Title

Functional Forms in Discrete/Continuous Choice Models With General Corner Solution

Authors
Vasquez Lavin, Felipe
Hanemann, W. Michael

Publication Date
2008-12-30

Functional Forms in Discrete/Continuous Choice Models With General Corner Solution

Felipe Vásquez Lavín
Department of Agricultural and Resource Economics,
University of California, Berkeley
vasquez@are.berkeley.edu

Michael Hanemann
Department of Agricultural and Resource Economics,
University of California, Berkeley
hanemann@are.berkeley.edu

December 30, 2008

Abstract

In this paper we present a new utility model that serves as the basis for modeling discrete/continuous consumer choices with a general corner solution. The new model involves a more flexible representation of preferences than what has been used in the previous literature and, unlike most of this literature, it is not additively separable. This functional form can handle richer substitution patterns such as complementarity as well as substitution among goods. We focus in part on the Quadratic Box-Cox utility function and examine its properties from both theoretical and empirical perspectives. We identify the significance of the various parameters of the utility function, and demonstrate an estimation strategy that can be applied to demand systems involving both a small and large number of commodities.

1 Introduction

Consumer behavior generally involves two types of decisions. On one hand, people decide which goods to purchase or not purchase (the goods actually purchased are typically only a small subset of all goods available in the market); on the other, people decide what quantity (how many units) to purchase of the commodities they have chosen to acquire. In principle, researchers would like to explain these two decisions using a unified utility model.

Researchers explain the first decision using a discrete choice model such as the multinomial logit, the nested logit model (McFadden, 1979), and recently the mixed logit model (Train, 1998, 2003). For the second decision, one can use a set of demand equations with either a continuous or discrete distribution depending on how much one is concerned about the integer nature of the quantities consumed. Some attempts to linked these two decisions include the applications by Bockstael et al. (1987), Hausmann et al. (1995), Parson and Kealy (1995) and Feather et al. (1995). They use a two-step estimation procedure that in the first step estimates a repeated logit model and stemming from this step, proxy variables for *price* and *quality* of the good are built and incorporated as exogenous variables in the estimation of a demand function that explains the total demand. Unfortunately, their procedure fails to integrate both decisions in a unified model of consumer utility maximization.

A unified model of consumer choice should account for other important characteristics of consumer behavior. First, consumer demand is affected by other attributes of the goods besides prices and income. We use the generic term of *quality* to represent these other atributes. Marketing applications have shown that brand selection and the quantity purchased of goods such as computers or soft drinks depend on the objective and subjective attributes of the different brands (Hendell, 1999; Chan, 2002; Dubé, 2004). In the context of outdoor recreation, environmental quality at different sites is an important variable explaining site visitation behavior (Herriges and Kling, 1999; Phaneuf and Smith, 2004).

Quality has been incorporated in demand models through a transformation of either the demand functions or the utility function (Bockstael et al., 1984). However, incorporating quality into the demand functions directly is not always recommended, especially with large demand systems, because the integrability of these systems is complex and there may sometimes not be a closed form solution for the implied indirect utility function. By contrast, the transformation of the utility function provides a clear relationship between preferences and the resulting demand functions with quality attributes.[1]

A second and very important feature of individual consumer choice behavior is the prevalence of corner solutions, wherein consumers are observed not to purchase any quantity of certain commodities. This stands in contrast to an interior solution, where the consumer consumes some positive quantity of every available commodity. It turns out that a corner solution has a relatively simple analytical structure if the consumer purchases a positive quantity of at most one or two commodities (including the numeraire good); Hanemann (1984) referred to this as an extreme corner solution. In a general corner solution, however, that condition does not

[1] There are at least three ways to incorporate quality into utility functions. First, each good in the utility function can be multiplied by its own quality index (scaling). Second, the quality index could be added to each good of the utility function (translating). Finally, the product of the quality index and the quantity can be incorporated into the utility function through the numeraire good (cross repackaging approach). All of these alternatives differ in terms of the implicit relationship between welfare and quality, demand and quality and empirical tractability. A comprehensive analysis of them can be found in Hanemann, (1982; 1984) and Bockstael et al. (1984).

hold: the consumer purchases positive quantities of more than two commodities but not of all commodities.

Analyzing demand functions for general corner solutions turns out to be rather complex. Since the work by Hanemann (1978, 1984), Wales and Woodland (1983) and Bockstael, Hanemann and Strand (1984), the approach used to formulate a likelihood function for maximum likelihood estimation of a demand system with a general corner solution has been based on the Kuhn-Tucker (KT) conditions for the solution to the consumer's utility maximization problem. The KT approach starts with the formulation of a utility function whose arguments include the level of consumption of each commodity over a period of time and also the quality or other attributes of some or all of the commodities, possibly represented in the form of a sub function that serves as overall index of quality for each relevant commodity. The maximization of this utility function subject to the budget and nonnegativity constraints generates the first order conditions (Karush-Kuhn-Tucker conditions) governing whether a positive or zero quantity of each good is consumed and, if the former, how large a quantity. Adding a random term to the utility function makes it possible to generate probability statements for the consumption bundle observed vector which serve as the building blocks for maximum likelihood estimation of the parameters of the utility function.

Any stochastic specification can be used for the error term. However, the choice of this specification determines the tractability of the likelihood function and therefore the number of commodities that can be handled in the estimation. For instance, using a multivariate normal distribution for the error term makes it hard to apply the model to a demand system with more than a small number of commodities because of the difficulty of evaluating high-dimensional multivariate normal probability integrals; the integrals have to be calculated numerically making the estimation process very slow. In contrast, using an extreme value distribution for the error term produces a simple closed form solution for the probabilities in the likelihood function.

An additional complication arises from the fact that one often wants to use the estimated utility model to predict commodity demands under different scenarios (price-attribute combinations) from those observed in the data. Moreover, one may wish to calculate welfare measures – Hicksian compensating or equivalent variations – for changes in prices or attributes. The complication here is that, in both of these cases, one need to predict the new general corner solution that will be chosen with the new combination of prices and attributes. This is complex because, if there are M commodities (including the numeraire) there are 2^{M-1} alternative possible solutions to the consumer's utility maximization (including an interior solution and all possible corner solutions). Evaluating all of these alternatives in a brute-force determination of the utility-maximizing optimum is computationally burdensome for large M.[2]

[2] Lee and Pitt (1986;1987) subsequently extended this approach to Kuhn-Tucker like conditions that apply to price-derivatives of the indirect utility function. But this does not reduce the dimensionality of the problem,

In the last few years there has been renewed interest in the KT model triggered by three major developments. First, it has proved feasible to apply the KT approach to demand functions with a significant number of commodities based on a clever use of the extreme value distribution (which provides a closed form representation of the likelihood function) and faster computers that permit estimation in a reasonable time (Von Haefen et al. 2004; Bhat, 2007).

Second, the development of simulation techniques allows researchers to avoid the problems associated with numerical evaluation when integrals in the likelihood function lack a closed form representation. Any likelihood function lacking a closed form solution can in principle be approximated using simulation. Furthermore, simulation introduces greater flexibility in the random structure of the model and facilitates the calculation of welfare measures. This flexibility can be exploited when researchers have micro level data since the random parameter models can be used to identify individual heterogeneity and to capture several patterns of substitution (see for example, Train 2003; Revelt and Train, 1999). This is also true in the context of Bayesian estimation of the discrete/continuous model such as Kim et al. (2002).

Finally, the use of an additively separable utility function has provided a simple way to calculate welfare measures without requiring an explicit comparison of all possible solutions to the utility maximization problem. Von Haefen and Phaneuf (2004) and Von Haefen et al. (2004) developed a methodology to estimate welfare measures in a KT framework with a large demand system and an additively separable utility function. This approach has dramatically increased the number of alternatives that can be handled with this framework from 3 or 4 (Wales et al., 1978; Lee and Pitt, 1986; Phaneuf et al., 2000) to 12 or 14 (Von Haefen et al. 2004), and even more than 50 (Von Haeffen et al., 2004; Mohn and Hanemann, 2005).

Bhat (2007) shows that the KT model with an additively separable utility function and an extreme value distribution for the error term is the natural extension of the traditional discrete choice model suggested by McFadden almost four decades ago. Therefore, the random parameter KT model is the natural extension of the mixed logit model (or random parameter logit model) suggested among others by Train and McFadden (1998), Train (1998), and Train (2003).

An additively separable utility function is unfortunately a restrictive functional form since it reduces the flexibility of the utility function in terms of substitution patterns. Nevertheless, the difficulty of finding an appropriate likelihood function and of calculating welfare measures with a nonadditively separable utility function has kept researcher from employing more general models.

