1 = 1$, and $\eta_1 = \exp(\xi_1)$. The Jacobian elements in this case are given as follows:

$$J_{ih} = \omega_{h+1}\left[-(1-z_{ih})\frac{\theta_{i+1,h+1}}{(\eta_{i+1} \mid \xi_{i+1})}\right] + p_{h+1}\left[L_1 + z_{ih}L_{i+1}\right]. \tag{34}$$

Similar to the earlier two formulations, the theoretical condition that the marginal utility of consumption for any alternative should always be positive must be ensured during model estimation. Thus, we should ensure $\eta_k > 0$ for each good *k*.

4. EMPIRICAL DEMONSTRATION

4.1. The Context

In 2010, transportation expenses accounted for nearly 20% of total household expenses and 12-15% of total household income (U.S. Bureau of Labor Statistics, 2012). In fact, this is the second largest family expense category after housing, with an average expenditure of $7,677 per year (or, equivalently, about $650 per month). It is little surprise, therefore, that the study of transportation expenditures has been of much interest in recent years (Gicheva *et al.*, 2007, Cooper, 2005, Hughes *et al.*, 2006, Thakuriah and Liao, 2006, Choo *et al.*, 2007a,b, Sanchez *et al.*, 2006). Several of these studies examine the factors that influence total household transportation expenditures and/or examine transportation expenditures in relation to expenditures on other commodities and services (such as in relation to housing, telecommunications, groceries, and eating out). But there has been relatively little research on identifying the many disaggregate-level components of transportation expenditures, with all transportation expenditures usually lumped into a single category. Besides, many of these earlier efforts use the almost ideal demand system (AIDS) proposed by Deaton and Muelbauer (1980), which assumes that all families expend their budgets in all possible expenditure categories (that is, the AIDS model does not allow corner solutions, as does our proposed model).

In the current paper, we demonstrate the use of the proposed model for an empirical case of household transportation expenditures in six disaggregate categories: (1) Vehicle purchase, (2) Gasoline and motor oil (termed as gasoline in the rest of the document), (3) Vehicle insurance, (4) Vehicle operation and maintenance (labeled as vehicle maintenance from hereon), (5) Air travel, and (6) Public transportation. In addition, we consider all other household expenditures in a single “outside good” category that lumps all non-transportation expenditures, so that total transportation expenditure is endogenously determined. Households expend some positive amount on the “outside good” category, while expenditures can be zero for one or more transportation categories for some households. A non-additively separable utility form is adopted to accommodate rich substitution patterns as well as to allow complementarity among the transportation expenditure categories.

Data for the analysis is drawn from the 2002 Consumer Expenditure (CEX) Survey, which is a national level survey conducted by the U.S. Census Bureau for the Bureau of Labor Statistics (U.S. Bureau of Labor Statistics, 2003). This survey has been administered regularly since 1980 and is designed to collect information on incomes and expenditures/buying habits of households in the United States. In addition, information on individual and household socio-economic, demographic, employment and vehicle characteristics is also collected. Details of the data and sample extraction process for the current analysis are available in Ferdous *et al*. (2010). Essentially, the 109 categories of expenditure and income defined by the CEX were consolidated, defining 17 broad categories of annual expenditures (including the six categories of transportation expenditures identified in the previous paragraph). Next, the 11 non-transportation categories were all grouped into a single "outside good" category, and the proportion of total expenditures (across the six transportation categories and the "outside good" category) spent in each of the six transportation categories and the "outside" non-transportation category were constructed as the dependent variables in the analysis.

The final sample for analysis includes 4100 households. About one-quarter of the sample reports expenditures on vehicle purchase. 94% of the sample incurs expenditures on gasoline, and 90% of the sample indicates vehicle maintenance expenses (the association between these two numbers is not surprising, because almost all households that do not expend money in gasoline also do not expend money on vehicle maintenance; further, most of these same households have some positive expenditure in the public transportation category). About 80% of the sample has vehicle-insurance related expenses, suggesting that a sizeable number of households operate motor vehicles with no insurance or have insurance costs paid for them (possibly by an employer or self-employed business). About one-third of the sample reports spending money on each of the two categories of public transportation and air travel. Only 2.6% of the households expend no money in transportation-related expenses. These households may undertake trips using non-motorized modes, or rely on someone else to travel. Altogether, expenditures on transportation-related items account for about 15% of household income, a figure that is quite consistent with reported national figures. Of the 4100 households, a random sample of 3600 households was used for model estimation and the remaining sample of 500 households was held for out-of-sample validation.

4.2 Model Specification and Estimation

The additively separable (B-ASUF) and non-additively separable (NASUF) models were estimated using the GAUSS matrix programming language.[10] We first estimated the best empirical specification for the MDCEV model (assuming the B-ASUF form) based on intuitive and statistical significance considerations, and then explored alternative specifications for the interaction parameters in the NASUF model for the three stochastic formulations proposed. The γ-profile of Equation (7) was used in all specifications, since it consistently provided a better model fit than the α-profile (this γ-profile is similar to the translog-type utility form). Also, the γ_1 value for the essential good was set to zero for estimation stability.

In the absence of interactions between the sub-utility functions of different alternatives, the DU-RM formulation collapses to the simple MDCEV model, while the RU-DM formulation collapses to an AS MDC model with IID normal (or probit) error terms (label this as the MDCP for MDC probit model). Thus, for model evaluation purposes, the analyst can compare the performance of the DU-RM model to its special case MDCEV and that of the RU-DM model to its special case MDCP. The RU-RM formulation utilizes a combination of extreme value error terms and normally distributed error terms for the consumer's mistakes and the analyst's errors, respectively. Thus, for this last formulation there is no direct B-ASUF-based model for comparison purposes. However, as discussed in Section 3.2.3, the RU-DM and DU-RM formulations are limiting case of the RU-RM formulation.

