

# **The Integer Programming Approach to a Generalized Corner-Solution Model:**

## **An Application to Recreation Demand**

Koichi Kuriyama

School of Political Science and Economics,

Waseda University,

1-6-1 Nishi-Waseda, Shinjuku-ku,

Tokyo 169-8050, JAPAN

TEL/FAX +81 3 5286 9822

E-mail [kkuri@waseda.jp](mailto:kkuri@waseda.jp)

W. Michael Hanemann

Department of Agricultural & Resource Economics,

University of California, Berkeley

March, 2006

Draft

**The Integer Programming Approach to a Generalized Corner-Solution Model:  
An Application to Recreation Demand\***

**Abstract**

We developed an integer programming approach that takes into account the integral character of recreation behavior data and that has the same advantages as the Kuhn-Tucker continuous demand approach, which is consistent with utility theory, and provides a single structural framework for simultaneously modeling site selection and the participation decision. We then used the local search algorithm and the greedy method to estimate the parameters of the utility function and to calculate welfare changes. The results of our empirical study suggest that the integer programming approach provides a better fit to our data than does the continuous demand approach.

**KEY WORDS** welfare analysis, demand system, approximate algorithms, beach recreation

**JEL Classification:** C61, D12, Q26

## I. Introduction

Typical approaches to the modeling of consumer demand assume the continuity of consumption and an interior solution to the consumer's utility maximization problem. In many settings, however, these assumptions are unrealistic. For example, recreational activity demand exhibits the characteristics of counting data, that is, the number of trips is always a non-negative integer value. Moreover, while some who engage in recreational activity visit more than one site, most individuals make a small number of recreational trips, if any.

At least two strategies have been proposed in the literature to deal with such integral data and the possibility of corner solutions: the linked model and the Kuhn-Tucker model (Herriges et al., 1999). The linked model segments the consumer's decision into two components, namely site selection and the participation decision. Studies that have employed the linked model include those of Bockstael et al. (1986, 1987); Yen and Adamowicz (1994); Feather et al. (1995); Hausman et al. (1995); Parsons and Kealy (1995). The linked model can be used to deal with site selection and the participation decision; however, these are not derived from a unified utility theory<sup>1</sup>. The Kuhn-Tucker model, on the other hand, relies upon a single structural framework to simultaneously model site selection and the participation decision and allow for corner solutions. The Kuhn-Tucker model was initially developed by Wales and Woodland (1983) and Hanemann (1978), and was refined by Bockstael et al. (1986). Phaneuf et al. (2000, hereafter PKH) subsequently applied this model to recreation demand data. While Herriges et al. (1999) compared the Kuhn-Tucker model with the linked model, von Haefen and Phaneuf (2003) provided a comparison of this model with the count model. PKH's strategy for the analytical calculation of compensating variation is not feasible when the size of a choice set is relatively large; however, von Haefen et al. (2004, hereafter vHPP) recently proposed a Monte-Carlo Markov Chain algorithm for constructing compensating variations for data comprising a large choice set.

Although the Kuhn-Tucker model is appealing, in that it can be used to deal with corner solutions, and is derived from a unified utility theory, this model also relies upon the assumption of continuous consumption, which is unrealistic in the context of recreation demand.

In this paper, we suggest an alternative approach: the integer programming model, which is an integral version of the Kuhn-Tucker model. While having the same advantages as the Kuhn-Tucker continuous demand approach, namely a single structural framework that allows the simultaneous modeling of site selection and the participation decision, as well as the possibility of a corner solution, and consistency with utility theory, it also satisfies the constraint that recreation behavior data have the properties of count data. The integer programming problem is known to be NP-hard (Non-deterministic Polynomial-time hard), and it may be very difficult to find an exact solution to the problem (Wolsey, 1998; Wolsey and Nemhauser, 1999; Vazirani, 2001)<sup>2</sup>. Alternatively, approximate algorithms, such as local search, the greedy method, the genetic algorithm, simulated annealing, and tabu search, have been proposed in the integer and combinatorial optimization literature (Sait, 2000; Aarts and Lenstra, eds, 1997). Therefore, we investigated the use of approximate algorithms in estimating the parameters of the utility function and calculating compensating variations from both theoretical and empirical perspectives. In the empirical section, we discuss the application of our integer programming approach to beach recreation data for the region of Southern California and compare the results from our proposed model with those from a corresponding Kuhn-Tucker continuous demand model. Our empirical analysis suggests that the integer programming approach fits our data better than does the continuous demand approach, and the welfare loss of the beach closure estimated using the continuous demand model were lower than those found using the integer programming model.

The remainder of the paper is structured as follows. In section II., we provide overviews

of both the Kuhn-Tucker continuous demand model and our proposed integer programming model. Section III. discusses our strategy for constructing welfare changes using integer programming algorithms. Section IV. provides the results of our empirical analysis of beach recreation data for the region of Southern California. Finally, section V. provides concluding comments.

## II. Model

### A. *The Continuous Demand Model*

This section provides a brief overview of the Kuhn-Tucker continuous demand model, which was initially developed by Wales and Woodland (1983) and Hanemann (1978). This model assumes that all goods can be consumed in non-negative, perfectly divisible units. This assumption of continuous consumption implies that a consumer's utility maximization problem can be written as a non-linear programming problem and solved by the Kuhn-Tucker approach. In particular, the consumer's problem is to

$$\begin{aligned} & \text{Max } U(\mathbf{x}, \mathbf{Q}, z, \beta, \varepsilon) \\ & \text{s.t. } \mathbf{p}'\mathbf{x} + z \leq y, \quad z \geq 0, \quad x_j \geq 0, \quad j = 1, \dots, M \end{aligned} \tag{1}$$

where  $U$  is assumed to be a quasi-concave, increasing, and continuously differentiable function of  $(\mathbf{x}, z)$ ,  $\mathbf{x} = (x_1, \dots, x_M)'$  is a vector of goods to be analyzed,  $\mathbf{Q} = (\mathbf{q}_1, \dots, \mathbf{q}_M)'$  is an  $M \times K$  matrix of quality attributes for the  $M$  goods,  $z$  is the Hicksian composite good,  $\beta$  is a vector of parameters, and  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_M)'$  is a vector of random components that are assumed to be known to the individual but unknown to the analyst. Finally,  $\mathbf{p} = (p_1, \dots, p_M)'$  is a vector of prices, and  $y$  denotes income.

