KYOTO UNIVERSITY

NRE Discussion Papers No. 2023-01
October 2023
Division of Natural Resource Economics, Graduate School of Agriculture, Kyoto University

The integer programing extreme value (IPEV) model
An application for estimation of the leisure trip demand

Koichi Kuriyama*, Yasushi Shoji, Takahiro Tsuge

October 20, 2023

Faculty/Graduate School of AGRICULTURE KYOTO UNIVERSITY

Division of Natural Resource Economics
Kitashirakawa Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan

The integer programing extreme value (IPEV) model:

An application for estimation of the leisure trip demand

Koichi Kuriyama*
Graduate School of Agriculture, Kyoto University, Japan
Division of Natural Resource Economics,
Oiwake-cho, Kitashirakawa, Sakyo-ku, Kyoto 606-8502, JAPAN
Phone: +81 75 753 6192, Fax: +81 75 753 6191
E-mail: kuriyama.koichi.8w@kyoto-u.ac.jp

Yasushi Shoji
Research Faculty of Agriculture, Hokkaido University, Japan
Kita 9 Nishi 9, Kita-Ku, Sapporo, 060-8589, Japan

Takahiro Tsuge
Graduate School of Global Environmental Studies,
Sophia University, Kioicho 7-1, Chiyoda-ku, Tokyo 102-8554, Japan

* corresponding author

Title:

The integer programing extreme value (IPEV) model: An application for estimation of the leisure trip demand

Abstract:

We developed an integer programing extreme value (IPEV) model that accounts the integer property of trip data and has the same advantages as the multiple discrete–continuous extreme value choice (MDCEV) model. The proposed model is consistent with utility theory and provides a single structural framework for simultaneously modeling the choice of alternatives and quantity decisions with the constraint of the integer value of consumption. We demonstrate that the proposed model has a closed-form probability expression. Finally, we apply the proposed model to the recreation demand for national parks in Japan. The empirical results suggest that the proposed model provides a better fit for the data than the previous model and that ignoring the integer property of demand might cause an underestimation of the welfare loss.

Keywords:

demand analysis, utility theory, integer demand, MDCEV, leisure trip demand

Acknowledgments:

The authors gratefully acknowledge financial support from Japan Society for the Promotion of Science under KAKENHI Grant Number 22H03812 and the Environment Research and Technology Development Fund (JPMEERF20224003) of the Environmental Restoration and Conservation Agency provided by the Ministry of Environment of Japan.

1. Introduction

Standard models of consumer demand assume continuity of consumption and an interior solution to the consumer's utility maximization problem. However, these assumptions are unrealistic in many settings. For example, leisure trip demand exhibits the characteristics of count data; that is, the number of trips is always a non-negative integer value. Moreover, for a given individual, some sites are visited once or more (interior solutions), while others are not visited (corner solutions). To analyze such data on the integer demand of multiple alternatives, econometric models that account for a mixture of corner and interior solutions with the constraint of non-negative integer demand are required. Similar corner solutions with integer demand emerge in transportation mode (e.g., whether to use a car or bus and how many times to ride), vehicle demand (e.g., which type of and how many vehicles to buy), food demand (e.g., which brand of cereal to buy and how many boxes), and labor demand (e.g., whether to employ part-time or full-time workers and how many workers to employ).

Under the assumption of continuous demand, the multiple discrete–continuous extreme value choice (MDCEV) model (Bhat, 2005, 2008) was proposed to analyze the demand for multiple alternatives with the corner solution. The MDCEV model provides a single structural framework that simultaneously models the consumer choice situation and purchase quantity decisions for multiple alternatives, with the possibility of a corner solution. Furthermore, the MDCEV model has a closed-form probability of observing demand; it is a general model that includes the logit model as a special case. However, the MDCEV model uses the Karush-Kuhn-Tucker (KKT) condition to estimate the parameters. Therefore, the MDCEV model requires the assumption of continuous demand, which may be unrealistic for leisure trips or vehicle choices.