The estimation of the three model formulations was undertaken to explicitly consider the constraint that the marginal utility of any good at any consumption point for each good *k* should always be positive. In the current empirical application, our attempts to use the constrained maximum likelihood module of GAUSS to estimate the models encountered estimation instability and convergence problems. Therefore, the models were estimated using the traditional maximum likelihood module of GAUSS, while checking for the positivity of the marginal utility at each iteration and heuristically updating parameters to cause the least departure from the

[10] Gauss^{TM}, Aptech Systems Inc., Chandler, AZ, USA, http://www.aptech.com.

iteration-search parameters and still ensuring positivity if positivity was not maintained (in most iterations, positivity was maintained automatically). The DU-RM NAS model was estimated imposing $\tilde{\pi}_k > 0$ for each good *k* (see Equation (6)), since the term $\tilde{\pi}_k$ is inside a logarithmic function. For the RU-DM and RU-RM NAS models, the baseline marginal utility is given by $\eta_k \left(\frac{x_k^*}{\gamma_k} + 1 \right)^{\alpha_k - 1}$. Because the term $\left(\frac{x_k^*}{\gamma_k} + 1 \right)^{\alpha_k - 1}$ is always positive, we have to constrain $\eta_k > 0$. In the estimation of the RU-DM and RU-RM NAS model formulations, we imposed the more restrictive condition $W_k > 0$ to ensure that the condition is fulfilled for all values of ξ_k (ξ_k is embedded in η_k; see Equation (24)). Quadrature techniques for log-normally distributed variables were used to evaluate the integral in Equation (26) for the RU-DM NAS model formulation (details are available from the authors). To evaluate the multivariate integral of Equation (32) for the RU-RM NAS model, we used the Halton sequence to draw realizations for $\xi = (\xi_1, \xi_2, ..., \xi_K)$ from a normal distribution, assuming in the empirical analysis that these error terms are independent and identically distributed across alternatives. Details of the Halton sequence and the procedure to generate this sequence are available in Bhat (2000, 2001, 2003). We tested the sensitivity of parameter estimates with different numbers of Halton draws per observation, and found the results to be very stable with as few as 75 draws. In this analysis we used 100 draws per household in the estimation.

4.3 Model Results

The estimation results are provided in Table 1. At the outset, we should note that the intent of this empirical analysis is not to contribute in a substantive way to an analysis of household expenditures. Rather, the emphasis is on demonstrating the applicability of the three different NASUF formulations proposed in this paper, and showing the advantage of the NASUF formulations relative to the B-ASUF formulations. To that extent, the focus is on in-sample and out-of-sample data fits of the NASUF and B-ASUF formulations, as well as on demonstrating the significant presence of NAS interaction parameters in our NASUF models.

Table 1 is organized in three main columns. The first main column provides the parameters estimates of the DU-RM NASUF model and its restrictive B-ASUF formulation (that is, the MDCEV formulation), while the second main column presents the results of the RU-DM

NAS model and its restrictive B-ASUF formulation (that is, the MDCP formulation). The third column provides the parameters estimates of the RU-RM NASUF model. As discussed in Section 3, one of the alternatives forms the base category for the introduction of the family-specific variables in the baseline utility in Table 1. This base alternative is the essential outside good, which is the non-transportation good category in the current analysis. If, in addition, some transportation categories do not appear for a variable in Table 1, it implies that these transportation categories also constitute the base expenditure category along with the non-transportation category. For example, for the effect of "Number of workers in the household", the base categories include the non-transportation category as well as the air travel and public transportation categories. A positive (negative) coefficient for a certain variable-category combination implies that an increase in the explanatory variable increases (decreases) the likelihood of budget being allocated to that expenditure category relative to the base expenditure categories.

Overall, the empirical results are intuitive. Also, while there are differences in the estimated coefficients between the AS and NAS models, the general pattern and direction of variable effects are similar. Regarding the baseline parameters ($\boldsymbol{\beta}$), the alternative specific constants in the baseline utility for all the transportation categories are negative, indicating the generally higher baseline utility of the "outside" non-transportation good category relative to each transportation category (this is a reflection of the higher expenditure on the outside good than on the transportation categories). Similar to the results found by Thakuriah and Liao (2005), as the number of workers in the household increases, so does the proportion of income allocated to all vehicle-related transportation expenses, presumably to support the transportation needs of multi-worker households (an exception is in the RU-DM model, in which the coefficient associated with vehicle insurance is negative but statistically insignificant). The effect of income was considered in a continuous linear form, in a piecewise linear form to introduce non-linearities, as well as in the form of dummy variables for specific income categories. At the end, a dummy variable specification with low income (less than 30K), mid-range income (30-70K), and high income (>70K) provided the best results. The effect of this discrete representation of income is incorporated with the low income category constituting the base category (and so the low income category does not appear in Table 1). The results indicate that, relative to families in the low income group, families in the middle and high income groups expend a higher proportion of their income on vehicle purchases and air travel. These families also spend a lower proportion of their income on gasoline relative to the low income group, suggesting that gasoline expenditures constitute a particularly high proportion of the income budgets of low income families. A detailed discussion of this result from a social and environmental justice perspective can be found in Deka (2004). Households with more vehicles tend to allocate a larger proportion of their income to all the transportation categories, except on public transportation. Finally, non-Caucasians, those residing in urban areas, and those living in the Northeast and West regions of the U.S. spend a higher proportion on public transportation than Caucasians, those residing in non-urban areas, and those living in the South and Midwest regions of the U.S, respectively.

The satiation parameters (γ_k) in Table 1 capture the variation in the extent of non-linearity across different expenditure categories. The satiation parameter is highest for the vehicle purchase category, indicating that households are likely to allocate a large proportion of their budget to acquiring a vehicle, if they expend any money in this category. The satiation parameter is lowest for gasoline, indicating that households allocate a relatively small proportion of their overall budget in gasoline consumption.

Several interaction parameters (θ_{km}) are statistically significant in the final model specification presented in Table 1. Many of these effects are complementary effects. Thus, the interaction parameters of the DU-RM NASUF model indicate a significant complementary effect in vehicle purchase and gasoline expenditures, and in vehicle purchase and vehicle maintenance expenditures. Also, as expected, there are complementary effects in the expenditures on gasoline, vehicle insurance, and vehicle maintenance, as well as between air travel and public transportation expenditures. This last complementary effect perhaps reflects the use of public transportation to get to/from the airport and the use of public transportation at the non-home end. On the other hand, there are particularly sensitive substitution effects in gasoline and air transportation expenditures, presumably a reflection of the choice between auto travel and air transportation mode travel for long-distance trips. For the RU-DM NASUF model, only complementarity effects were statistically significant, which align with the results of the DU-RM NASUF models. The RU-RM model interaction parameters show significant complementarity effects similar to those from the DU-RM and RU-DM models, along with a strong substitution effect between vehicle purchase and public transportation expenditures. This latter substitution effect is more intuitive than the complementary effect between vehicle purchase and public transportation expenditures, as implied by the RU-DM model.