The first-order Kuhn-Tucker condition for the utility maximization problem is as fol-

lows:

$$\begin{aligned}
U_j(\mathbf{x}^*, \mathbf{Q}, y - \mathbf{p}'\mathbf{x}^*, \beta, \varepsilon) &\leq p_j \cdot U_z(\mathbf{x}^*, \mathbf{Q}, y - \mathbf{p}'\mathbf{x}^*, \beta, \varepsilon), \\
x_j^* &\geq 0, \\
x_j^* [U_j - p_j \cdot U_z] &= 0,
\end{aligned} \tag{2}$$

for  $j = 1, \dots, M$ , where  $U_j = \partial U / \partial x_j$ ,  $U_z = \partial U / \partial z$ , and  $\mathbf{x}^*$  is a solution vector of the consumer problem (1). Following PKH, we assume that  $\partial U_z / \partial \varepsilon_k = 0 \ \forall k$ ,  $\partial U_j / \partial \varepsilon_k = 0 \ \forall k \neq j$ , and  $\partial U_j / \partial \varepsilon_j > 0 \ \forall j$ ; then, the first-order condition can be rewritten as

$$\begin{aligned}
\varepsilon_j &\leq g_j(\mathbf{x}^*, \mathbf{p}, y, \mathbf{Q}, \beta), \\
x_j^* &\geq 0, \\
x_j^* [g_j(\mathbf{x}^*, \mathbf{p}, y, \mathbf{Q})] &= 0,
\end{aligned} \tag{3}$$

for  $j = 1, \dots, M$ , where  $g_j$  is the solution to  $U_j(\mathbf{x}^*, \mathbf{Q}, y - \mathbf{p}'\mathbf{x}^*, \beta, g_j) = p_j \cdot U_z(\mathbf{x}^*, \mathbf{Q}, y - \mathbf{p}'\mathbf{x}^*, \beta, g_j)$ . Furthermore, an individual who chooses to consume only the first  $k$  goods ( $x_j^* > 0 \ j = 1, \dots, k$ ,  $x_j^* = 0 \ j = k + 1, \dots, M$ ) contributes the following likelihood function:

$$l(\beta | \mathbf{x}^*, \mathbf{p}, y, \mathbf{Q}) = \int_{-\infty}^{g_{k+1}} \cdots \int_{-\infty}^{g_M} f_\varepsilon(g_1, \dots, g_k, \varepsilon_{k+1}, \dots, \varepsilon_M) \times \text{abs} |J_k| d\varepsilon_M \cdots d\varepsilon_{k+1}, \tag{4}$$

where  $f_\varepsilon$  is the joint density function and  $J_k$  is the Jacobian for the transformation of  $\varepsilon$  to  $(x_1, \dots, x_k, \varepsilon_{k+1}, \dots, \varepsilon_M)'$ .

## ***B. Integer Programming Model***

When the assumption of continuous consumption is not satisfied, the consumer's problem cannot be solved by means of traditional Kuhn-Tucker conditions. For example, this assumption is violated by the demand for recreation. Recreational trips can only be num-

bered in non-negative integers and, thus, the consumer's problem should be written as a non-linear integer programming problem. In this case, the consumer solves

$$\begin{aligned} & \text{Max } U(\mathbf{x}, \mathbf{Q}, z, \beta, \varepsilon) \\ & \text{s.t. } \mathbf{p}'\mathbf{x} + z \leq y, \quad z \geq 0, \quad \mathbf{x} \in \mathbb{Z}_+, \end{aligned} \tag{5}$$

where  $\mathbb{Z}_+$  is the set of non-negative integers. Let  $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_M)'$  and  $\tilde{z}$  characterize the solution to the above problem (5). If  $\tilde{x}_j$  is a corner solution ( $\tilde{x}_j = 0$ ), then the first-order Kuhn-Tucker condition  $U_j \leq p_j \cdot U_z$  still holds in the count demand model, because the corner solution satisfies the constraint that the solution be a non-negative integer. On the other hand, if  $\tilde{x}_j$  is an interior solution, the solution to the first-order condition  $U_j = p_j \cdot U_z$  may not be an integer; therefore, the Kuhn-Tucker condition cannot be used to solve problem (5) when  $\tilde{x}_j$  is an interior solution.

The integer programming problem is generally known to be NP-hard, and it may be very difficult to find an exact solution to this problem. As alternatives to the above approach, approximation algorithms, such as local search, the greedy method, the genetic algorithm, simulated annealing, and tabu search have been proposed. We focus on the local search algorithm, which tries iteratively to identify local optima within a given neighborhood (Aarts and Lenstra, eds, 1997; Wolsey, 1998). In particular, this algorithm proceeds as follows. Consider the neighborhood  $N^1(\mathbf{x}^1) = \{\mathbf{s} \in \mathbb{Z}_+ | d(\mathbf{x}^1, \mathbf{s}) \leq \alpha\}$ , where  $d(\mathbf{x}^1, \mathbf{s})$  is the Euclidean distance. If there is another allocation  $\mathbf{x}^2 \in N^1(\mathbf{x}^1)$  that yields a utility higher than  $\mathbf{x}^1$ , then consider the updated neighborhood  $N^2(\mathbf{x}^2) = \{\mathbf{s} \in \mathbb{Z}_+ | d(\mathbf{x}^2, \mathbf{s}) \leq \alpha\}$ , and iterate this process. If there is no allocation that satisfies  $\mathbf{x}^{n+1} \in N^n(\mathbf{x}^n)$  and  $U(\mathbf{x}^{n+1}) > U(\mathbf{x}^n)$ , then define the consumption  $\mathbf{x}^n$  as the approximate solution of the integer programming problem.