In this study, we propose an integer programing extreme value (IPEV) model, which is an integer demand version of the MDCEV model. The proposed IPEV model, which applies the integer programing method to the MDCEV model, accounts for non-negative integer demand with multiple alternatives. Our proposed model is appealing in that it satisfies the constraint imposed by the non-negative integer amount of consumption, in addition to having the same advantages as the MDCEV continuous demand approach. The integer programing problem is known to be nondeterministic polynomial-time hard (NP-hard), and finding an exact solution to the problem might be difficult (Wolsey, 1998; Wolsey and Nemhauser, 1999; Vazirani, 2001). In the integer and combinatorial optimization literature, approximate algorithms, including local search and greedy methods, have been proposed (Sait and Youssef, 2000; Aarts and Lenstra, 2003). This study investigates the implementation of approximate algorithms to the proposed IPEV model for the estimation of parameters, demand prediction, and welfare analysis. In the empirical section, we consider the application of the IPEV model to the leisure trip demand of national parks in Japan and compare the results of our proposed model with those of the continuous demand MDCEV model.

The contribution of this study can be summarized as follows:

- This study develops an econometric IPEV model that does not require the KKT condition used by the MDCEV model to estimate the parameters. Instead, the proposed IPEV model uses a local search algorithm to determine the optimization condition with the constraint of a non-negative integer demand.
- We show that the proposed IPEV model has the same advantages as the MDCEV model. The IPEV model can analyze the demand for multiple alternatives with corner and interior solutions; it has a closed-form probability of observing demand, and it is a general model that includes the conditional logit model as a special case.
- This study proposes an integer programing approach for demand prediction and welfare analysis of integer demand.
- The empirical results suggest that our proposed IPEV model provides a better fit for leisure trip data than the MDCEV model. Furthermore, we show that ignoring the integer property of demand may cause an underestimation of welfare loss.

The remainder of this paper is organized as follows. Section 2 reviews the literature. Section 3 provides an overview of both the continuous demand model and the proposed integer programing model. Section 4 presents the proposed strategy for demand prediction and welfare analysis using integer-programing algorithms. Section 5 describes the simulation experiment using the proposed model. Section 6 presents the results of our empirical analysis of leisure trip data from the national parks in Japan. Finally, Section 7 provides the concluding remarks.

2. Related works

Several approaches have been proposed for demand analysis of multiple goods using corner solutions. The first approach is the multivariate count data models (Zhang et al., 2017), which are the multivariate version of the count data models of a single good, including Poisson and negative binomial regression models (Cameron and Trivedi, 2013). Multivariate count data models have the advantage of a closed-form probability expression for estimating integer demand for multiple goods. However, these multivariate count models are not based on underlying utility theory. Therefore, these models may not be suitable for economic welfare analyses, although a theoretical foundation for count data models has been developed (Hellerstein and Mendelsohn, 1993).

The second approach is using discrete choice models that describe choice behavior from alternatives. If it were possible to model all possible combinations of good selection and quantity decisions, we could have estimated the utility function using a discrete choice model with the entire set of combinations as the choice set (Jara-Díaz,, 2007). However, due to the vast number of combinations for both site choice and visit frequency, estimation using discrete choice models is practically unfeasible.

The third approach is linked model, which segments the consumer's decision into two components: the choice of alternative and quantity decisions (Bockstael et al., 1986). This model combines the estimation results of the discrete choice and count regression models. Therefore, the linked model can handle the properties of an integer value of consumption. However, this model is not derived from the unified utility theory.[1] Utility-consistent approaches to the linked model are also proposed. For example, Bhat et al. (2014) proposes combing the multivariate count data and multiple discrete–continuous models, while Bhat et al. (2015) considers a combination between the count data and discrete choice models. These models are utility-consistent econometric approaches that can model the property of discrete and non-negative demand. However, these utility-theoretic linked models are not a single structural framework; they require the two components of site choice and visit frequency.