As mentioned in Section 3.2.3, the RU-RM NASUF combines the RU-DM and DU-RM postulates of consumer behavior via the parameter μ. In the current empirical analysis, we obtained $\mu = 0.379$. The parameter is statistically different from zero (with a t-stat of 58.51 as shown in Table 1) and statistically different from one (with a t-stat of 95.60). The μ parameter is closer to zero than it is to one, indicating that the predominant source of stochasticity (62%) is due to the analyst's errors in characterizing the consumer's utility function. To a lesser extent (38%), stochasticity arises also from the random "mistakes" consumers make during utility maximization.

4.4. Model Evaluation

In this section, we compare the model performance of the B-ASUF and NASUF models, both in the estimation sample of 3600 households as well as a validation sample of 500 households.

In terms of model fit in the estimation data, the log-likelihood value at convergence of the DU-RM NASUF model is -36,645, while that of the MDCEV model is -37,045. A likelihood ratio test between these two models returns a value of 799, which is larger than the chi-squared statistic value with 7 degrees of freedom at any reasonable level of significance, indicating the substantially superior fit of the DU-RM NASUF model compared to the MDCEV model. Similarly, the log-likelihood value at convergence of the RU-DM NASUF model is -35,086, while the same figure for the MDCP model is -35,269. The likelihood ratio test between the RU-DM and MCDP models is 366, which again indicates a statistically significant difference in data fit between the models. The log-likelihood value at convergence of the RU-RM NASUF model is -34,168, which is considerably higher than the corresponding value for the MDCEV and MDCP models. The RU-RM model log-likelihood is also far superior to the log-likelihood values of the DU-RM and RU-DM models, underscoring the presence of stochasticity on the part of both the analyst and the consumer. Between the DU-RM and RU-DM models, the latter performs better than the former.

To further compare the performance of the B-ASUF models (that is, the MDCEV and MDCP models) with the NASUF models, we computed an out-of-sample log-likelihood function (OSLLF) using the validation sample of 500 observations. The OSLLF is computed by computing the predictive log-likelihood in the out-of-estimation (*i.e.*, validation) sample. As indicated by Norwood *et al.* (2001), the model with the highest value of OSLLF is the preferred one, since it is most likely to generate the set of out-of-sample observations. Table 2 reports the OSLLF values for the entire validation sample (of 500 households) as well as for different socio-demographic segments within the sample. As can be observed from the first row, the OSLLF value for the DU-RM model is better than for the MDCEV model, and the OSLLF value for the RU-DM model is better than for the MDCP model. This result is also maintained, in general, for all socio-demographic segments. Also, in general, the RU-RM formulation outperforms all other formulations, except in a few isolated segments with few observations.

In summary, the data fits of the NASUF models are superior to that of the B-ASUF models in both the estimation and validation samples.

5. 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 or even complements for one another. Traditional MDC models developed in the literature assume that the marginal utility of a good is independent of the consumption amounts of other goods because of the use of an additively-separable utility form. This prevents the possibility of complementarity among goods and rich substitution patterns. The current paper develops model formulations that allow complementarity effects and richer substitution patterns than the traditional MDC models. The proposed formulations are easy to estimate and reduce inconsistency problems relative to the non-additive form used by Vásquez-Lavín and Hanemann (2008). An important consideration in such extended MDC models is the introduction of stochasticity. We introduce three different ways to incorporate stochasticity to develop three possible models for non-additively separable utility functions (NASUFs). In the first stochastic formulation, labeled as the deterministic utility–random maximization or DU-RM decision postulate, consumers are assumed to make random mistakes in maximizing utility. In the second stochastic formulation, labeled as the random utility-deterministic maximization or RU-DM decision postulate, consumers are assumed to know all relevant factors impacting their choices and make an error-free maximization of overall utility, but the analyst is not aware of all the factors influencing consumer's choice. The third stochastic formulation combines the two previous postulates into a random utility-random maximization (RU-RM) decision postulate.

The proposed model formulations should have several applications. In the current paper, we demonstrate the application of the formulations to the empirical case of household transportation expenditures in six disaggregate categories: (1) Vehicle purchase, (2) Gasoline and motor oil, (3) Vehicle insurance, (4) Vehicle operation and maintenance, (5) Air travel, and (6) Public transportation. In addition, we consider other household expenditures in a single "outside good" category that lumps all non-transportation expenditures, so that total transportation expenditure is endogenously determined. Households expend some positive amount on the "outside good" category, while expenditures can be zero for one or more transportation categories for some households. Data for the analysis is drawn from the 2002 Consumer Expenditure (CEX) Survey, which is a national level survey conducted by the US Census Bureau for the Bureau of Labor Statistics. The results of the DU-RM, RU-DM and RU-RM non-additively separable formulations suggest statistically significant complementary and substitution effects in the utilities of selected pairs of transportation categories, and show the substantially superior data fit of the proposed formulations relative to ones that assume an additively separable utility structure. The proposed non-additive separable models performed better in a validation sample as well.

In summary, the paper has successfully formulated and applied different forms of MDC models capable of handling complementarity and rich substitution patterns among alternatives. One area for further research is to develop more formal and rigorous methods to ensure the positivity of the marginal utility (the condition in equation 5) for each observation at each estimation iteration. Currently, we aided the estimation procedure by heuristically (and somewhat in an *ad hoc* manner) updating parameters to cause the least departure from the iteration-search parameters and still ensuring positivity (if positivity was not maintained automatically) of $\tilde{\pi}_k$ for each good k in the DU-RM model and η_k for each good in the RU-DM and RU-RM models.

ACKNOWLEDGEMENTS

The authors acknowledge the helpful comments of David Bunch and two anonymous reviewers on an earlier version of the paper. The first author would also like to acknowledge support from the U.S. Department of Transportation through the Data-Supported Transportation Operations and Planning (D-STOP) Tier 1 University Transportation Center and from the Alexander von Humboldt Foundation, Germany. Finally, the authors are grateful to Lisa Macias for her help in formatting this document.