In this paper we explore the estimation of a generalized KT model with a nonadditively separable utility function. Depending on the value of the parameters, this utility function can take almost any particular functional form used in the literature such as the translog, linear,

both in estimating the likelihood function and in predicting demands or calculating welfare measures for new price-attribute combinations.

and quadratic utility functions. We discuss estimation strategies and welfare calculation for both small and large demand systems. In our application we estimate both additively and nonadditively separable utility functions and compare welfare measures derived from them.

The next section develops the general KT model and presents a discussion of functional forms used in the literature with emphasis on Bhat's suggestion. Section 3 shows the generalization to a nonadditively separable utility function, its stochastic formulation, and the likelihood function. Section 4 presents an application of the model to a large demand system.

2 Kuhn-Tucker framework

In this model consumers have a continuously differentiable, strictly increasing, and strictly quasiconcave utility function (Hanemann 1978, 1984; Wales and Woodland, 1984; Von Haefen et al., 2004) denoted by

$$U = U(\mathbf{x}, \mathbf{q}, z, \boldsymbol{\beta}, \boldsymbol{\varepsilon}),$$

where $\mathbf{x}$ is a M-dimensional vector of consumption levels, $\mathbf{q}$ is a Mxk matrix of attributes for the vector of commodities and z is a Hicksian good. $\boldsymbol{\beta}$ is a vector of parameters and $\boldsymbol{\varepsilon}$ is a vector of unobserved components. Given a vector of prices ($\mathbf{p}$) for the commodities and a level of income (y) for the individual, the utility maximization problem is

$$\max_{\mathbf{x},\ z} U(\mathbf{x}, \mathbf{q}, z, \boldsymbol{\beta}, \boldsymbol{\varepsilon}) \text{ s.t. } \mathbf{p}'\mathbf{x} + z = y,\ \ \mathbf{x} \geq \mathbf{0}. \tag{1}$$

The first order Kuhn-Tucker conditions are

$$\begin{aligned} \frac{\partial U}{\partial x_k} &\leq \frac{\partial U}{\partial z} p_k \qquad k = 1, ..., M \\ x_k \left(\frac{\partial U}{\partial x_k} - \frac{\partial U}{\partial z} p_k \right) &= 0 \qquad\quad k = 1, ..., M \\ x_k &\geq 0. \end{aligned}$$

Assuming an additive error term in these equations, they can be rewritten as

$$\begin{aligned} \varepsilon_k &\leq g_k\left(\mathbf{x}, \mathbf{q}, \mathbf{p}, \boldsymbol{\beta}\right) & (2) \\ x_k\left(\varepsilon_k - g_k\left(\mathbf{x}, \mathbf{q}, \mathbf{p}, \boldsymbol{\beta}\right)\right) &= 0 & (3) \\ x_k &\geq 0 \quad j = 1, ..., M, \end{aligned}$$

where $g_k(.)$ is a function that contains only the deterministic component of the *FOC*. Let $\hat{x}_n$ represent the observed combination of zero and positive consumption of k goods for individual n, that is $\hat{x}_{nk} = (x_1, ...x_k, 0, ..., 0)$, then the probability of observing an individual consuming only the first k goods is

$$f(\hat{x}_{nk}) = \int_{-\infty}^{g_M} \dots \int_{-\infty}^{g_{k+1}} f_\varepsilon(g_1, \dots g_k, \varepsilon_{k+1}, \dots, \varepsilon_M) \times |J_k| \, d\varepsilon_{k+1} \dots d\varepsilon_M, \tag{4}$$

where $f_\varepsilon(.)$ is the joint density function of the error terms and J_k is the Jacobian of the transformation. For the goods that are consumed we know by equation (3) that $\varepsilon_k = g_k(\mathbf{x}, \mathbf{q}, \mathbf{p}, \boldsymbol{\beta})$, however for the rest of the goods that are not consumed we only know that $\varepsilon_k \leq g_k(\mathbf{x}, \mathbf{q}, \mathbf{p}, \boldsymbol{\beta})$ by equation (2).

Given a distribution function for the error term and a functional form for the utility function, it is possible in principle to construct the likelihood function. However, whether or not the likelihood function is computationally tractable, depends on how we represent the interaction among commodities in the utility function. This interaction is modeled through the choice of a functional form or an error structure. The simplest case uses a functional form and an error structure that does not allow any interaction – i.e., substitution – among the alternatives commodities. Using a distribution function that allows dependence among error terms creates a problem for the likelihood function because for many distributions there is not a closed form solution for the integrals in equation (4) and this severely limits the number of goods that can be handled in the model. Similarly, more sophisticated functional forms increase the difficulties in the definition of the likelihood function and the estimation process when there are many commodities. For instance, if we assume the error terms have a joint normal distribution with zero mean and not a diagonal covariance matrix, then the integrals in the expression $f(\hat{x}_n)$ (equation 4) do not have a closed form and have to be calculated numerically. Other error structures, such as a GEV distribution makes the calculation of the density function complex as the number of alternatives increases.

These facts seem to explain why the literature has shown combinations of a few number of goods with a multivariate normal distribution or a large number of goods but with an extreme value distribution. For example, in this last case the probability of observing an individual consuming only the first k goods is given by

$$f(\hat{x}_{nk}) = |J_k| \prod_{k}^{M} \left(e^{(-g_k(\mathbf{x}, \mathbf{q}, \mathbf{p}, \boldsymbol{\beta})/\mu)}/u \right)^{I_{x_k > 0}} e^{\left(-e^{(-g_k(\mathbf{x}, \mathbf{q}, \mathbf{p}, \boldsymbol{\beta})/\mu)} \right)},$$

whose closed form solution enables us to estimate the model with a large number of goods.

Previous applications of the KT model used a normal distribution and a quadratic utility function but with only 4 goods (Woodland et al., 1983). The model has been also estimated using a generalized extreme value distribution, an additively separable utility function and four goods by Phaneuf et al. (1999). More recently, there have been some applications using an extreme value distribution, an additively separable utility function and a larger number of goods (Von Haefen et al., 2004; Von Haefen and Phaneuf, 2004; Mohn and Hanemann, 2005).

Calculation of welfare measures requires solving a constrained optimization problem. This fact has constrained the application of KT models with large demand systems to utility functions that are additively separable. The compensating variation (CV) for a change from $(\mathbf{p}_0, \mathbf{q}_0)$ to $(\mathbf{p}_1, \mathbf{q}_1)$ is calculated implicitly using the indirect utility function as

$$V(\mathbf{p}_1, \mathbf{q}_1, y - CV, \boldsymbol{\beta}, \boldsymbol{\varepsilon}) = V(\mathbf{p}_0, \mathbf{q}_0, y, \boldsymbol{\beta}, \boldsymbol{\varepsilon}). \tag{5}$$

This implicit definition of CV does not have a closed form solution, therefore researchers use a mathematical algorithm to find the value of CV that satisfies the equality condition. For the new vector $(\mathbf{p}_1, \mathbf{q}_1)$ we need to find the optimal consumption pattern that maximizes utility. For a small number of alternatives Phaneuf et al. (2000) compare all the 2^M possible consumptions patterns and find the alternative that provides the higher level of utility. Given this consumption pattern and the new utility level they use a numerical bisection routine to find the value of CV in (5). Both procedures perform well if the dimension of the problem is small, however for large demand system the comparison of all consumption patterns is not possible and the bisection routine could also be very slow in find the value of CV. Von Haefen, Phaneuf and Parson (2004) suggest an algorithm to find the optimal consumption pattern using the properties of the additively separable utility function and the nonnegativity condition for the numeraire good. The level of z is the only information needed to determine the consumption of all other goods. After the optimal quantities have been found, they also use the numerical bisection approach to find the CV. Von Haefen (2004) suggests a more efficient approach to find the CV which uses the expenditure function instead of the indirect utility function. In this case the optimal quantities are obtained with a similar routine as in Von Haefen et al. but finding the optimal quantities that minimize the total expenditure. The welfare measure is calculated as

$$CV = y - e(\mathbf{p}_1, \mathbf{q}_1, U_0, \boldsymbol{\beta}, \boldsymbol{\varepsilon}), \tag{6}$$

where the initial level of utility U_0 is obtained evaluating the utility function at the initial values of the explanatory variables using the parameters $\boldsymbol{\beta}$ given in the estimation process.