Consider first  $N(\tilde{\mathbf{x}}) = \{\mathbf{s} \in \mathbb{Z}_+ | d(\tilde{\mathbf{x}}, \mathbf{s}) \leq 1\}$ , which is the neighborhood of  $\tilde{\mathbf{x}}$  having a Euclidean distance less than or equal to one. From the point of view of the local search

algorithm in neighborhood  $N(\tilde{\mathbf{x}})$ ,  $\tilde{\mathbf{x}}$  corresponds to approximate optimal consumption if and only if

$$\begin{aligned} U(\tilde{\mathbf{x}}, \mathbf{Q}, y - \mathbf{p}'\tilde{\mathbf{x}}, \beta, \varepsilon) &\geq U(\tilde{\mathbf{x}}_j^+, \mathbf{Q}, y - \mathbf{p}'\tilde{\mathbf{x}}_j^+, \beta, \varepsilon), \\ U(\tilde{\mathbf{x}}, \mathbf{Q}, y - \mathbf{p}'\tilde{\mathbf{x}}, \beta, \varepsilon) &\geq U(\tilde{\mathbf{x}}_j^-, \mathbf{Q}, y - \mathbf{p}'\tilde{\mathbf{x}}_j^-, \beta, \varepsilon), \end{aligned} \quad \text{for all } j = 1, 2, \dots, M \quad (6)$$

where  $\tilde{\mathbf{x}}_j^+ = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_{j-1}, \tilde{x}_j+1, \tilde{x}_{j+1}, \dots, \tilde{x}_M)$ ,  $\tilde{\mathbf{x}}_j^- = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_{j-1}, \tilde{x}_j-1, \tilde{x}_{j+1}, \dots, \tilde{x}_M)$ . Next, consider the utility differences  $d_j^+(\varepsilon|\tilde{\mathbf{x}}, \mathbf{Q}, y - \mathbf{p}'\tilde{\mathbf{x}}, \beta) = U(\tilde{\mathbf{x}}, \mathbf{Q}, y - \mathbf{p}'\tilde{\mathbf{x}}, \beta, \varepsilon) - U(\tilde{\mathbf{x}}_j^+, \mathbf{Q}, y - \mathbf{p}'\tilde{\mathbf{x}}_j^+, \beta, \varepsilon)$  and  $d_j^-(\varepsilon|\tilde{\mathbf{x}}, \mathbf{Q}, y - \mathbf{p}'\tilde{\mathbf{x}}, \beta) = U(\tilde{\mathbf{x}}, \mathbf{Q}, y - \mathbf{p}'\tilde{\mathbf{x}}, \beta, \varepsilon) - U(\tilde{\mathbf{x}}_j^-, \mathbf{Q}, y - \mathbf{p}'\tilde{\mathbf{x}}_j^-, \beta, \varepsilon)$ . Note that either  $d_j^+$  or  $d_j^-$  is the respective utility difference for the increment or decrement of one unit in  $\tilde{x}_j$ , and that other goods ( $\tilde{x}_k, \forall k \neq j$ ) do not change in quantity; therefore, these utility differences can be written as  $d_j^+ = \int_{\tilde{x}_j+1}^{\tilde{x}_j} (U_j - p_j U_z) dx_j$  and  $d_j^- = \int_{\tilde{x}_j-1}^{\tilde{x}_j} (U_j - p_j U_z) dx_j$ . From the assumptions of  $\partial U_z / \partial \varepsilon_k = 0 \forall k$ ,  $\partial U_j / \partial \varepsilon_k = 0 \forall k \neq j$ , and  $\partial U_j / \partial \varepsilon_j > 0 \forall j$ , it is possible to show that  $\partial d_j^+ / \partial \varepsilon_j < 0$ ,  $\partial d_j^- / \partial \varepsilon_j > 0$ , and  $\partial d_j^+ / \partial \varepsilon_k = \partial d_j^- / \partial \varepsilon_k = 0, \forall k \neq j$ . Thus, the approximate optimal condition (6) can be rewritten as

$$h_j^-(\tilde{\mathbf{x}}, \mathbf{p}, y, \mathbf{Q}, \beta) \leq \varepsilon_j \leq h_j^+(\tilde{\mathbf{x}}, \mathbf{p}, y, \mathbf{Q}, \beta), \quad (7)$$

where  $h_j^+(\tilde{\mathbf{x}}, \mathbf{p}, y, \mathbf{Q}, \beta)$  is the solution to  $d_j^+(h_j^+|\tilde{\mathbf{x}}, \mathbf{Q}, y - \mathbf{p}'\tilde{\mathbf{x}}, \beta) = 0$  and  $h_j^-(\tilde{\mathbf{x}}, \mathbf{p}, y, \mathbf{Q}, \beta)$  is the solution to  $d_j^-(h_j^-|\tilde{\mathbf{x}}, \mathbf{Q}, y - \mathbf{p}'\tilde{\mathbf{x}}, \beta) = 0$ . Figure 1 illustrates the approximate optimal condition, as given in relation (7). The curves represent utility given the consumptions  $\tilde{\mathbf{x}}$ ,  $\tilde{\mathbf{x}}_j^+$ , and  $\tilde{\mathbf{x}}_j^-$ . Consider the case of  $\partial U / \partial \varepsilon_j > 0$ . Given the assumptions of  $\partial U_z / \partial \varepsilon_k = 0 \forall k$ ,  $\partial U_j / \partial \varepsilon_k = 0 \forall k \neq j$ , and  $\partial U_j / \partial \varepsilon_j > 0 \forall j$ , the curve  $U(\tilde{\mathbf{x}}_j^+)$  is steeper than  $U(\tilde{\mathbf{x}})$  and  $U(\tilde{\mathbf{x}}_j^-)$  but gentler than  $U(\tilde{\mathbf{x}})$ . This figure shows that  $U(\tilde{\mathbf{x}})$  is greater than or equal to  $U(\tilde{\mathbf{x}}_j^+)$  and  $U(\tilde{\mathbf{x}}_j^-)$  when the unobserved component  $\varepsilon_j$  satisfies the approximate optimal condition  $h_j^- \leq \varepsilon_j \leq h_j^+$ . The condition for the case of  $\partial U / \partial \varepsilon_j \leq 0$  may similarly be derived. When interior solutions exist for the first  $k$  goods while corner solutions obtain for the other goods,

an individual contributes the following likelihood function:

$$l(\beta|\tilde{\mathbf{x}}, \mathbf{p}, y, \mathbf{Q}) = \int_{h_1^-}^{h_1^+} \cdots \int_{h_k^-}^{h_k^+} \int_{-\infty}^{g_{k+1}} \cdots \int_{-\infty}^{g_M} f_\varepsilon(\boldsymbol{\varepsilon}) d\varepsilon_M \cdots d\varepsilon_1. \quad (8)$$