REFERENCES

Ahn, J., Jeong, G., Kim, Y., 2008. A forecast of household ownership and use of alternative fuel vehicles: A multiple discrete-continuous choice approach. *Energy Economics* 30(5), 2091-2104.

Bhat, C.R., 2000. A multi-level cross-classified model for discrete response variables, *Transportation Research Part B* 34(7), 567-582.

Bhat, C.R., 2001. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model, *Transportation Research Part B* 35(7), 677-693.

Bhat, C.R., 2003. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. *Transportation Research Part B* 37(9), 837-855.

Bhat, C.R., 2005. A multiple discrete-continuous extreme value model: Formulation and application to discretionary time-use decisions. *Transportation Research Part B* 39(8), 679-707.

Bhat, C.R., 2008. The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B* 42(3), 274-303.

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

Bhat, C.R., Castro, M., Khan, M., 2013. A new estimation approach for the multiple discrete-continuous probit (MDCP) choice model. *Transportation Research Part B* 55(1), 1-22.

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

Chikaraishi, M., Zhang, J., Fujiwara, A., Axhausen, K.W., 2010. Exploring variation properties of time use behavior based on a multilevel multiple discrete-continuous extreme value model. *Transportation Research Record* 2156, 101-110.

Choo, S., Lee, T., Mokhtarian, P.L., 2007a. Do transportation and communications tend to be substitutes, complements, or neither? U.S. consumer expenditures perspective, 1984-2002. *Transportation Research Record* 2010, 121-132.

Choo, S., Lee, T., Mokhtarian, P.L., 2007b. Relationships between US consumer expenditures on communications and transportation using Almost Ideal Demand System modeling: 1984-2002. *Transportation Planning and Technology* 30(5), 431-453.

Cooper, M., 2005. The impact of rising prices on household gasoline expenditures. Consumer Federation of America, www.consumerfed.org/.

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

Deka, D., 2004. Social and environmental justice issues in urban transportation. In *The Geography of Urban Transportation*, S. Hanson and G. Giuliano (Eds.), Guilford Press, New York.

Ferdous, N., Pinjari, A.R., Bhat, C.R., Pendyala, R.M., 2010. A comprehensive analysis of household transportation expenditures relative to other goods and services: An application to United States consumer expenditure data. *Transportation* 37(3), 363-390.

Gicheva, D., Hastings, J., Villas-Boas, S., 2007. Revisiting the income effect: Gasoline prices and grocery purchases. Working Paper No. 13614, National Bureau of Economic Research, Cambridge, MA.

Holt, M.T., Goodwin, B.K., 2009. Chapter 2 The almost ideal and translog demand systems. In *Quantifying Consumer Preferences (Contributions to Economic Analysis, Volume 288)*, D.J. Slottje (Ed.), Emerald Group Publishing Limited, 37-59.

Hughes, J.E., Knittel, C.R., Sperling, D., 2006. Evidence of a shift in the short-run price elasticity of gasoline demand. *The Energy Journal* 29(1), 93-114.

Jorgenson, D.W., Lau, L.J., 1975. The structure of consumer preferences. *Annals of Economic and Social Measurement* 4(1), 49-101.

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

Lee, S., Allenby, G.M., 2009. A direct utility model for market basket data. *Fisher College of Business Working Paper*, 1443390. Available at: http://dx.doi.org/10.2139/ssrn.1443390.

Lee, S., Kim, J., Allenby, G.M., 2010. A category-level model of asymmetric complements. *Fisher College of Business Working Paper*. Available at: http://dx.doi.org/10.2139/ssrn.1596218.

Mehta, N., 2007. Investigating consumers' purchase incidence and brand choice decisions across multiple produce categories: A theoretical and empirical analysis. *Marketing Science* 26(2), 196-217.

Norwood, B., Ferrier, P., Lusk, J., 2001. Model selection criteria using likelihood functions and out-of-sample performance. Proceedings of the NCR-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO. [http://www.farmdoc.uiuc.edu/nccc134].

Pinjari, A.R., Bhat, C.R., 2011. Computationally efficient forecasting procedures for Kuhn-Tucker consumer demand model systems: Application to residential energy consumption analysis. Technical paper, Department of Civil and Environmental Engineering, University of South Florida.

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

Ryan, D.L., Wales, T.J., 1998. A simple method for imposing local curvature in some flexible consumer-demand systems. *Journal of Business and Economic Statistics* 16, 331-338.

Sanchez, T.W., Makarewicz, C., Hasa, P.M., Dawkins, C.J., 2006. Transportation costs, inequities, and trade-offs. Presented at the 85th Annual Meeting of the Transportation Research Board, Washington, D.C., January.

Sauer, J., Frohberg, K., Hockmann, H., 2006. Stochastic efficiency measurement: The curse of theoretical consistency. *Journal of Applied Economics* 9(1), 139-165.

Song, I., Chintagunta, P.K., 2007. A discrete-continuous model for multicategory purchase behavior of households. *Journal of Marketing Research* 44(4), 595-612.

Thakuriah, P., Liao, Y., 2005. An analysis of variations in vehicle-ownership expenditures. *Transportation Research Record* 1926, 1-9.

Thakuriah, P., Liao, Y., 2006. Transportation expenditures and ability to pay: Evidence from consumer expenditure survey. *Transportation Research Record* 1985, 257-265.

U.S. Bureau of Labor Statistics (BLS). 2003. 2002 Consumer expenditure interview survey public use microdata documentation. U.S. Department of Labor Bureau of Labor Statistics, Washington, D.C., http://www.bls.gov/cex/csxmicrodoc.htm#2002.

U.S. Bureau of Labor Statistics (BLS). 2012. Consumer Expenditures in 2010: Lingering Effects of the Great Recession, Report 1037. Available at: http://www.bls.gov/cex/csxann10.pdf

Vásquez-Lavín, F., Hanemann, M., 2008. Functional forms in discrete/continuous choice models with general corner solution. Department of Agricultural & Resource Economics, University of California Berkeley, CUDARE Working Paper 1078.

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

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

Wang, D., Li, J., 2011. A two-level multiple discrete-continuous model of time allocation to virtual and physical activities. *Transportmetrica* 7(6), 395-416.