2.1 Previous functional forms in the KT model

One of the most commonly used functional forms for the utility function is

$$U(\mathbf{x}, \mathbf{Q}, z, \boldsymbol{\beta}, \boldsymbol{\varepsilon}) = \sum_k^M \Psi_j \ln(\phi_k x_k + \theta) + \frac{1}{\rho} z^{\rho},$$

with

$$\ln \Psi_k = \boldsymbol{\delta}'\mathbf{s} + \mu\varepsilon_k, \quad \ln \phi_k = \boldsymbol{\gamma}'\mathbf{q}_k, \quad \rho = 1 - \exp(\rho^*), \quad \ln \theta = \theta^* \text{ and } \ln \mu = \mu^*.$$

x_k is the level of consumption of good k, $\mathbf{s}$ is a vector of individual characteristics, $\mathbf{q}$ is a vector of attributes of the good and $z = y - \mathbf{p}'\mathbf{x}$. ε_k is an extreme value error term with scale parameter μ and θ is a translating parameter. For this functional form the KT conditions are given by

$$\frac{\partial U}{\partial x_k} \leq \frac{\partial U}{\partial z} p_k$$

$$\frac{\Psi_k}{(\phi_k x_k + \theta)} \phi_k - \left(y - \mathbf{p}'\mathbf{x}\right)^{\rho - 1} p_j = 0$$

$$\varepsilon_k \leq \frac{1}{\mu}\left[-\boldsymbol{\delta}'\mathbf{s} + \ln\left(\frac{p_k}{\phi_k}\right) + \ln\left(\phi_k x_k + \theta\right) + (\rho - 1)\ln\left(y - \mathbf{p}'\mathbf{x}\right)\right] = g_k.$$

This functional form is used by Von Haefen et al. (2004) and Herriges et al. (1999). A slightly different utility function has been suggested by Phaneuf et al. (2000) and Mohn and Hanemann (2005), given by

$$U(\mathbf{x}, \mathbf{Q}, z, \boldsymbol{\beta}, \boldsymbol{\varepsilon}) = \sum_k^M \Psi_k \ln\left(x_k + \theta\right) + \ln z$$

and whose *FOC* are

$$\varepsilon_k \leq \ln\left(p_k\right) + \ln\left(x_k + \theta\right) - \ln\left(y - \mathbf{p}'\mathbf{x}\right) - \boldsymbol{\delta}'\mathbf{s}.$$

Bhat generalizes these functional forms using a Box-Cox transformation of the quantities consumed in the model, the utility function is

$$U = \sum_{k=1}^M \frac{\gamma_k}{\alpha_k} \Psi_k \left\{ \left(\frac{x_k}{\gamma_k} + 1\right)^{\alpha_k} - 1 \right\}$$

where the Box-Cox transformation is applied over $x_k^* = (x_k/\gamma_k + 1)$, i.e., we have that $(x_k^*)^{(\alpha_k)} = \left((x_k^*)^{\alpha_k} - 1\right)/\alpha_k$. α_k, γ_k and Ψ_k are parameters to be estimated. Consistency conditions require $\Psi_k \geq 0$ and $\alpha_k \leq 1$. Unlike previous functional forms there is not an outside good in this formulation.

This model allows corner solutions given the presence of the parameter α_k. If $\alpha_k = 1$ the model reduces to the extreme corner solution discussed by Hanemann (1984) and Chiang and Lee (1992) with utility function equal to $U = \sum_{k=1}^M \Psi_k x_k$. When $\alpha_k = 0$ the utility function is $U = \sum_{k=1}^M \gamma_k \Psi_k \ln\left(\frac{x_k}{\gamma_k} + 1\right)$ which is similar to functional forms described above if we define $\gamma_k \Psi_k = \Psi_k^*$, $\theta = 1$ and $\frac{1}{\gamma_k} = \phi_k$.

Bhat suggests the following interpretation for the parameters; Ψ_k is the baseline marginal utility since $\frac{\partial U}{\partial x_k} = \Psi_k \left(\frac{x_k}{\gamma_k} + 1\right)^{\alpha_k - 1}$, therefore, if $x_k = 0$ then $\frac{\partial U}{\partial x_k} = \Psi_k$. The parameters γ_k and α_k are satiation parameters where the former shifts the position of the point at which the indifference curve hits the positive orthant allowing corner solutions while the latter defines the satiation level of a good.

As in the other cases the stochastic part of the model is included in the definition of $\Psi_k(s_k, \varepsilon_k) = \Psi(s_k)e^{\varepsilon_k} = \exp(\beta' s_k + \varepsilon_k)$. Following Bhat's the maximization problem is

$$\max_{x_k} U = \sum_{k=1}^{M} \frac{\gamma_k}{\alpha_k} \exp(\beta' s_k + \varepsilon_k) \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\}$$

subject to

$$\sum_{k=1}^{M} p_k x_k = \sum_{k=1}^{M} e_k = E$$

whose FOC are given by

$$\frac{\exp(\beta' s_k + \varepsilon_k)}{p_k} \left(\frac{e_k}{p_k \gamma_k} + 1 \right)^{\alpha_k - 1} - \lambda \leq 0$$

this is satisfied with equality when $e_k^* > 0$ and with inequality when $e_k^* = 0$. Individuals are constrained to consume at least one good, therefore the marginal utility of income can be defined as $\lambda = \frac{\exp(\beta' s_1 + \varepsilon_1)}{p_1} \left(\frac{e_1}{p_1 \gamma_1} + 1\right)^{\alpha_1 - 1}$ using good 1 without loss of generality. Replacing it into the FOC we obtain

$$\frac{\exp(\beta' s_k + \varepsilon_k)}{p_k} \left(\frac{e_k}{p_k \gamma_k} + 1 \right)^{\alpha_k - 1} - \frac{\exp(\beta' z_1 + \varepsilon_1)}{p_1} \left(\frac{e_1}{p_1 \gamma_1} + 1 \right)^{\alpha_1 - 1} \leq 0, \tag{7}$$

this can be reduced to

$$V_k + \varepsilon_k \leq V_1 + \varepsilon_1$$

where

$$V_k = \beta' s_k + \varepsilon_k - \ln p_k + (\alpha_k - 1) \ln \left(\frac{e_k}{p_k \gamma_k} + 1 \right)$$

With this set up the individual's contribution to the likelihood function is

$$l_i = |J_k| \int_{\varepsilon_1 = -\infty}^{\infty} \int_{\varepsilon_M = -\infty}^{g_M} \cdots \int_{\varepsilon_{k+1} = -\infty}^{g_{k+1}} * f(\varepsilon_1, g_2, ..., g_k, \varepsilon_{k+1}, ..., \varepsilon_M) \, d\varepsilon_{k+1} ... d\varepsilon_M d\varepsilon_1$$

with $g_i = V_1 - V_i + \varepsilon_1$. The elements of the Jacobian are

$$J_{ih} = \frac{\partial \left[V_1 - V_i + \varepsilon_1 \right]}{\partial e_h^*}; \; i, h = 2, ..., k.$$

Additionally, γ_k and α_k cannot be identified separately so we have to normalize the model with respect to one of these parameters, then V_k is either $V_k = \beta' s_k + \varepsilon_k - \ln p_k + (\alpha_k - 1)\ln\left(\frac{e_k}{p_k} + 1\right)$ or $\beta' s_k + \varepsilon_k - \ln p_k - \ln\left(\frac{e_k}{p_k \gamma_k} + 1\right)$.

Using an extreme value distribution for the error terms, the individual's contribution to the likelihood becomes

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

where λ denotes the standard extreme value density and Λ denotes the cumulative distribution. The Jacobian reduces to

$$|J| = \left(\prod_{i=1}^{K} c_i\right)\left(\sum_{i=1}^{K} \frac{1}{c_i}\right) \quad \text{with } c_i = \frac{1-\alpha_i}{e_i^* + \gamma_i p_i}$$

and after solving the integral in the likelihood function Bhat's solution is

$$l_i = \frac{1}{\mu^{K-1}}\left(\prod_{i=1}^{K} c_i\right)\left(\sum_{i=1}^{K} \frac{1}{c_i}\right) \frac{\prod_{i=1}^{K} e^{v_i/\mu}}{\sum_{k=1}^{M} e^{v_k/\mu}}(K-1)!$$

This elegant formulation collapses to the simple conditional logit model formulation when K=1, i.e., $l_i = \frac{e^{v_i/\mu}}{\sum_{k=1}^{K} e^{v_k/\mu}}$.