FIGURE 1. APPROXIMATE OPTIMAL CONDITION

### C. Preference Specification

The empirical application employs the following additively separable utility function suggested by vHPP:

$$\begin{aligned} U(\mathbf{x}, \mathbf{Q}, z, \beta, \boldsymbol{\varepsilon}) &= \sum_j^M \left[ \Psi(\varepsilon_j) \ln(\phi(\mathbf{q}_j)x_j + \theta_j) \right] + \frac{1}{\rho} z^\rho, \\ \ln \Psi(\varepsilon_j) &= \delta + \varepsilon_j, \\ \ln \phi(\mathbf{q}_j) &= \boldsymbol{\gamma}' \mathbf{q}_j, \end{aligned} \quad (9)$$

where  $\theta_j > 0$ ,  $\rho \leq 1$ ,  $\delta$ , and  $\boldsymbol{\gamma}$  are parameters. Note that this specification introduces quality through the simple repackaging parameter  $\phi$  (Griliches, 1964), and that weak complementarity is satisfied for all parameter values ( $\partial U / \partial q_j = 0$ , if  $x_j = 0$ ,  $\forall j$ ; see Mäler (1974)). In addition, assume that the  $\varepsilon_j$  are independent and identically distributed draws from the type-I extreme value distribution with inverse scale parameter  $\mu$ , for all  $j$ .

The likelihood function of the continuous demand model is then as follows:

$$l(\beta|\mathbf{x}^*, \mathbf{p}, y, \mathbf{Q}) = |\mathbf{J}| \prod_j \left( \left[ \frac{1}{\mu} \exp\left(\frac{-g_j}{\mu}\right) \right]^{1[x_j^* > 0]} G(g_j/\mu) \right), \quad (10)$$

where

$$1[x_j^* > 0] = \begin{cases} 1 & \text{if } x_j^* > 0 \\ 0 & \text{otherwise} \end{cases}, \quad (11)$$

$$G(s) = \exp[-\exp(-s)], \quad (12)$$

$$g_j = -\delta + \ln p_j - \gamma' \mathbf{q}_j - (1 - \rho) \ln(y - \mathbf{p}' \mathbf{x}^*) + \ln(\phi(\mathbf{q}_j)x_j^* + \theta_j). \quad (13)$$

For the integer programming model, assume that the  $j$ th demand has a corner solution when the observed  $j$ th demand is zero, and that the observed demand can be approximated as the solution obtained using the local search algorithm. Then, the likelihood function for this model is given by

$$l(\beta | \tilde{\mathbf{x}}, \mathbf{p}, y, \mathbf{Q}) = \prod_j (1[\tilde{x}_j > 0] \cdot [G(h_j^+/\mu) - G(h_j^-/\mu)] + 1[\tilde{x}_j = 0] \cdot G(g_j/\mu)), \quad (14)$$

where

$$1[\tilde{x}_j > 0] = \begin{cases} 1 & \text{if } \tilde{x}_j > 0 \\ 0 & \text{otherwise} \end{cases}, \quad 1[\tilde{x}_j = 0] = \begin{cases} 1 & \text{if } \tilde{x}_j = 0 \\ 0 & \text{otherwise} \end{cases}, \quad (15)$$

$h_j^+(\tilde{\mathbf{x}}, \mathbf{p}, y, \mathbf{Q}, \beta)$  is the solution to  $U(\tilde{\mathbf{x}}, \mathbf{Q}, y - \mathbf{p}' \tilde{\mathbf{x}}, \beta, h_j^+) - U(\tilde{\mathbf{x}}_j^+, \mathbf{Q}, y - \mathbf{p}' \tilde{\mathbf{x}}_j^+, \beta, h_j^+) = 0$ , and  $h_j^-(\tilde{\mathbf{x}}, \mathbf{p}, y, \mathbf{Q}, \beta)$  is the solution to  $U(\tilde{\mathbf{x}}, \mathbf{Q}, y - \mathbf{p}' \tilde{\mathbf{x}}, \beta, h_j^-) - U(\tilde{\mathbf{x}}_j^-, \mathbf{Q}, y - \mathbf{p}' \tilde{\mathbf{x}}_j^-, \beta, h_j^-) = 0$ . In particular,

$$h_j^+ = -\delta + \ln [\tilde{z}^\rho - (\tilde{z} - p_j)^\rho] - \ln \rho - \ln [\ln(\phi(\mathbf{q}_j)(\tilde{x}_j + 1) + \theta_j) - \ln(\phi(\mathbf{q}_j)\tilde{x}_j + \theta_j)], \quad (16)$$

and

$$h_j^- = -\delta + \ln [(\tilde{z} + p_j)^\rho - \tilde{z}^\rho] - \ln \rho - \ln [\ln(\phi(\mathbf{q}_j)\tilde{x}_j + \theta_j) - \ln(\phi(\mathbf{q}_j)(\tilde{x}_j - 1) + \theta_j)]. \quad (17)$$

### III. Welfare Estimation

#### A. Continuous Demand Model

Let  $V(\mathbf{p}, y, \mathbf{Q}, \beta, \varepsilon)$  be the solution to the utility maximization problem for the case of continuous demand. Under the continuous demand approach, the compensating variation (*CV*) associated with the change in price and quality from  $(\mathbf{p}^0, \mathbf{Q}^0)$  to  $(\mathbf{p}^1, \mathbf{Q}^1)$  is defined by

$$V(\mathbf{p}^0, y, \mathbf{Q}^0, \beta, \varepsilon) = V(\mathbf{p}^1, y - CV, \mathbf{Q}^1, \beta, \varepsilon). \quad (18)$$

Unless preferences are homothetic or quasi-linear in income, no closed-form solution for the *CV* exists; thus, an iterative search procedure is required in such a situation. However, when the dimension of the choice set is large, solving the *CV* analytically, as done by PKH, is not feasible. As an alternative, vHPP proposed a computationally tractable strategy that numerically solves the Kuhn-Tucker conditions for the optimal consumption levels. Their algorithm is as follows:

vHPP's numerical bisection algorithm for searching  $(\mathbf{x}^*, z^*)$

1. At iteration  $i$ , set  $z_a^i = (z_l^{i-1} + z_u^{i-1})/2$ . To initialize the algorithm, set  $z_l^0 = 0$  and  $z_u^0 = y$ .
2. Conditional on  $z_a^i$ , solve for  $\mathbf{x}^i$  and  $z^i$  using the Kuhn-Tucker first-order condition and the budget constraint.
3. If  $z^i > z_a^i$ , set  $z_l^i = z_a^i$  and  $z_u^i = z_u^{i-1}$ . Otherwise, set  $z_l^i = z_l^{i-1}$  and  $z_u^i = z_a^i$ .
4. Iterate until  $abs(z_l^i - z_u^i) \leq c$  where  $c$  is arbitrarily small.