LIST OF FIGURES

LIST OF TABLES

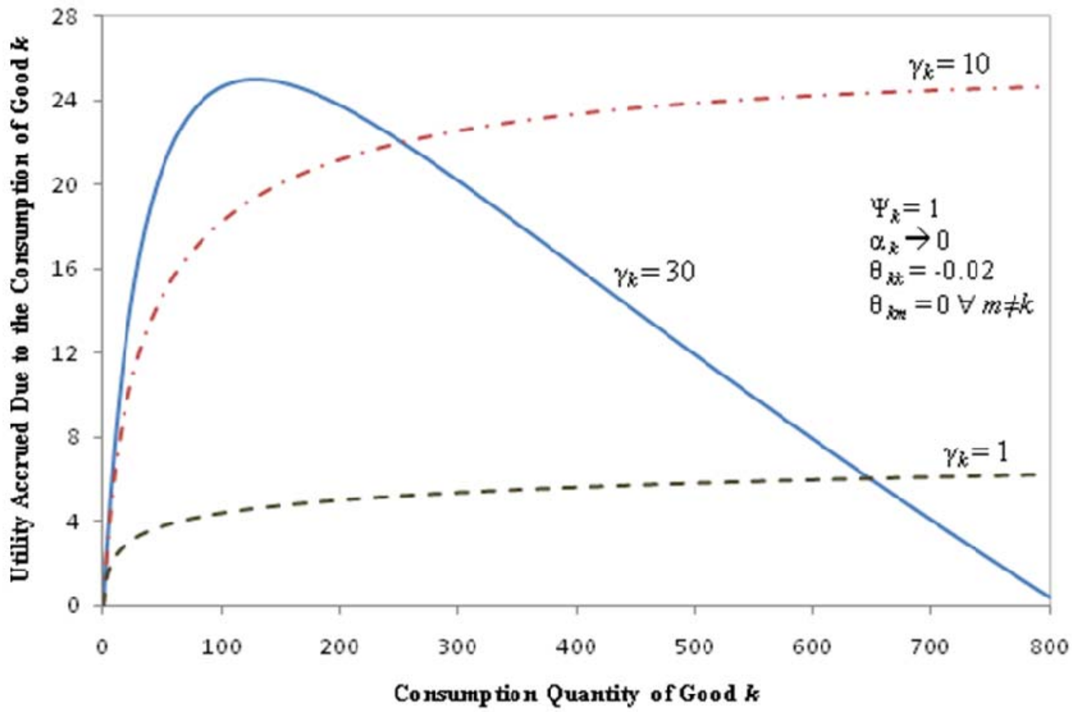

Figure 1. Effect of γ_k, due to Negative θ_{kk}, on Good *k*'s Subutility Function Profile

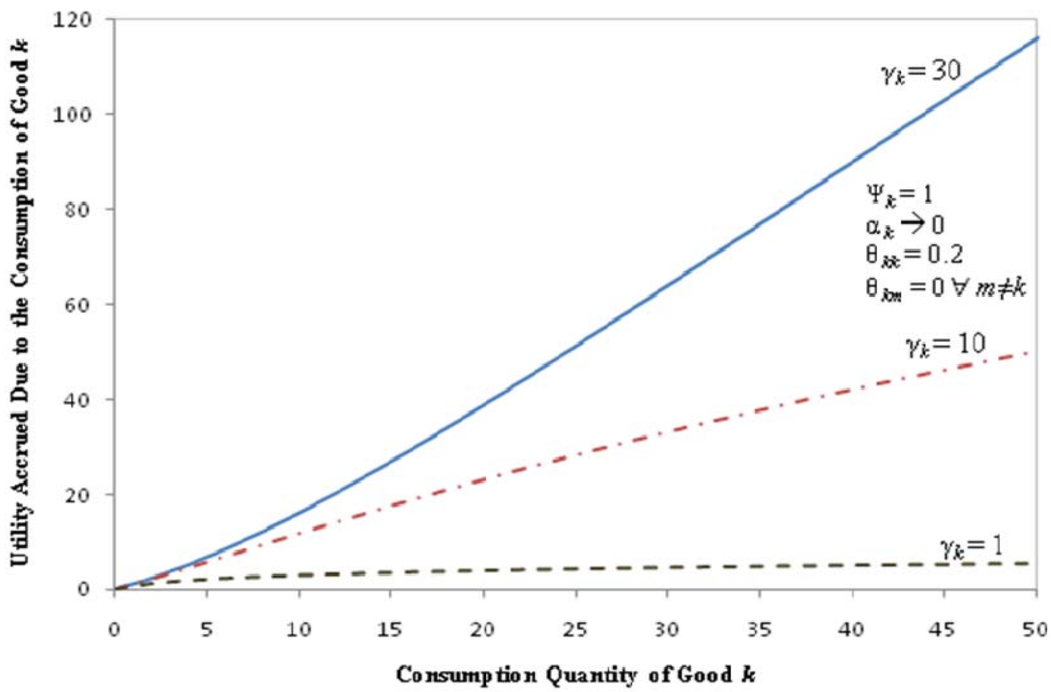

Figure 2. Effect of γ_k, due to Positive θ_{kk}, on Good *k*'s Subutility Function Profile