3 A non-additively separable utility function

In this section we present a Quadratic Box-Cox functional form, which is the natural extension of the additively separable linear Box-Cox functional form developed by Bhat (2007). We present this non-additively separable utility function, discuss the interpretation of the parameters in the model, and compare our findings with Bhat's results. The utility function is

$$\begin{aligned} U &= \sum_{k=1}^{M} \frac{\gamma_k}{\alpha_k}\Psi_k\left(\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right) \\ &+ \frac{1}{2}\sum_{k=1}^{M}\sum_{m=1}^{M} \beta_{km}\left(\frac{\gamma_k}{\alpha_k}\left(\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k} - 1\right)\right)\left(\frac{\gamma_m}{\alpha_m}\left(\left(\frac{x_m}{\gamma_m}+1\right)^{\alpha_m} - 1\right)\right), \end{aligned}$$

Again, x_k is the quantity of good k and α_k, γ_k and Ψ_k are parameters to be estimated in the model with the same assumptions as before, i.e., $\Psi_k \geq 0$ and $\alpha_k \leq 1$.

This utility function includes as particular cases most of the other utility functions used in the KT approach. For instance, if $\beta_{km} = 0$ for all k, this utility function becomes the linear Box-Cox utility function, with the properties discussed above.

If $\alpha_k \to 0$ the utility becomes the translog function

$$U = \sum_{k=1}^{K} \gamma_k \Psi_k \ln\left(\frac{x_k}{\gamma_k} + 1\right) + \frac{1}{2}\sum_{k=1}^{K}\sum_{m=1}^{M} \beta_{km}\gamma_k\gamma_m \ln\left(\frac{x_k}{\gamma_k} + 1\right)\ln\left(\frac{x_m}{\gamma_m} + 1\right),$$

and if $\alpha_k = 1$ we obtain the quadratic utility function used by Wales and Woodland (1983).

$$U = \sum_{k=1}^{K} \Psi_k x_k + \frac{1}{2}\sum_{k=1}^{K}\sum_{m=1}^{M} \beta_{km} x_k x_m$$

Now we analyze the role of the parameters Ψ_k, γ_k, and α_k in the model using this utility function.

3.1 Interpretation of parameters

3.1.1 Marginal utility: Ψ_k

Taking a derivative with respect to x_k produces the marginal utility of consumption

$$\frac{\partial U}{\partial x_k} = \Psi_k \left(\frac{x_k}{\gamma_k} + 1\right)^{\alpha_k - 1} + \sum_{m=1}^{M} \beta_{km} \left(\frac{x_k}{\gamma_k} + 1\right)^{\alpha_k - 1} \frac{\gamma_m}{\alpha_m}\left(\left(\frac{x_m}{\gamma_m} + 1\right)^{\alpha_m} - 1\right) \tag{8}$$

at $x_k = 0$ we have

$$\left.\frac{\partial U}{\partial x_k}\right|_{x_k=0} = \Psi_k + \sum_{m_{-k}}^{M} \beta_{km} \frac{\gamma_m}{\alpha_m}\left(\left(\frac{x_m}{\gamma_m} + 1\right)^{\alpha_m} - 1\right)$$

then the marginal utility depends on the coefficients α_m, β_{km}, γ_m, and Ψ_k and the level of consumption of the other goods. In this case

$$\frac{\partial^2 U}{\partial x_k \partial x_m} = \beta_{km}\left(\frac{x_k}{\gamma_k} + 1\right)^{\alpha_k - 1}\left(\frac{x_m}{\gamma_m} + 1\right)^{\alpha_m - 1}, \ m \neq k$$

and the effect depends on the sign of β_{km}. Similarly to Bhat's linear Box-Cox additively separable utility function the contribution to the utility is zero when $x_k = 0$, since $\left(\left(\frac{x_k}{\gamma_k} + 1\right)^{\alpha_k} - 1\right) = 0$, then all the interactions including x_k are eliminated. But, unlike his formulation, the marginal utility in this general utility function includes the effect of consumption of other goods on the desirability of x_k, for instance, if the good already consumed is complementary with x_k, ($\beta_{km} > 0$) then it is more likely that the x_k will be consumed. Ψ_k is the baseline utility only if all other goods are not consumed, in other words Ψ_k represents the desirability of x_k before any consumption decision has been made.

3.1.2 Corner solution and satiation parameter: γ_k.

For our analysis we consider only two goods and assume the following values for the parameters $\Psi_1 = \Psi_2 = 1$, $\alpha_1 = \alpha_2 = 0.5$, $\gamma_2 = 1$, $\beta_{11} = \beta_{22} = -0.01$, $\beta_{12} = -0.001$. The indifference curves for three different values of gamma are given in figure 1.

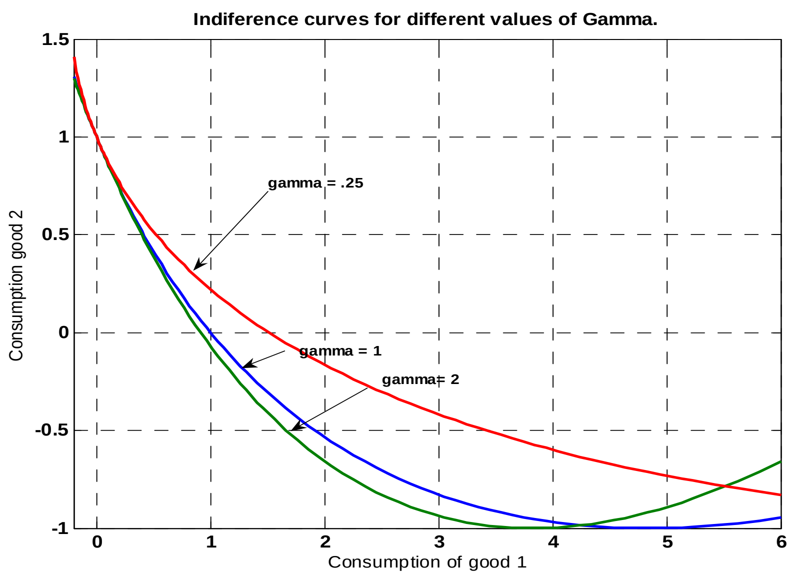

Figure 1

As in the linear Box-Cox additively separable functional form, γ_k shifts the position of the intersection between the indifference curve and the x-axis, allowing for corner solutions since the budget line could be tangent to the indifference curve at the point where this curve intersects the axis. The indifference curves become steeper as the value of γ_k increases which can be interpreted as a stronger preference for good 1. The contribution to the utility function of consumption of good 1 when $\alpha_k \to 0$ is given by

$$\begin{aligned} U &= \gamma_1 \ln\left(\frac{x_1}{\gamma_1}+1\right) + \frac{\beta_{11}\gamma_1^2}{2}\left(\ln\left(\frac{x_1}{\gamma_1}+1\right)\right)^2 \\ &\quad + \beta_{12}\gamma_1\gamma_2 \ln\left(\frac{x_1}{\gamma_1}+1\right)\ln\left(\frac{x_2}{\gamma_2}+1\right) + \frac{\beta_{22}\gamma_2^2}{2}\left(\ln\left(\frac{x_2}{\gamma_2}+1\right)\right)^2 \end{aligned}$$

since

$$\lim_{\alpha_k \to 0} \frac{\gamma_k}{\alpha_k} \Psi_k \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\} = \gamma_k \Psi_k \ln \left(\frac{x_k}{\gamma_k} + 1 \right)$$

Figure 2 plots the utility contribution for values of γ_k equal to 1, 5, and 100.

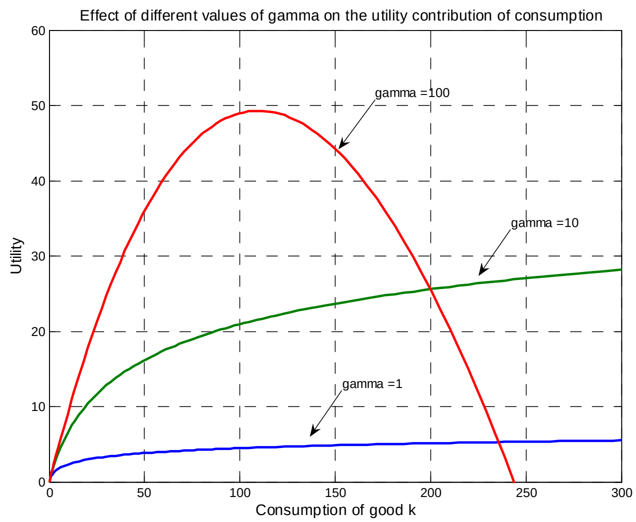

Figure 2

According to figure 2, the parameter γ_k also defines the satiation points in the model since it determines the marginal utility of consumption and the level where we have a negative contribution to utility.