This procedure will yield the optimal consumption  $(\mathbf{x}^*, z^*)$  and optimal utility level  $U^*$ , which are evaluated conditional on  $(\mathbf{p}, \mathbf{Q}, y, \beta, \varepsilon)$ . This algorithm could also be nested and then used to solve for the *CV* that equates utility before and after the price and quality

changes. For the continuous demand model, vHPP showed that the expectation of the compensating variation  $E[CV]$  can be estimated using the first-order condition and a simulation technique for the unobserved components  $\varepsilon$ . Specifically, if good  $j$  is consumed in a strictly positive quantity ( $x_j^* > 0$ ), the first-order condition of equation (3) implies  $\varepsilon_j = g_j(\mathbf{x}^*, \mathbf{p}, y, \mathbf{Q}, \beta)$ . Otherwise, given the i.i.d. type-I extreme value distribution assumption,  $\varepsilon_j < g_j(\mathbf{x}^*, \mathbf{p}, y, \mathbf{Q}, \beta)$  can be simulated. Thus, the unobserved components  $\varepsilon$  can be constructed as

$$\varepsilon_j = g_j(\mathbf{x}^*, \mathbf{p}, y, \mathbf{Q}, \beta) \text{ for } x_j^* \text{ is inner solution,} \quad (19)$$

$$\varepsilon_j = -\ln(-\ln(G(g_j/\mu)\tau))\mu \text{ for } x_j^* \text{ is corner solution,} \quad (20)$$

where  $G$  is the i.i.d. type-I extreme value distribution with a scale of one and  $\tau$  is a random draw from the uniform distribution.

## ***B. Integer Programming Model***

Let  $\tilde{V}(\mathbf{p}, y, \mathbf{Q}, \beta, \varepsilon)$  be the solution of the utility maximization integer programming problem (5). Under the count demand approach, the compensating variation ( $CV$ ) associated with the change in the price and quality from  $(\mathbf{p}^0, \mathbf{Q}^0)$  to  $(\mathbf{p}^1, \mathbf{Q}^1)$  is defined by

$$\tilde{V}(\mathbf{p}^0, y, \mathbf{Q}^0, \beta, \varepsilon) = \tilde{V}(\mathbf{p}^1, y - CV, \mathbf{Q}^1, \beta, \varepsilon). \quad (21)$$

The strategy of vHPP for solving for the  $CV$  cannot be applied to the count demand model, because their algorithm uses the Kuhn-Tucker first-order condition, which may not be satisfied under the constraint of integral demand. Alternatively, we propose to use the greedy method to search for the  $CV$ . The development of the greedy method has been documented in the integer programming literature (Aarts and Lenstra, eds, 1997; Wolsey, 1998). This algorithm always makes the choice that seems best at the moment, while finding a solution

by means of an iterative process. The greedy method is an approximate algorithm and does not always yield an optimal solution; however, this approach is known to perform well in many maximization problems where the target function is globally concave. The greedy strategy for finding the solution  $(\tilde{\mathbf{x}}, \tilde{z})$  to the integer programming problem is as follows:

1. At iteration  $i$ , construct  $\mathbf{x}_j^{i+1} = (\tilde{x}_1^i, \tilde{x}_2^i, \dots, \tilde{x}_{j-1}^i, \tilde{x}_j^i+1, \tilde{x}_{j+1}^i, \dots, \tilde{x}_M^i) \quad \forall j = 1, 2, \dots, M$  and  $z^{i+1} = y - \mathbf{p}'\mathbf{x}_j^{i+1}$ . To initialize the algorithm, set  $\tilde{\mathbf{x}}^0 = \mathbf{0}$  and  $\tilde{z}^0 = y$ .
2. If there exists an  $\mathbf{x}_j^{i+1}$  such that  $U(\mathbf{x}_j^{i+1}, \mathbf{Q}, z_j^{i+1}, \beta, \varepsilon) > U(\mathbf{x}_k^{i+1}, \mathbf{Q}, z_k^{i+1}, \beta, \varepsilon) \quad \forall k \neq j$ , set  $\tilde{\mathbf{x}}^{i+1} = \mathbf{x}_j^{i+1}$  and  $\tilde{z}^{i+1} = z_j^{i+1}$ . Otherwise, stop the iteration.
3. Iterate this process.

This procedure will yield the approximate optimal consumption  $(\tilde{\mathbf{x}}, \tilde{z})$  under the constraints that consumption be non-negative and integral. Because income is a continuous variable, a combination of the greedy method and vHPP's numerical bisection algorithm could be used to solve the CV, which is defined by equation (21). Consider the cases of a price decrease or a quality improvement. The search algorithm for the CV is as follows:

1. Calculate the baseline utility  $\tilde{V}^0 = \tilde{V}(\mathbf{p}^0, y, \mathbf{Q}^0, \beta, \varepsilon)$  using the greedy method.
2. At iteration  $n$ , set  $CV_a^n = (CV_l^{n-1} + V_u^{n-1})/2$ . To initialize the algorithm, set  $CV_l^0 = 0$  and  $CV_u^0 = y$ .
3. Conditional on  $(\mathbf{p}^1, y - CV_a^n, \mathbf{Q}^1, \beta, \varepsilon)$ , solve for  $\tilde{\mathbf{x}}^n$  and  $\tilde{z}^n$  using the greedy method.
4. Construct the utility  $\tilde{V}^n = \tilde{V}(\mathbf{p}^1, y - CV_a^n, \mathbf{Q}^1, \beta, \varepsilon)$ . If  $V^n > V^0$ , set  $CV_l^n = CV_a^n$  and  $CV_u^n = CV_u^{n-1}$ . Otherwise, set  $CV_l^n = CV_l^{n-1}$  and  $CV_u^n = CV_a^n$ .
5. Iterate until  $abs(\tilde{V}^n - \tilde{V}^0) \leq c$  where  $c$  is arbitrarily small.