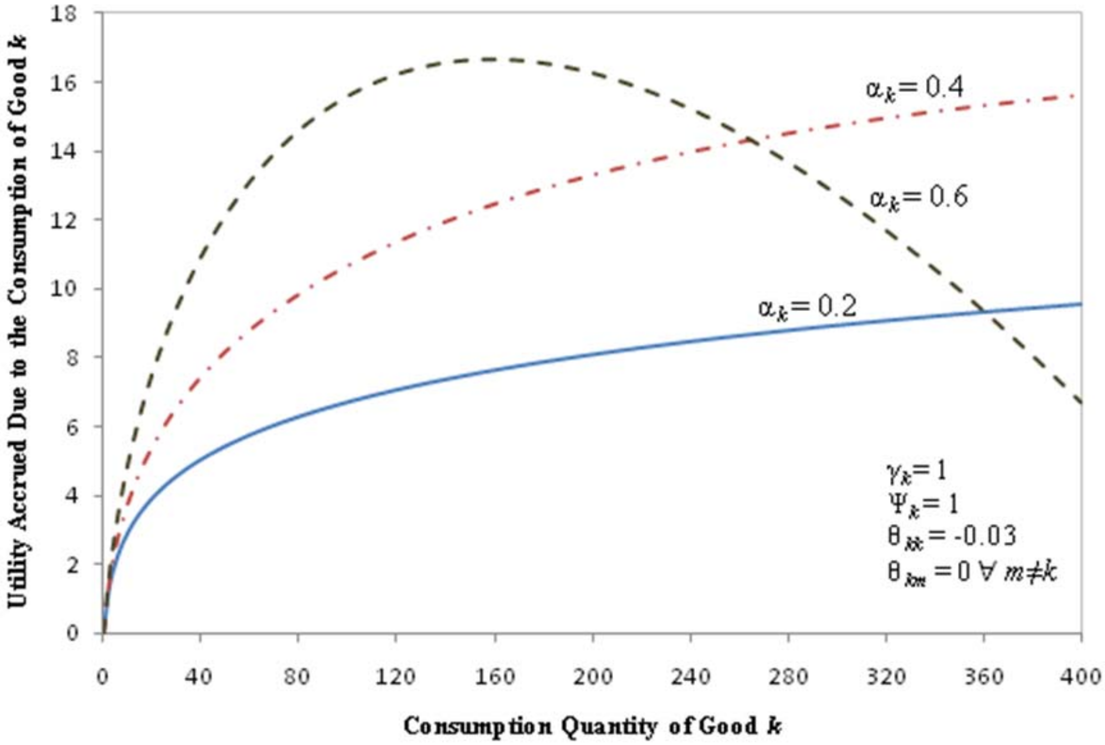

Figure 3. Effect of α_k, due to Negative θ_{kk}, on Good *k*'s Subutility Function Profile

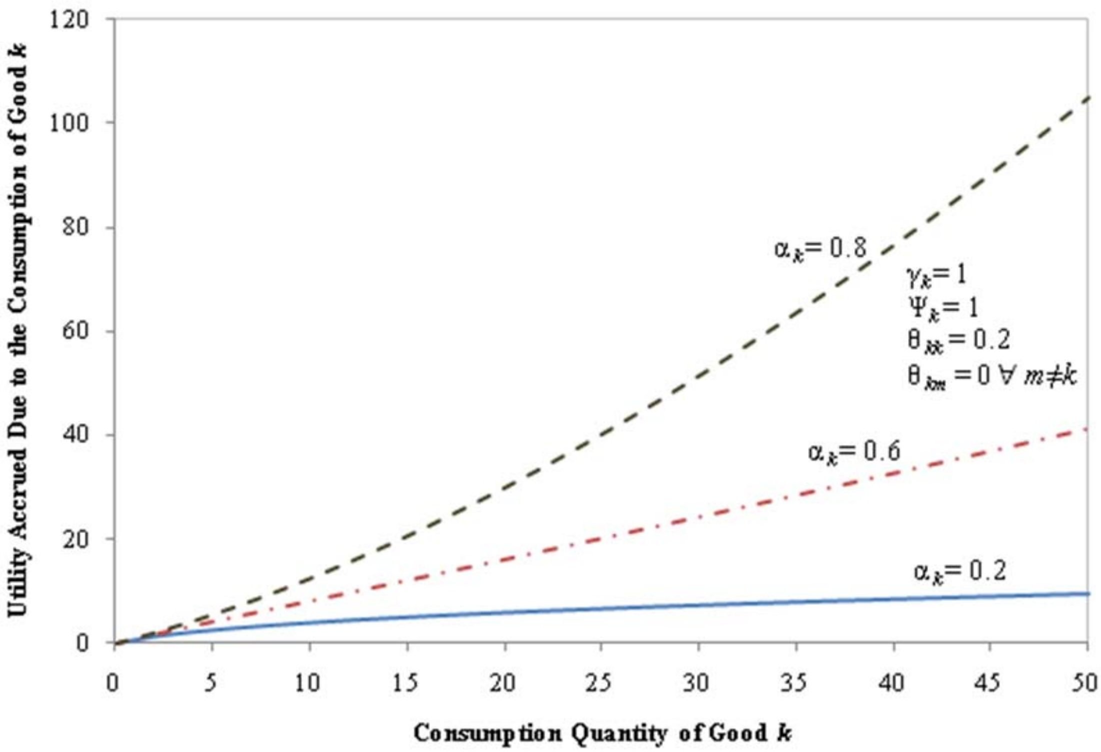

Figure 4. Effect of α_k, due to Positive θ_{kk}, on Good *k*'s Subutility Function Profile

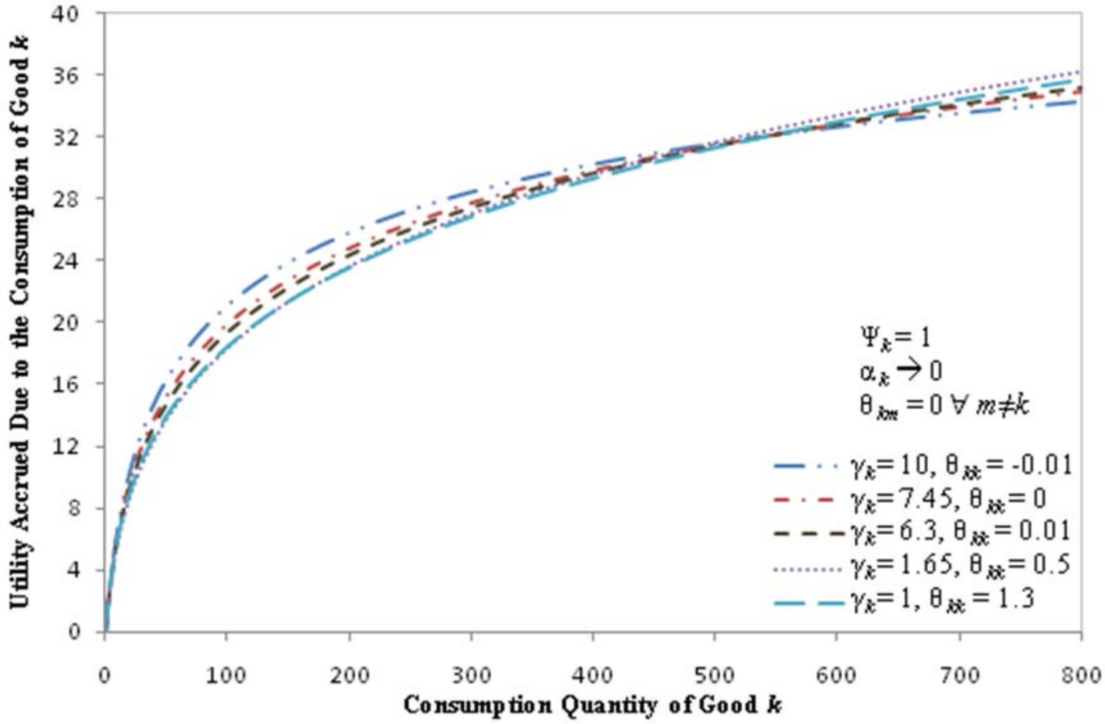

Figure 5. Alternative Subutility Profiles (for Good *k*) with Different θ_{kk} and γ_k Values