3.1.3 Satiation parameter: α_k

Again, interpretation of this parameter is similar to the linear Box-Cox results. α_k determines how fast the marginal utility decreases with increasing consumption of good k. for Example if $\alpha_k = 1 \; \forall k$ there is no satiation, the utility function is

$$U = \sum_{k=1}^{M} \Psi_k x_k + \frac{1}{2} \sum_{k=1}^{M} \sum_{m=1}^{M} \beta_{km} \left(\frac{x_k}{\gamma_k} \right) \left(\frac{x_m}{\gamma_m} \right)$$

With Bhat's utility function only the first term in the equation exists. Therefore, unlike the simple case we do not have perfect substitute goods since there are some substitution given by the second term of the equation. Figure 3 shows that the higher the alpha the steeper the slope of the function representing the utility contribution of consumption. In other words, a higher alpha implies slower satiation in consumption.

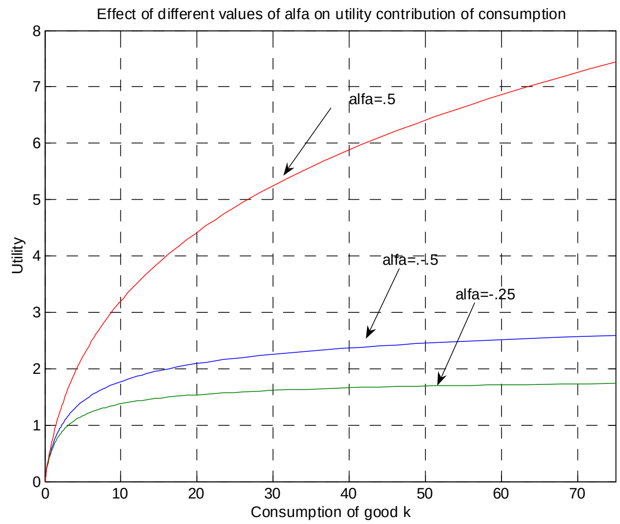

Figure 3

3.2 Likelihood function

With the general utility function the calculation of the Jacobian and the construction of the likelihood function become more complicated. The maximization problem is

$$\begin{aligned}\max U &= \sum_{k=1}^{M}\frac{\gamma_k}{\alpha_k}\Psi_k\left(\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k}-1\right)\\ &+\frac{1}{2}\sum_{k=1}^{M}\sum_{m=1}^{M}\beta_{km}\left(\frac{\gamma_k}{\alpha_k}\left(\left(\frac{x_k}{\gamma_k}+1\right)^{\alpha_k}-1\right)\right)\left(\frac{\gamma_m}{\alpha_m}\left(\left(\frac{x_m}{\gamma_m}+1\right)^{\alpha_m}-1\right)\right)\end{aligned}$$

subject to

$$\sum_{k=1}^{M}p_k x_k=\sum_{k=1}^{M}e_k=E$$

For convenience most of the literature has defined the error term inside Ψ_k either by a multiplicative approach as $\Psi_k = e^{\beta' s_k}e^{\varepsilon_k}$ (Bhat, 2007; Von Haefen et al. 2004; Hanemann, 1984) or an additive approach $\Psi_k = \beta' s_k + \varepsilon_k$ (Wales et al. 1983; Hanemann, 1984) and the assumptions on ε_k are either a extreme value distribution or a normal distribution. As we showed above in the simple case the assumption $\Psi_k = e^{\beta' s_k}e^{\varepsilon_k}$ is convenient since the FOC are reduced to an expression easily simplified by applying the logarithm operator (see equation 7) and using it directly in the likelihood function. However, in our new formulation this multiplicative assumption does not help us simplify the model.

There are two solutions for this problem. First, we could follow the approach by Wales and Woodland and define $\Psi = \beta' s_k + \varepsilon_k$. Second, we could keep the definition of $\Psi_k = e^{\beta' s_k} e^{\varepsilon_k}$ but define an outside good which will simplify the calculation of the FOC. We show details of each alternative below.

3.2.1 Linear stochastic structure

In the first case the FOC can be expressed as

$$\begin{aligned} 0 \quad \leq \quad & \frac{\beta' s_k + \varepsilon_k}{p_k} \left(\frac{e_k}{p_k \gamma_k} + 1 \right)^{\alpha_k - 1} + \frac{1}{2} \sum_{m=1}^{M} \frac{\beta_{km}}{p_k} \left(\frac{e_k}{p_k \gamma_k} + 1 \right)^{\alpha_k - 1} \frac{\gamma_m}{\alpha_m} \left(\left(\frac{e_m}{p_m \gamma_m} + 1 \right)^{\alpha_m} - 1 \right) \\ & - \frac{\beta' s_1 + \varepsilon_1}{p_1} \left(\frac{e_1}{p_1 \gamma_1} + 1 \right)^{\alpha_1 - 1} - \frac{1}{2} \sum_{m=1}^{M} \frac{\beta_{1m}}{p_1} \left(\frac{e_1}{p_1 \gamma_1} + 1 \right)^{\alpha_1 - 1} \frac{\gamma_m}{\alpha_m} \left(\left(\frac{e_m}{p_m \gamma_m} + 1 \right)^{\alpha_m} - 1 \right) \end{aligned}$$

with equality if $e_k > 0$ and inequality if $e_k = 0$. The last component in this expression is the marginal utility of income λ.

Lets define $a_k = \left(\frac{e_k}{p_k \gamma_k} + 1 \right)^{\alpha_k - 1}$ and $a_1 = \left(\frac{e_1}{p_1 \gamma_1} + 1 \right)^{\alpha_1 - 1}$ then the equation is

$$\begin{aligned} 0 \quad \leq \quad & \frac{\varepsilon_k}{p_k} a_k - \frac{\varepsilon_1}{p_1} a_1 + \frac{\beta' z_k}{p_k} a_k + \frac{1}{2 p_k} a_k \sum_{m=1}^{M} \frac{\beta_{km} \gamma_m}{\alpha_m} \left(\left(\frac{e_m}{p_m \gamma_m} + 1 \right)^{\alpha_m} - 1 \right) - \frac{\beta' z_1}{p_1} a_1 \\ & - \frac{1}{2 p_1} a_1 \sum_{m=1}^{M} \frac{\beta_{1m} \gamma_m}{\alpha_m} \left(\left(\frac{e_m}{p_m \gamma_m} + 1 \right)^{\alpha_m} - 1 \right) \\ 0 \quad \leq \quad & \frac{\varepsilon_k}{p_k} a_k - \frac{\varepsilon_1}{p_1} a_1 + V_k - V_1 = \varepsilon_k - \frac{p_k}{p_1} \frac{a_1}{a_k} \varepsilon_1 + \frac{p_k}{a_k} V_k - \frac{p_k}{a_k} V_1 \end{aligned}$$

$$\begin{aligned} \varepsilon_k \quad < \quad & \frac{p_k}{a_k} V_1 - \frac{p_k}{a_k} V_k + \frac{p_k}{p_1} \frac{a_1}{a_k} \varepsilon_1 \\ \varepsilon_k \quad < \quad & a_k^* V_1 - a_k^* V_k + a_k^* a_1^* \varepsilon_1 = V_1^* - V_k^* + \varepsilon_1^* \end{aligned}$$

this area is given by

$$\int_{-\infty}^{V_1^* - V_k^* + \varepsilon_1^*} e^{-\varepsilon_k} e^{-e^{-\varepsilon_k}} d\varepsilon_k = e^{-e^{-(V_1^* - V_k^* + \varepsilon_1^*)}}$$

The likelihood function follows a similar patterns as before,

$$l_i = |J_k| \int_{\varepsilon_1 = -\infty}^{\varepsilon_1 = \infty} \prod_{i=2}^{K} \frac{1}{\mu} \lambda \left[\frac{V_1^* - V_i^* + \varepsilon_1^*}{\mu} \right] \times \prod_{s=K+1}^{M} \Lambda \left[\frac{V_1^* - V_s^* + \varepsilon_1^*}{\mu} \right] \frac{1}{\mu} \lambda \left[\frac{\varepsilon_1}{\mu} \right] d\varepsilon_1$$

Unfortunately there is no simple solution for the Jacobian that can be generalized to any consumption pattern. Furthermore, the integral in the likelihood function does not have a closed form solution. Nevertheless, the integral can be calculated numerically or by simulation and, since there is only a single integral in the likelihood function, this is not burdensome. Because it is possible to include random parameters in the model the additional effort to calculate this integral is not significant. For example, if we include a random parameter structure such that $\beta_n = b + \sigma\eta$, where b is the mean effect and σ a deviation with respect to this mean, then the likelihood function is

$$L(\beta) = \int_{\beta} \int_{\varepsilon_1=-\infty}^{\varepsilon_1=\infty} \prod_{i=2}^{K} \frac{1}{\mu} e^{-\left(V_1^* - V_i^* + \varepsilon_1^*\right)} e^{-e^{-\frac{1}{\mu}\left(V_1^* - V_i^* + \varepsilon_1^*\right)}} \prod_{s=K+1}^{M} e^{-e^{-\frac{1}{\mu}\left(V_1^* - V_s^* + \varepsilon_1^*\right)}} \frac{1}{\mu} e^{-\varepsilon_1} e^{-e^{-\varepsilon_1}} d\varepsilon_1 f(\beta) d\beta$$

which can be calculated using simulation.