From the approximate optimal condition of the integer programming problem (7) and the simulation technique for the error components  $\varepsilon$ , the expectation of the compensating variation  $E[CV]$  can be estimated. If the optimal value of good  $j$  is corner solution, given the i.i.d. type-I extreme value distribution assumption, then  $\varepsilon_j$  can be simulated using (20). If the optimal value of good  $j$  is an interior solution, then the  $\varepsilon_j$  that satisfies the approximate optimal condition (7) can be simulated as

$$\varepsilon_j = -\ln\left(-\ln\left((1-\tau)G(h_j^+/\mu) + \tau \cdot G(h_j^-/\mu)\right)\right)\mu, \quad (22)$$

where  $G$  is the i.i.d. type-I extreme value distribution with a scale of one and  $\tau$  is a random draw from the uniform distribution<sup>3</sup>.

## IV. Application

### A. Data

To illustrate the integer programming approach, we applied this model to beach recreation data for the region of Southern California. Each year, the beaches of Southern California attract 175 million visitors, who spend more than \$1.5 billion during their visits (California EPA, 2001). A panel of Southern California respondents was recruited between November 1999 and January 2000; these recruits were asked to keep a record of all their trips to local beaches between December 1999 and November 2000. Six bi-monthly diary surveys (waves 1-6) were conducted, and, at the end of each wave, the panel members were interviewed by phone. The resulting sample consisted of 4,367 trips taken by the 922 respondents, who supplied sufficient data for our analysis. To be modeled, a trip needed to have been made to a valid destination, and the panelist who took the trip had to have supplied all of the demographic variables that are included in the model and were needed

to calculate travel costs. Multiple beach destination trips or multi-day trips were excluded from the analysis. Table 1 describes the distribution of our trip data for all waves. Note that, while 33% of the respondents visited no beaches, 37% visited more than one site. Thus, the generalized corner-solution model was applicable to our data.

TABLE 1. DISTRIBUTION OF TRIPS

Fifty-two beach sites were considered in the collection of the data. The attributes of these beaches and of other selected variables are described in Table 2. The travel cost was specified to be a function of the respondent's reported income, the estimated cost of vehicle operation (\$0.145/mile), travel time, and the distance between the respondent's residence and each beach in the choice set. The one-way travel distance and travel time between each respondent's address and the beach address were calculated using the computer program PC-Miler. To be specific, the cost was calculated as

$$\text{cost} = 2 * [\text{one way travel dist} * 0.145 + (\text{one way travel time}) * (0.5 * \text{hourly wage})].$$

Beach water and sand quality data were obtained from the Southern California non-profit group Heal the Bay. These data contained ratings for beach water quality on dry days, on a scale of A+ to F, at numerous monitoring stations throughout Southern California, for a selection of dates between June 1998 and April 2001.

TABLE 2. VARIABLE LIST

## ***B. Results***

We compared two models: the Kuhn-Tucker model with continuous demand (KT) and the integer programming model (IP). Table 3 provides the maximum-likelihood parameter estimates. All of the estimated parameters had the expected signs and were statistically

significant at the 1% level. In both models, the impacts of the variables “condoshotels,” “bikepath,” “sandy,” “tidepools,” and “water” took on positive signs, while the variables “rocky” and “oilpumps” had negative effects on utility. Such log-likelihood values might initially suggest that the integer programming model (IP) provides a better fit to the data than does the continuous demand model (KT); however, the estimated parameters were, in fact, similar across the two models.

TABLE 3. ESTIMATED PARAMETERS

### ***C. Welfare Estimation***

We analyzed two scenarios designed to provide policy implications for beach recreation: (A) a closing of the Newport, Huntington City, and Huntington State beaches; and (B) a 20% reduction in the water quality level at all beaches. The first policy scenario concerns possible beach closure resulting from an oil spill or water quality deterioration. For example, on February 7, 1990, the tanker vessel *American Trader* ran aground and spilled 416,598 gallons of crude oil into the Pacific Ocean offshore of Huntington Beach. In consequence, approximately 14 miles of Newport and Huntington beaches were closed for a period of up to 34 days. For details of the *American Trader* case, see Chapman and Hane-mann (2001). In addition, Huntington Beach was also closed during July and August 1999, due to elevated levels of total coliform, fecal coliform, and enterococci bacteria. In our sample, 127 visitors took a total of 636 trips to Newport Beach; 84 visitors took 260 trips to Huntington City Beach; and 64 visitors took 179 trips to Huntington State Beach. The second scenario considers the possible welfare loss due to water pollution. For example, there were 2,584 beach closures and advisories in Los Angeles and Orange County in 2002 (Natural Resources Defense Council, 2003). The causes of these beach closures and advisories included water pollution by storm drain run-off, discharges from point sources, and

sewage spills.

Table 4 reports predicted trips for the baseline case (status quo), as well as for the two scenarios. The values in parentheses are standard errors that were estimated using the one-hundred-repetition procedure proposed by Krinsky and Robb (1986)<sup>4</sup>. Note that the mean of observed trips was 4.736; thus, Table 4 indicates that the integer programming model (IP), as well as continuous demand model (KT), is capable of accurately predicting baseline trips when the unobserved components are simulated by means of equations (19), (20), and (17). This means that our proposed strategy for finding the solution of the integer programming problem performs well. In the case of the beach closure scenario, there were no differences in the trips predicted by the continuous demand and integer programming models. On the other hand, the predicted demand estimated using the continuous demand model for the scenario of water quality deterioration was slightly higher than that resulting from the integer programming model estimation; however, this difference was not significant.

TABLE 4. PREDICTED TRIPS

The compensating variation for each of the two scenarios is reported in Table 5. The absolute values of the welfare losses in both scenarios were estimated to be lower under the continuous demand model than under the integer programming model. The difference in the estimated compensating variations for the water quality deterioration scenario was not significant; however, the difference in the welfare losses for the beach closing scenario was significant at the 5% level. Thus, the continuous demand model may have underestimated the welfare losses, even though the difference between compensating variations for the continuous demand model and the integer programming model were not relatively large.