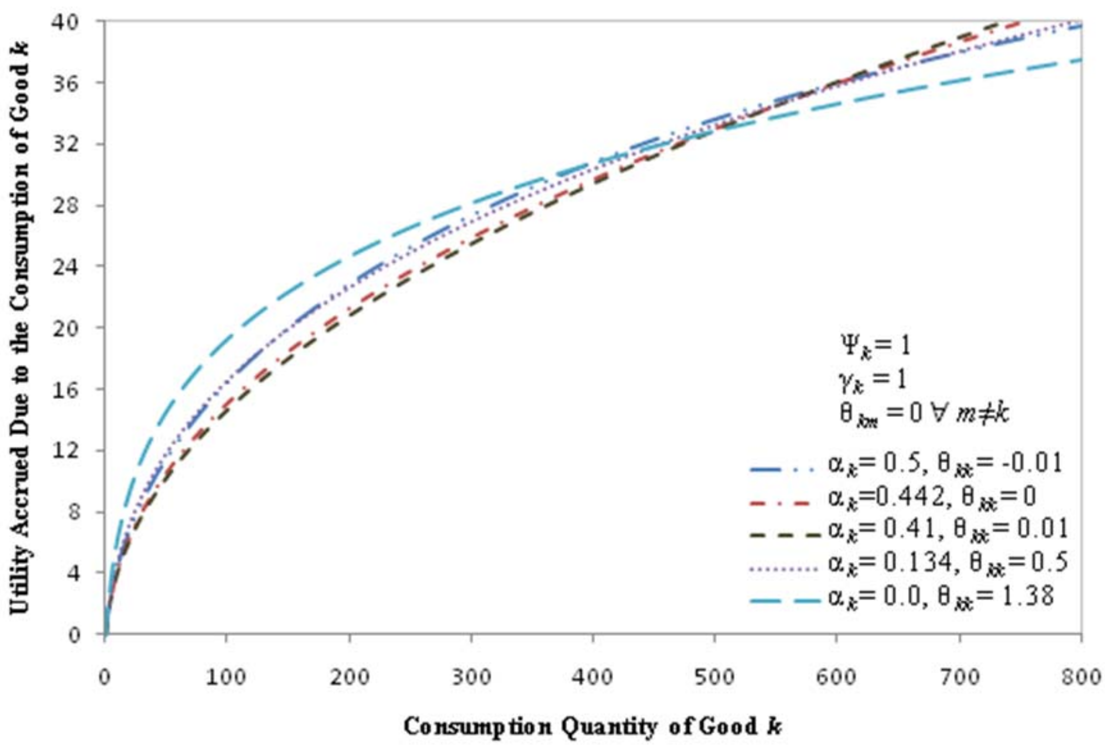

Figure 6. Alternative Subutility Profiles (for Good *k*) with Different θ_{kk} and α_k Values

Table 1. Model Estimation Results

Variables	MDCEV and DU-RM Models				MDCP and RU-DM Models				RU-RM NAS	
	MDCEV		DU-RM NAS		MDCP		RU-DM NAS			
	Parameter	t-stat	Parameter	t-stat	Parameter	t-stat	Parameter	t-stat	Parameter	t-stat
Baseline Utility Parameters (β)										
Baseline Constants										
Veh. purchase	-7.126	-70.59	-8.059	-19.59	-5.865	-117.32	-5.279	-88.92	-5.926	-106.70
Gasoline/oil	-2.523	-37.62	-2.955	-38.11	-3.280	-79.61	-2.366	-56.75	-3.453	-100.43
Veh. insurance	-3.975	-72.08	-4.565	-28.01	-4.116	-106.60	-3.829	-84.00	-4.329	-120.54
Veh. maintenance	-3.446	-60.82	-4.247	-30.02	-3.893	-90.77	-3.486	-78.87	-4.169	-135.08
Air travel	-6.144	-72.87	-5.487	-50.12	-5.334	-125.56	-4.646	-101.04	-5.931	-76.82
Public transp.	-5.819	-42.16	-5.596	-52.95	-5.171	-78.27	-4.489	-52.26	-5.893	-38.35
Number of workers in household										
Veh. purchase	0.182	4.41	0.194	3.59	0.085	3.70	0.060	1.94	0.079	3.62
Gasoline	0.209	7.74	0.264	5.78	0.175	8.40	0.184	6.64	0.165	10.64
Veh. Insurance	0.081	2.89	0.111	2.52	0.058	3.40	-0.003	-0.14	0.039	2.30
Veh. Maintenance	0.192	7.36	0.288	6.02	0.139	8.74	0.116	5.32	0.098	7.71
Annual HH income 30-70K										
Veh. purchase	0.808	7.97	1.368	4.24	0.446	9.69	0.580	9.57	0.513	10.37
Gasoline	-0.284	-5.60	-0.337	-3.32	-0.198	-4.27	-0.346	-6.23	-0.219	-7.51
Air travel	0.756	8.80	0.414	7.26	0.400	9.10	0.511	9.97	0.330	4.43
Annual HH income >70K										
Veh. purchase	0.805	6.34	1.395	4.07	0.430	6.28	0.525	6.10	0.509	7.88
Gasoline	-0.793	-10.89	-0.964	-5.51	-0.656	-8.26	-1.006	-11.18	-0.636	-13.91
Veh. insurance	-0.337	-5.26	-0.379	-2.94	-0.308	-5.18	-0.356	-4.66	-0.251	-5.34
Air travel	1.189	11.31	0.695	7.16	0.587	8.13	0.670	8.70	0.290	2.80
Number of vehicles in household										
Veh. purchase	0.304	11.75	0.340	10.59	0.171	10.77	0.126	6.82	0.149	11.68
Gasoline	0.305	15.70	0.350	12.65	0.263	15.83	0.247	14.26	0.177	20.44
Veh. insurance	0.275	14.04	0.317	10.50	0.220	15.14	0.166	9.46	0.151	12.76
Veh. maintenance	0.269	13.62	0.326	11.86	0.198	14.87	0.141	9.25	0.105	11.96
Air travel	0.073	2.56	0.100	7.30	0.056	3.29	0.007	0.38	-0.030	-1.20
Public transp.	-0.122	-3.82	-0.555	-15.84	-0.051	-3.71	-0.131	-8.74	-0.698	-25.25

Table 1. Model Estimation Results (cont.)