3.2.2 Outside good

The alternative way to overcome the difficulty in the definition of the likelihood function is to assume the existence of an outside good which does not have an error term. In that case the utility function will be

$$U^* = U + \frac{1}{\rho} z^{\rho}$$

Where U is the Quadratic Box-Cox utility function. With this assumption the FOC are simpler since we can take advantage of the FOC for z to define the marginal utility of income, i.e.,

$$\frac{\partial U^*}{\partial z} = z^{\rho-1} - \lambda = 0 \Rightarrow z^{\rho-1} = \lambda$$

and we can replace this value in the FOC for other goods.

$$\begin{aligned} 0 \quad \leq \quad & \frac{\exp(\beta' s_k + \varepsilon_k)}{p_k} \left(\frac{e_k}{p_k \gamma_k} + 1\right)^{\alpha_k - 1} \\ & + \frac{1}{2} \sum_{m=1}^{M} \beta_{km} \left(\left(\frac{e_k}{p_k \gamma_k} + 1\right)^{\alpha_k - 1}\right) \left(\frac{\gamma_m}{\alpha_m} \left\{\left(\frac{e_m}{p_m \gamma_m} + 1\right)^{\alpha_m} - 1\right\}\right) - z^{\rho-1} \end{aligned}$$

and

$$\varepsilon_k \leq \beta' s_k + \ln\left(-\frac{p_k}{2}\left(\frac{e_k}{p_k \gamma_k} + 1\right)^{\alpha_k - 1} \sum_{m=1}^{M} \beta_{km} \left(\frac{\gamma_m}{\alpha_m} \left\{\left(\frac{e_m}{p_m \gamma_m} + 1\right)^{\alpha_m} - 1\right\}\right) + p_k z^{\rho-1}\right),$$

which has the same likelihood function as the simple case except we do not have any integrals in its definition because the outside good does not have an error term. In the simple case of the linear Box-Cox utility function, Bhat gives two arguments against the outside good formulation.

First, it is arbitrary to assume one good does not have an error term while all the others are stochastic variables. Second, in his simple linear formulation the likelihood function does not reduce to the simple discrete choice model when $J = 1$. From our perspective these arguments may be less important than the advantage of having a simple likelihood function which facilitates the estimation of this more complex model.

3.3 Welfare measures

With this new utility function we face an additional problem in calculating welfare measures. In this case it is no longer possible to use Von Haefen et al. (2004) algorithm to obtain the optimal quantities after a change in the vector $(\mathbf{p}, \mathbf{q})$. The optimal quantities depend on all other goods since from the FOC we cannot separate $\mathbf{z}$ from the vector of quantities $\mathbf{x}$. We found that the use of a constrained optimization routine solve this problem. Using this routine we minimize the total expenditure needed to reach the initial level of utility U_0 at the new price and quality conditions.

To understand this problem we repeat the relevant part of Von Haefen et al. algorithm for our estimation and explain where we use the constrained optimization method.

1. In the first step Von Haefen et al. simulate the unobserved heterogeneity of the model that comes from the random parameter assumption and from the traditional error component. The process starts taking random draws from the distribution of $\boldsymbol{\beta}_t$, and accepting or rejecting these simulated parameters using a Metropolis-Hasting algorithm. Then it simulates the error terms conditional on the value of the parameters $\boldsymbol{\beta}_t$ given in the first step. Since we have not used random parameters yet no simulation is needed for $\boldsymbol{\beta}_t$, (this can be easily incorporated in the model later), so we only need a simulation for the error terms $\boldsymbol{\varepsilon}_t$ from the conditional distribution $f(\boldsymbol{\varepsilon}_t \,|\boldsymbol{\beta}_t, x_t)$. Following von Haefen et al., for each element of $\boldsymbol{\varepsilon}_t$ we set $\varepsilon_{tj} = g_{tj}(.)$ if the j-th good is consumed and $\varepsilon_{tj} = -\ln(-\ln(\exp(-\exp(-g_{jt}(.)))\upsilon_{tj}))$ where υ_{tj} is a uniform random draw if the j-th good is not consumed.

2. For each set of the simulated heterogeneity given in (1) we need to compute the minimum level of expenditure $e(\mathbf{p}_1, \mathbf{q}_1, U_0, \boldsymbol{\beta}, \boldsymbol{\varepsilon})$ required to reach the initial utility level U_0. Von Haefen et al. find the optimal quantities that minimize the total expenditure using the following subroutine:

   (a) at iteration i the level of z is set at $z_a^i = \left(z_l^{i-1} + z_u^{i-1}\right)/2$, where $z_l^0 = 0$ and $z_u^0 = y$.

   (b) with this suggested value for z solve the FOC derived from an expenditure minimization given by $\frac{\partial u_j(x_j)}{\partial x_j} \leq \frac{\partial u_z(z)}{\partial z} p_j, \quad x_j \geq 0 \quad \forall_j$ and obtain the new quantities x^i.

(c) given results in (b) find the new value for z equal to $z^i = y - \sum_k p_k x_k$

(d) if $z^i > z^i_a$ set $z^{i-1}_l = z^i_a$ and $z^i_u = z^{i-1}_u$. And if $z^i < z^i_a$ set $z^i_l = z^{i-1}_l$ and $z^i_u = z^i_a$.

(e) iterate until $abs(z^i_l - z^i_u) < c$ for c arbitrarily small.

3. The total expenditure given at the last iteration of this subroutine is replaced in equation (6) to obtain the welfare measure. Finally the average of the simulated welfare measures is used as an estimate of the $E(CV)$.

The second part of this algorithm is only appropriate if the utility function is additively separable. In our case the FOC conditions do not allow us to separate the definitions of z from the other goods. To find the optimal bundle of goods we use a constrained optimization routine that finds the values of x and z that minimize the total expenditure subject to the nonnegativity constraints and the initial level of utility. With the value of x and z we calculate the new level of expenditure needed to reach U_0 and replace it into equation (6) to obtain the corresponding welfare measure.

4 Application

Our application uses a data containing a panel of 1063 Alaskan anglers who took 6815 fishing trips in the summer of 1986. We have information on the number of trips taken by each angler in each of the 22 weeks of the season and on the characteristics of each of these trips including date, duration, destination zone and type of fish targeted. For each individual n, we have information on his travel cost (TC_{nj}) to each of the 29 sites included in the sample. This TC was calculated as the round trip travel cost from the origin zone to the corresponding site i.

We also have information on the quality of fishing (Q_{jt}) at each site on each decision occasion. This is based on detailed sportfishing advisories published each week by the Alaska Department of Fish & Game (ADFG) describing fishing conditions at sites around the state. They define the fishing in qualitative terms using adjectives and descriptors, rather than predicting a specific catch rate per hour of effort. Accordingly, we view this as an ordinal measure of fishing quality. Based on the advisories, our variable is coded as an eight-level indicator of fishing quality, starting from 1 (“no fish are available”) to 8 (“excellent fishing”). The fish advisories are specific to 11 different types of fish. These species are classified into 3 macrospecies (Salmon, Saltwater, Freshwater) and a no target alternative. For each macrospecies there are a number of subspecies. In the macrospecies salmon there are four subspecies; King, Red, Silver and Pink. Freshwater has five subspecies; Rainbow Trout, Dolly Varden, Lake Trout, Grayling, and others. Saltwater has only three subspecies, Halibut, Razor Clams, and others.

For each subspecies there are several sites at which fishing is available. Our definition of a "site" therefore is a combination of macrospecies/subspecies/site. There are 181 potential goods, an individual chooses to visit a subset of these combinations. Although there are 29 destination sites in the sample not all of them are available for each subspecies. The maximum number of sites available for each subspecies is presented in table 1 in the appendix. Two other variables describing the sites include a measurement of harvest output in previous year (*LogHarv85*) and the variable *Developed* which takes the value 1 when there is some degree of development like boats and tourist facilities in site i and 0 otherwise.