TABLE 5. COMPENSATING VARIATIONS

## V. Concluding Comments

In this paper, we considered integer programming models of recreation demand from both theoretical and empirical perspectives. We developed a methodology for estimating the parameters of a utility function and the corresponding compensating variation that uses approximate solution approaches to integer programming problems. Our proposed model is appealing in that it satisfies the constraint imposed by the integral nature of recreation behavior data, in addition to having the same advantages as the Kuhn-Tucker continuous demand approach, which provides a single structural framework that simultaneously models site selection and the participation decision, while allowing for the possibility of a corner solution, in addition to being consistent with utility theory. It is generally difficult to find exact solutions to integer programming problems; therefore, we have suggested the use of alternative approximate approaches, including the local search algorithm and the greedy method, both of which have been widely used in other studies of integer programming. With an empirical application using data for Southern California beaches, we compared our proposed model with the continuous demand model. The empirical results suggested that the integer programming approach generally provides a better fit for our data than does the continuous demand approach. Moreover, the absolute values of the compensating variations estimated using the continuous demand model were lower than those found using the integer programming model.

We suggest three additional paths of investigation into the integer programming approach. First, the assumption that preferences are additively separable should be relaxed and the implications of this modification examined. The assumption of additive separability is not necessary for the application of our proposed strategy for solving compensating variations, while this assumption is essential for vHPP's numerical bisection algorithm. Therefore, further conceptual and empirical study should address problems with non-additively

separable utility functions using the integer programming approach. Second, other integer programming strategies should be investigated. We have used the local search algorithm and the greedy method, which are straightforward techniques and relatively easy to implement. Although our empirical study showed that our proposed strategy for finding the approximate solution of an integer programming problem performs well, more complex and sophisticated algorithms, such as the genetic algorithm, simulated annealing, and tabu search, have also been proposed in integer programming studies for this purpose. The application of these algorithms should be investigated as a confirmation of the accuracy of our proposed strategy. Finally, the robustness of integer programming estimation should be investigated. Several ways of estimating preference parameters exist; these include the simulated method of moments, as used by Dube (2004) and Hendel (1999), a discrete choice framework with all combinations of multiple-choice alternatives, and the integer programming approach that we have proposed. Our proposed strategy is more general than are the other approaches because it relaxes the assumption of a zero income effect, which is essential in the other cases; however, a comparative estimation robustness check is nonetheless of importance.

## REFERENCES

- Aarts, Emilie, and Jan Karel Lenstra, eds (1997) *Local Search in Combinatorial Optimization* Wiley-Interscience Series in Discrete Mathematics and Optimization (Chichester, England; New York: John Wiley & Sons)
- Bockstael, Nancy, W. Michael Hanemann, and Catherine L. Kling (1987) 'Estimating the value of water quality improvements in a recreation demand framework.' *Water Resources Research* 23(5), 951–960

- Bockstael, Nancy, W. Michael Hanemann, and Ivar E. Strand (1986) 'Measuring the benefits of water quality improvements using recreation demand models, vol. 2'
- California EPA (2001) 'California beach closure report 2000'
- Chapman, David J., and W. Michael Hanemann (2001) 'Environmental damages in court: The american trader case.' In *The Law and Economics of the Environment*, ed. Anthony Heyes (Edward Elgar Publishing) pp. 319–367
- Dube, Jean-Pierre (2004) 'Multiple discreteness and product differentiation: Demand for carbonated soft drinks.' *Marketing Science* 23(1), 66–81
- Feather, Peter, Daniel Hellerstein, and Theodore Tomasi (1995) 'A discrete-count model of recreational demand.' *Journal of Environmental Economics and Management* 29(2), 214–227
- Griliches, Zvi (1964) 'Notes on the measurement of price and quality changes.' In 'Models of Income Determination, NBER Studies in Income and Wealth, Volume 28.' (Princeton, NJ.: Princeton University Press.) pp. 381–418
- Hanemann, W. Michael (1978) 'A theoretical and empirical study of the recreation benefits from improving water quality in the boston area.' PhD dissertation, Department of Economics, Harvard University
- Hausman, Jerry A, Gregory K Leonard, and Daniel 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), 1–30
- Hendel, Igal (1999) 'Estimating multiple-discrete choice models: An application to computerization returns.' *Review of Economic Studies* 66(2), 423–446

- Herriges, Joseph A., Catherine L. Kling, and Daniel J. Phaneuf (1999) 'Corner solution models of recreation demand: A comparison of competing frameworks.' In *Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice*, ed. Joseph A. Herriges and Catherine L. Kling (New Horizons in Environmental Economics. Cheltenham, U.K. and Northampton, Mass.: Edward Elgar) pp. 163–197
- Krinsky, Itzhak, and A. Leslie Robb (1986) 'On approximating the statistical properties of elasticities.' *Review of Economics and Statistics* 68(4), 715–719
- Mäler, Karl-Goran (1974) *Environmental Economics: A Theoretical Inquiry* (Baltimore, MD: Johns Hopkins University Press for Resources for the Future.)
- Natural Resources Defense Council (2003) 'Testing the waters: A guide to water quality at vacation beaches'
- Parsons, George R, and Mary Jo Kealy (1995) 'A demand theory for number of trips in a random utility model of recreation.' *Journal of Environmental Economics and Management* 29(3), 357–367
- Phaneuf, Daniel J., Catherine L. Kling, and Joseph A. Herriges (2000) '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
- Sait, Sadiq M. (2000) *Iterative Computer Algorithms with Applications in Engineering: Solving Combinatorial Optimization Problems* (Wiley-IEEE Computer Society Pr.)
- Smith, V. Kerry (1997) 'Combining discrete choice and count data models: A comment'
- Train, Kenneth (2003) *Discrete Choice Methods with Simulation* (New York: Cambridge University Press)