Variables	MDCEV and DU-RM Models				MDCP and RU-DM Models				RU-RM NAS	
	MDCEV		DU-RM NAS		MDCP		RU-DM NAS			
	Parameter	t-stat	Parameter	t-stat	Parameter	t-stat	Parameter	t-stat	Parameter	t-stat
Baseline Utility Parameters (β)										
Non-Caucasian HH – Public transp.	0.417	5.29	0.559	7.28	0.340	10.22	0.347	6.67	0.712	9.75
Urban location – Public transp.	0.490	3.96	0.580	6.09	0.261	4.88	0.287	3.78	0.487	3.36
North East Region – Public transp.	0.722	9.04	0.944	11.10	0.510	14.73	0.585	10.79	0.873	11.32
Western Region – Public transp.	0.590	8.28	0.709	8.73	0.292	8.60	0.370	7.49	0.398	5.59
Translation Parameters (γ_k)										
Veh. purchase	20.888	15.31	21.429	10.95	70.739	12.20	66.645	10.86	80.185	9.82
Gasoline	0.196	17.49	0.179	9.57	0.510	17.73	0.348	33.50	0.744	18.24
Veh. insurance	0.613	27.13	0.607	17.58	1.176	26.99	1.791	29.67	1.568	26.30
Veh. maintenance	0.284	21.08	0.270	17.55	0.879	23.94	1.153	28.82	1.809	23.95
Air travel	0.677	19.58	0.500	14.43	1.879	22.90	1.280	20.05	8.314	16.48
Public transp.	0.237	19.64	0.160	17.47	0.918	30.57	0.577	26.02	1.330	18.33
Interaction Parameters (θ_{km})										
Veh. purchase and gasoline	-	-	1.278×10^{-3}	3.23	-	-	1.126×10^{-3}	37.29	-	-
Veh. purchase and veh. insurance	-	-	-	-	-	-	0.406×10^{-3}	22.78	0.300×10^{-4}	4.36
Veh. purchase and veh. maintenance	-	-	0.338×10^{-3}	2.26	-	-	0.467×10^{-3}	19.41	0.131×10^{-3}	9.98
Veh. purchase and air travel	-	-	-	-	-	-	-	-	-	-
Veh. purchase and public transp.	-	-	-	-	-	-	0.212×10^{-3}	5.70	-0.890×10^{-4}	-4.85
Gasoline and veh. insurance	-	-	2.023×10^{-2}	4.53	-	-	1.954×10^{-2}	32.58	2.436×10^{-3}	10.17
Gasoline and veh. maintenance	-	-	5.095×10^{-2}	7.00	-	-	2.151×10^{-2}	37.93	0.909×10^{-3}	4.95
Gasoline and air travel	-	-	-5.023×10^{-3}	-5.81	-	-	-	-	-	-
Gasoline and public transp.	-	-	-	-	-	-	-	-	-	-
Veh. insurance and veh. maintenance	-	-	4.103×10^{-3}	2.90	-	-	8.879×10^{-3}	25.34	0.366×10^{-3}	4.19
Veh. insurance and air travel	-	-	-	-	-	-	-	-	-	-
Veh. insurance and public transp.	-	-	-	-	-	-	-	-	-	-
Veh. maintenance and air travel	-	-	-	-	-	-	-	-	-	-
Veh. maintenance and public transp.	-	-	-	-	-	-	-	-	-	-
Air travel and public transp.	-	-	8.623×10^{-3}	14.45	-	-	1.199×10^{-3}	7.48	9.204×10^{-3}	35.87
μ parameter	-		-		-		-		0.379	58.51
Number of parameters	33		40		33		41		40	
Log-likelihood at convergence	-37,045		-36,645		-35,269		-35,086		-34,168	

Table 2. Out-of-sample log-likelihood function (OSLLF) in the Validation Sample

Sample details	Number of observations	MDCEV and DU-RM Models		MDCP and RU-DM Models		RU-RM NASUF
		MDCEV	DU-RM NASUF	MDCP	RU-DM NASUF	
Full validation sample	500	-5575.23	-5518.59	-5271.59	-5263.30	-5179.57
Number of workers in HH						
0	14	-147.99	-148.82	-139.89	-142.17	-136.35
1	109	-1139.69	-1126.78	-1075.22	-1149.70	-1059.11
2	240	-2667.62	-2623.82	-2527.78	-2521.63	-2433.19
>2	137	-1619.94	-1619.16	-1528.63	-1520.63	-1515.07
Household income ($/year)						
< 30K	10	-100.62	-101.93	-95.85	-95.28	-100.27
30K-70K	168	-1862.08	-1845.04	-1742.09	-1743.00	-1702.33
>70K	322	-3612.53	-3571.61	-3433.48	-3425.01	-3362.76
Number of vehicles						
0	9	-98.68	-98.00	-95.69	-96.03	-100.06
1	81	-854.90	-846.73	-805.70	-808.38	-783.27
2	173	-1763.61	-1746.95	-1671.01	-1689.54	-1690.87
More than 2	237	-2858.05	-2826.90	-2698.78	-2666.61	-2571.36
Race						
Non-Caucasian	47	-527.42	-520.27	-491.55	-483.70	-494.47
Caucasian	453	-5047.80	-4998.31	-4779.76	-4779.63	-4630.92
Residential location						
Urban	469	-5217.53	-5167.21	-4933.27	-4929.99	-4855.93
Rural	31	-357.72	-351.37	-337.88	-333.33	-321.51