There are nine variables describing individuals. These variables are *Skill, Leisure, Avlong, Own, Site focus, boatown, trophy, release, cabin* and *crowding*. *Skill* is an index of the individual's experience in sport fishing which ranges from 1 for a novice to 4 for an expert angler. *Leisure* measures the amount of leisure time available for an angler. *Avlong* is the average length of all fishing trips taken by the individual in Alaska over the summer. *Own* is a dummy variable that takes the value 1 if the individual owns a cabin, a boat or an RV and 0 otherwise. *Site focus* is a dummy variable taking the value 1 if the choice of a site was more important to an individual than the choice of a target species. *Boatown* takes the value 1 if the individual owns a boat and 0 otherwise. *Trophy* takes the value 1 if the individual prefers trophy sport fishing and 0 otherwise. *Release* takes the value 1 if the individual practices catch and release fishing and the value 0 otherwise. *Cabin* is a dummy variable that takes the value 1 when the individual owns a cabin at site i and 0 otherwise. *Crowding* is a measure of subjective crowding conditions at site i and it was computed as the product of an individual's crowding tolerance (positive if the individual likes crowded places and negative otherwise) and a measure of crowding conditions at the site i, that ranges from 0 for not crowded to 2 for very crowded. This is based on information provided by ADFG. Note *crowding* and *cabin* varies for both the individual and the site.

Given the size of the choice set, we estimate the both the additively separable utility function and a translog utility function including an outside good. The translog utility function is

$$U(\mathbf{x}, \mathbf{Q}, z, \boldsymbol{\beta}, \boldsymbol{\varepsilon}) = \sum_{j}^{M} \Psi_j \ln\left(\phi_j x_j + \theta\right) + \frac{1}{2}\sum_{i}^{M}\sum_{j}^{M} \beta_{ij} \ln\left(\phi_i x_i + \theta\right) \ln\left(\phi_j x_j + \theta\right) + \ln\left(z\right),$$

with

$$\ln \Psi_j = \boldsymbol{\delta}'\mathbf{s} + \mu\varepsilon_j, \;\; \ln \phi_j = \boldsymbol{\gamma}'\mathbf{q}_j, \;\; \ln \theta = \theta^* \text{ and } \ln \mu = \mu^*,$$

the vector $\mathbf{s}$ include the variables varying across individuals while $\mathbf{q}$ includes the quality variables varying across sites. The interaction parameters β_{ij} can be represented in the matrix

$$B = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1M} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{M1} & \beta_{M2} & \cdots & \beta_{MM} \end{bmatrix},$$

which is a symmetric and negative semi definite matrix of parameters. The KT conditions are given by

$$\frac{\Psi_j}{\left(\phi_j x_j + \theta\right)}\phi_j + \sum_i^M \frac{\phi_j \beta_{ij}}{\phi_j x_j + \theta} \ln\left(\phi_i x_i + \theta\right) - \frac{p_j}{\left(y - \mathbf{p}'\mathbf{x}\right)} = 0$$

$$\varepsilon_j = \frac{1}{\mu}\left(\ln\left(p_j\left(\phi_j x_j + \theta\right) - \phi_j\left(y - \mathbf{p}'\mathbf{x}\right)\sum_i^M \beta_{ij}\ln\left(\phi_i x_i + \theta\right)\right) - \ln\phi_j - \ln\left(y - \mathbf{p}'\mathbf{x}\right) - \boldsymbol{\delta}'\mathbf{s}\right),$$

whose Jacobian is

$$\frac{\partial \varepsilon_j}{\partial x_j} = \frac{1}{\mu}\left(\frac{p_j\phi_j}{A_j} + \frac{\phi_j p_j \sum_i^M \beta_{ij}\ln\left(\phi_i x_i + \theta\right)}{A_j} - \frac{\beta_{jj}\phi_j^2\left(y - \mathbf{p}'\mathbf{x}\right)}{\left(\phi_j x_j + \theta\right)A_j} + \frac{p_j}{y - \mathbf{p}'\mathbf{x}}\right),$$

$$\frac{\partial \varepsilon_j}{\partial x_i} = \frac{1}{\mu}\left(\frac{\phi_j p_i \sum_i^M \beta_{ij}\ln\left(\phi_i x_i + \theta\right)}{A_j} - \frac{\phi_j\phi_i\left(y - \mathbf{p}'\mathbf{x}\right)\beta_{ji}}{\left(\phi_i x_i + \theta\right)A_j} + \frac{p_i}{y - \mathbf{p}'\mathbf{x}}\right),$$

and $A_j = p_j\left(\phi_j x_j + \theta\right) - \phi_j\left(y - \mathbf{p}'\mathbf{x}\right)\sum_i^M \beta_{ij}\ln\left(\phi_i x_i + \theta\right)$. To assure the negativity and symmetry conditions for the matrix B we make the following adjustment (Wiley et al., 1973)

$$B = -CC'$$

where

$$C = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{12} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{11} + a_{12} + ... + a_{1(M-1)} & a_{22} + a_{23} + ... + a_{2(M-1)} & \cdots & 0 \end{bmatrix}$$

and we define the parameters a_{ij} as a function of observable variables (D_{ij}) common to alternatives i and j, that is,

$$a_{ij} = \eta' D_{ij},$$

In this application we use a dummy variable to indicate the two alternatives belong to the same fish subspecies. This is intended to reduce the number of parameters that we need to estimate when we have a large demand system, but it is not necessary with a small demand system.

4.1 Data and Results

Table 2 present the results of the KT model with the additively separable utility function while table 3 presents results of the translog utility function. All the covariates varying among individuals are included in the set $\mathbf{s}$ and enter into the utility function through $\Psi = \exp\left(\boldsymbol{\delta}'\mathbf{s} + \boldsymbol{\varepsilon}\right)$. Results show that five out of nine parameters are significant in the first equation. The other four coefficients on *own, site focus, release* and *crowding* are not significant. Crowding is at the edge of being significant. These parameters suggest an increase in the index of quality or in the level of harvest makes that site more attractive to anglers. Both coefficients are significant.

In the translog utility function the qualitative results are similar to the previous case, except that now *release* becomes significant and *Harvest* is not significant and has an unexpected sign. The 13 dummy variables (D1 to D13) represent the 12 different subspecies and the no target alternative. In other words, $D_J = 1$ if the two goods belong to the same fish group and 0 otherwise. All these coefficients are significant. Other variables describing simultaneously the two goods under consideration could also be incorporated, for example we could use another dummy variable to indicate the two goods belong to the same macrospecies.

Finally, Table 4 presents welfare measures for several scenarios. For example, closing site 1 (Gulkana River) for all salmon species (King and Red in this case), for all freshwater species (Rainbow Trout and Grayling) and for all subspecies. Additionally we include two scenarios for closing one of the most important fishing sites in Alaska; the Kenai River. The first scenario closes the Kenai River for all species (the four types of salmon in this case) and the second scenario closes the river only for king salmon. The welfare loss of closing the Kenai River for king salmon is significantly greater than closing the Gulkana River for this subspecies. This shows the relevance of the Kenai River in the recreational fishing activities in Alaska due to better fishing qualities of the river.

The translog formulation produces smaller welfare measures than the welfare measures of the traditional utility function which can be explained by the substitution patterns allowed in the translog utility function. Even though the magnitudes of the difference in welfare measures are not large, the results show that making the effort to estimate more flexible functional form could help us enrich the substitution patterns of the model and to correctly asses welfare implications of different policies.

5 Conclusions

In this paper we analyze and estimate a more flexible utility function for the Kuhn-Tucker approach to multiple/discrete continuous choice models. Unlike previous literature our utility function is not additively separable and it can be applied to estimate parameters of a large demand system. We estimated this model considering 180 commodities which, to our knowledge, is the largest amount of goods that have been used in this approach. We think we have overcome the problem of welfare calculation using a sequential quadratic programming method to find the optimal quantities needed to perform welfare analysis. The paper contributes to model the multiple discrete continuous choice model, adding greater flexibility to include richer substitution patterns.

References

[1] Bhat, C. (2007) "The Multiple Discrete-Continuous Extreme Value (MDCEV) Model: Role of Utility Function Parameters, Identification Considerations, and Model Extension" Forthcoming *Transportation Research*.

[2] Bockstael, N., M. Hanemann and I. Strand (1984). "Measuring the Benefits of Water Quality Improvements using Recreation Demand Models" US Environmental Protection Agency Washington D.C. 20460

[3] Bockstael, N., M. Hanemann and C. Kling (1987) "Estimating the Value of Water Quality Improvements in a Recreational Demand Framework" *Water Resources Research 23(5):951-960.*

[4] Chan, T. (2002) "Estimating a Continuous Choice Model with an Application to Demand for Soft Drinks," unpublished (R&R at Rand).

[5] Chiang J. and L. Lee (1992) Discrete Continuous Models of Consumer Demand with Binding Nonnegativity Constraints" *Journal of Econometrics* 54: 79:93.