- Vazirani, Vijay V. (2001) *Approximation Algorithms* (Berlin; New York: Springer)
- von Haefen, Roger, and Daniel Phaneuf (2003) 'Estimating preferences for outdoor recreation: A comparison of continuous and count data demand system frameworks.' *Journal of Environmental Economics and Management* 45(3), 612–630
- von Haefen, Roger H., Daniel J. Phaneuf, and George R. Parsons (2004) 'Estimation and welfare analysis with large demand systems.' *Journal of Business and Economic Statistics* 22(2), 194–205
- Wales, Terence J., and Alan Woodland (1983) 'Estimation of consumer demand systems with binding non-negativity constraints.' *Journal of Econometrics* 21(3), 263–285
- Wolsey, Laurence A. (1998) *Integer Programming* Wiley-Interscience Series in Discrete Mathematics and Optimization (New York: Wiley-Interscience)
- Wolsey, Laurence A., and George L. Nemhauser (1999) *Integer and Combinatorial Optimization* (Wiley-Interscience)
- Yen, Steven T, and Wiktor L Adamowicz (1994) 'Participation, trip frequency and site choice: A multinomial-poisson hurdle model of recreation demand.' *Canadian Journal of Agricultural Economics* 42(1), 65–76

## Notes

\*We thank Roger H. von Haefen, Linwood Pendleton, David Layton, Craig Mohn, James Hilger, Catherine Kling, Edward Morey and participants at the ASSA/AERE Winter Meeting in Philadelphia, PA, January 2005 for helpful comments and suggestions. The authors gratefully acknowledge the financial support from the National Oceanic and Atmospheric Administration (NOAA).

<sup>1</sup> Hausman et al. (1995) claimed that their model is consistent with utility theory; however, Smith (1997) demonstrated that their proof of consistency relies upon an assumption that holds only under restrictive circumstances. See also Herriges et al. (1999).

<sup>2</sup> Hendel (1999) and Dube (2004) assumed that preferences were additively separable and quasi-linear in their solution of the integer programming model. Such strong assumptions make solving the integer programming problem considerably easier. Given additive separability and zero income effects, the demand for each good is completely independent and one need not worry about interaction effects among goods.

<sup>3</sup> Note that  $\varepsilon_j$  is truncated, as specified in relation (7). See Train (2003) for details regarding random draws from truncated distributions.

<sup>4</sup> The number of iterations was restricted by the burden of the welfare computation for the integer programming model. For example, the one hundred iterations of this model that were used to estimate the standard errors given in Table 5 required more than ten hours on a Pentium 4 CPU 1.5GHz computer.

TABLE 1. DISTRIBUTION OF TRIPS

number of trips take per respondent	number of observations		number of sites visited per respondent	number of observations	
0	305	33%	0	305	33%
1	131	14%	1	279	30%
2	117	13%	2	167	18%
3	68	7%	3	80	9%
4	49	5%	4	44	5%
5	41	4%	5	29	3%
6	30	3%	6	13	1%
7	24	3%	7	2	0%
8	13	1%	8	0	0%
9	22	2%	9	1	0%
10-	122	13%	10-	2	0%
Total	922	100%	Total	922	100%

TABLE 2. VARIABLE LIST

Variables	Description	Mean	Standard Deviation
trips	number of trips	4.736	9.968
cost	travel cost	50.242	42.173
income	household income (\$)	62,148	43,694
condoshotels	1/0, 1 = condo(s) or hotel(s) visible from beach	0.327	0.474
bikepath	1/0, 1 = bike path available	0.442	0.502
sandy	1/0, 1 = beach is sandy	0.865	0.345
rocky	1/0, 1 = beach is rocky	0.462	0.503
tidepools	1/0, 1 = tide pools at the beach	0.269	0.448
oilpumps	1/0, 1 = on-shore oil pumps visible from beach	0.039	0.194
water	the average grade of beach water quality	3.456	1.023

TABLE 3. ESTIMATED PARAMETERS<sup>a,b</sup>

		The Kuhn-Tucker Continuous Demand (KT)		The Integer Programming (IP)	
Psi ( $\ln \Psi$ )	$\delta$	-5.5151 ***		-5.7842 ***	
		(-14.50)		(-5.78)	
Quality Index ( $\ln \phi$ )	condoshotels	0.7627 ***		0.7893 ***	
		(14.62)		(14.66)	
	bikepath	0.2972 ***		0.3115 ***	
		(4.77)		(4.90)	
	sandy	1.5179 ***		1.5534 ***	
		(9.19)		(9.28)	
	rocky	-0.2492 ***		-0.2522 ***	
		(-3.52)		(-3.44)	
	tidepools	0.7409 ***		0.7573 ***	
		(8.65)		(8.50)	
	oilpumps	-1.4512 ***		-1.4268 ***	
		(-7.43)		(-7.12)	
	water	0.1025 ***		0.1063 ***	
		(3.31)		(3.59)	
Rho	$\ln(1 - \rho)$	-0.5545 ***		-0.5454 ***	
		(-9.43)		(-3.51)	
Translating	$\ln \theta$	3.2329 ***		3.1359 ***	
		(15.00)		(14.96)	
Scale	$\mu$	0.9666 ***		0.9818 ***	
		(61.90)		(55.35)	
Log-Likelihood		-7958.6		-7888.1	

<sup>a</sup> Significance at the 1% level is indicated by three asterisks.

<sup>b</sup> t-statistics in parentheses

TABLE 4. PREDICTED TRIPS <sup>a</sup>

	The Kuhn-Tucker Continuous Model ( $x_{KT}$ )	The Integer Programming ( $x_{IP}$ )	$x_{IP} - x_{KT}$
baseline	4.74 (0.00)	4.74 (0.00)	0.00 (0.00)
closing Newport beach and Huntington beaches	3.57 (0.00)	3.57 (0.00)	0.00 (0.00)
20% reduction in the water quality level	4.51 (0.06)	4.48 (0.07)	-0.03 (0.10)

<sup>a</sup> The values in parentheses are standard errors calculated using the procedure reported by Krinsky and Robb (1986) and based on one hundred iteration.

TABLE 5. COMPENSATING VARIATIONS <sup>a</sup>

	The Kuhn-Tucker Continuous Model ( $CV_{KT}$ )	The Integer Programming ( $CV_{IP}$ )	$CV_{IP} - CV_{KT}$
closing Newport beach and Huntington beaches	-20.97 (0.58)	-22.35 (0.69)	-1.38 (0.73)
20% reduction in the water quality level	-7.19 (1.86)	-7.29 (1.92)	-0.10 (2.79)

<sup>a</sup> The values in parentheses are standard errors calculated using the procedure reported by Krinsky and Robb (1986) and based on one hundred iteration.

FIGURE 1. APPROXIMATE OPTIMAL CONDITION