[6] Dubé, Jean-Pierre (2004). "Multiple Discreteness and Product Differentiation: Strategy and Demand for Carbonated Soft Drinks," *Marketing Science 23(1):66-81.*

[7] Hanemann, M. (1978) A Methodological and Empirical Study of the Recreation Benefits from Water Quality Improvements. Ph. D. Dissertation, Department of Economics, Harvard University.

[8] Hanemann, M. (1982) "Quality and Demand Analysis" In New Directions in Econometric Modeling and Forecasting in US agriculture. Edited by Gordon C. Rausser. New York Elsevier/North Holland. p.p. 55-98.

[9] Hanemann, M. (1984). "Discrete/Continuous Models of Consumer Demand" *Econometrica 52(3):541-562.*

[10] Hausman, J., G. Leonard and D. McFadden (1995) " A utility Consistent, Combined Discrete Choice and Count Data Model Assessing Recreational Use Losses Due to Natural Resource Damage" *Journal of Public Economics 56:1-30.*

[11] Hendel, Igal (1999). "Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns," *Review of Economic Studies, 66(1999): 423-446.*

[12] Herriges, J. and C. Kling (1999). Valuing Recreation and the Environment. Edward Elgar.

[13] Herriges, J, C. Kling and D. Phaneuf (1999). "Corner Solutions Models of Recreation Demand: A Comparison of Competing Frameworks: in Valuing Recreation and the Environment, Herriges J and C. Kling ed Edward Elgar.

[14] Kim, J., G. Allenby and P. Rossi (2002) "Modeling Consumer Demand for Variety" Marketing Science 21(3):229-250.

[15] Lee, L. and M. Pitt (1986). "Microeconometric Demand System with Binding Nonnegativity Constraint: The Dual Approach" *Econometrica 54(5):1237-1242.*

[16] Lee, L. and M. Pitt (1987). "Microeconometric Models of Rationating, Imperfect Markets and Nonegativity Constraints" *Journal of Econometrics* 3:89-110.

[17] McFadden, D. (1978). "Modeling the Choice of Residential Location" In Spatial Interaction Theory and Planning Models ed. by A. Karlquist, L. Lundquist, F. Snickars and J.L. Weibull. Amsterdam: North Holland.

[18] McFadden, D. (1979). "Quantitative Methods for Analyzing Travel Behavior of Individuals: Some Recent Developments." *Behavioral Travel Modeling.* David Hensher and P. Stopher, eds. pp. 279-318. London: Croom Helm,1979.

[19] Mohn, Craig and M. Hanemann (2005) "Caught in a Corner Solution: Using the Kuhn-Tucker Conditions to Value Montana Sportfishing" Working paper, University of California.

[20] Parson, G., P. and M. Kealy (1995). "A Demand Theory for Number of Trips in a Random Utility Model of Recreation" *Journal of Environmental Economics and Management 29:357–367.*

[21] Parson, G., P. Jakus and T. Tomasi (1999). "A Comparison of Welfare Estimates from Four Models for Linking Seasonal Recreational Trips to Multinomial Logit Models of Site Choice" *Journal of Environmental Economics and Management 38, 143-157.*

[22] Phaneuf, D. and K. Smith (2004) "Recreation Demand Models" Prepared for Handbook of Environmental Economics. K. Mäler and J Vincent, Editors.

[23] Phaneuf, D., C. Kling and J.Herriges (1999). "Estimation and Welfare Calculations in a Generalized Corner Solution Model with an Application to Recreation Demand" *Review of Economics and Statistics 82(1):83-92.*

[24] Revel, D. and K. Train (1999). "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level" *The Review of Economics and Statistics* 80(4):647-657.

[25] Train, K. (1998). "Recreation Demand Models with Taste Differences Over People" *Land Economics* 74(2):230-39.

[26] Train K. (2003). *Discrete Choice Model with Simulation* Cambridge University Press.

[27] Von Haefen, R. (2004) "Empirical Strategies for Incorporating Week Complementarity into Continuous Demand System Models" Working paper.

[28] Von Haefen, R. and D. Phaneuf (2004) "Kuhn Tucker Demand System Approaches to Non-Market Valuation" prepared for Applications of Simulation Methods in Environmental and Resource Economics, Anna Alberini and Riccardo Scarpa, editors.

[29] Von Haefen, R., D. Phaneuf and G. Parson (2004). "Estimation and Welfare Analysis with Large Demand System" *Journal of Business & Economic Statistics 22(2):194-205.*

[30] Wales T. and A. Woodland (1983). "Estimation of Consumer Demand Systems With Binding Nonnegative Constraints" *Journal of Econometrics 21:263-285.*

[31] Wiley, D., W. Schmidt and W. Bramble (1973) "Studies of a Class Covariance Structure Models" *Journal of the American Statistical Association 68:317-323.*

Table 1: Maximum number of sites available for each subspecies

Species	Number of sites	Species	Number of sites
King	21	Lake Trout	7
Red	17	Arctic Grayling	11
Silver	24	Other Freshwater	9
Pink	11	Halibut	7
Rainbow Trout	19	Razor Clam	2
Dolly Varden	17	Other Saltwater	7
		No target	29
total			181

Table 2. ML estimates of KT model

variable	estimates	s.e.	Est./s.e.
Scale μ	0.335	0.016	21.596
Constante δ_0	-3.197	0.583	-5.487
Skill δ_1	2.835	0.432	6.568
Leisure δ_2	-0.052	0.258	-0.203
Avlong δ_3	-0.775	0.198	-3.921
own δ_4	-0.024	0.072	-0.330
site focus δ_5	0.030	0.073	0.417
boatown δ_6	1.758	0.108	16.266
trophy δ_7	0.646	0.133	4.847
release δ_8	-0.082	0.120	-0.681
cabin δ_9	2.338	0.153	15.302
Crowding δ_{10}	6.065	3.294	1.841
theta θ	2.452	0.117	20.950
quality γ_1	23.675	1.204	19.657
harvest γ_2	17.326	1.230	14.090
Devp γ_3	0.037	0.044	0.849
rho ρ	-3.005	1.089	-2.760
Mean log-likelihood	-21.3683		

Table 3. ML estimates of Translog KT model

variable	estimates	s.e.	Est./s.e.
Scale μ	-0.5692	0.0125	-45.676
Constante δ_0	-8.6695	0.0873	-99.341
Skill δ_1	1.0125	0.2427	4.173
Leisure δ_2	0.0074	0.0672	0.111
Avlong δ_3	-0.6443	0.126	-5.115
own δ_4	-0.034	0.0407	-0.835
site focus δ_5	-0.0126	0.0447	-0.281
boatown δ_6	0.6715	0.0766	8.762
trophy δ_7	0.1627	0.0687	2.369
release δ_8	0.1658	0.0557	2.975
cabin δ_9	0.9502	0.0642	14.795
Crowding δ_{10}	1.8153	1.5288	1.187
theta θ	1.2025	0.0789	15.237
quality γ_1	29.4915	1.8626	15.834
harvest γ_2	-1.6007	0.8734	-1.833
Devp γ_3	-0.0272	0.0518	-0.524
D1	0.0013	0.0001	19.317
D2	0.002	0.0001	16.833
D3	0.0014	0.0001	16.736
D4	0.0045	0.0003	15.631
D5	0.0021	0.0001	18.541
D6	0.0035	0.0003	13.761
D7	0.0057	0.0004	14.867
D8	0.0031	0.0002	16.324
D9	0.0044	0.0003	13.267
D10	0.0028	0.0002	12.020
D11	0.0142	0.0014	10.471
D12	0.0069	0.0006	11.585
D13	0.0011	0.0001	17.421
Mean log-likelihood	-20.1912		

Table 4. Mean WTP KT model

	Separable utility function			Translog		
	mean	Total	Per choice occasion	mean	Total	Per choice occasion
Gulkana river all Species	-8.37	-8898.36	-1.11	-7.51	-7981.35	-0.996
Gulkana river all Salmon	-1.85	-1968.59	-0.25	-1.47	-1559.49	-0.19
Gulkana river King salmon	-1.14	-1210.81	-0.15	-0.99	-1053.01	-0.13
Gulkana river Red salmon	-0.71	-757.70	-0.09	-0.47	-504.19	-0.06
Gulkana river All FW	-5.60	-5950.95	-0.74	-7.13	-7576.15	-0.95
Gulkana River Rainbow trout	-3.68	-3914.65	-0.49	-4.94	-5247.84	-0.65
Gulkana River Grayling	-1.92	-2036.22	-0.25	-1.51	-1600.28	-0.20
Kenai River all species	-4.14	-4400.67	-0.55	-3.04	-3233.42	-0.40
Kenai River King	-1.64	-1740.26	-0.22	-1.48	-1575.98	-0.20