The fourth approach is the Kuhn-Tucker (KT) model,[2] which uses the KKT condition for utility maximization to estimate the parameters of the utility function (von Haefen and Phaneuf, 2005). The KT model is based on utility maximization behavior. The KT model has the advantage of providing a single structural framework to simultaneously model alternative selection and quantity decisions using a mixture of interior and corner solutions. However, the KT model does not have a closed-form probability expression. It is necessary for the KT model to numerically calculate the likelihood. Furthermore, the KT model assumes continuous demand and is inconsistent with discrete choice models. To address the integer demand for the KT model, an integer-programing approach was proposed (Kuriyama and Hanemann, 2006; Lee and Allenby, 2014). Nevertheless, as Bhat (2022) shows, the integer programing approach to the KT model has no closed-form probability expressions. Furthermore, determining the relationship between integer-programing KT and discrete choice models is difficult.

The fifth approach is the MDCEV model (Bhat, 2005, 2008) that also uses the KKT condition to estimate the utility function parameters. Therefore, the MDCEV is consistent with the utility theory, similar to the KT model. The difference between the KT and MDCEV models lies in the assumption that the outside good is deterministic or random (Bhat, 2008). For the KT model, the utility of the outside good (the Hicks composite good) is assumed to be deterministic. For the MDCEV model, the utility of all goods, including outside good, is assumed to have error components with an extreme-value distribution. The MDCEV model has the advantage of having a closed-form probability expression for the observed demand. Furthermore, as Bhat (2008) shows, the MDCEV is a general model that includes a conditional logit model and a KT model for special cases. However, the MDCEV model assumes continuous demand. Therefore, the estimation results of the integer demand using the MDCEV model may be biased. Recently, Bhat

[1] Hausman et al. (1995) claimed that the linked model is utility-consistent for the utility function of Gorman generalized polar form. However, Smith (1997) showed that their proof of consistency relies upon an assumption that holds only under restrictive circumstances. See also Herriges et al. (1999).

[2] The KT demand model was proposed by Hanemann (1978) and Wales and Woodland (1983). Phaneuf et al. (2000) provided an empirical application of this model to estimate recreation demand.

(2022) proposed the multiple discrete-count extreme value (MDCNTEV) model, which is a utility-theoretic model for multiple discrete-count data that links MDCEV and count models. The MDCNTEV model can deal with the property of integer demand of the total consumption. Nevertheless, it has the MDCEV component that requires the assumption of a continuous fractional split in the total demand.

In contrast, this study proposes the IPEV model, which is an integer-programing approach to the MDCEV model. It has the same advantages as the MDCEV model. This is a utility-theoretic econometric model for multiple alternatives with a mixture of interior and corner solutions. It has a closed-form expression for the probability of observing consumption. It is a general model that includes the conditional logit model for special cases. Moreover, the proposed IPEV model is consistent with non-negative integer demand. Although discrete choice models could also be consistent with non-negative integer demand, utilizing them requires all possible combinations of expenditures as a choice set. In contrast, the proposed IPEV model enables estimating utility functions by comparing only in the neighborhood of observed consumption. The difference between discrete choice models and the proposed model therefore lies in whether all feasible combinations are used, or only the neighborhood is considered. The utility-theoretic linked models (Bhat et al., 2014; Bhat et al., 2015) can also estimate utility function using non-negative count data. However, the advantage of the proposed model over utility-theoretic linked models—that require multiple components of site choice and visit frequency—is that the proposed model provides a single structural framework for non-negative integer demand, without multiple components.

3. Model

3.1 The continuous demand MDCEV model

This section summarizes the continuous demand MDCEV model proposed by Bhat (2008). Consider a generalized variant of the translated CES utility function with an outside good:

$$U(x) = \sum_{k=1}^{K} U_k(x_k) = \frac{\psi_1}{\alpha_1} x_1^{\alpha_1} + \sum_{k=2}^{K} \frac{\gamma_k \psi_k}{\alpha_k} \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\},$$
$$\psi_1 = \exp(\varepsilon_1),$$
$$\psi_k = \exp(\beta' z_k + \varepsilon_k), \tag{1}$$

where $U(x)$ is a quasi-concave, increasing, and continuously differentiable function with respect to the consumption goods $x = (x_1, \dots, x_K)$; x_1 is the numeraire or an outside good that is always consumed in positive quantities; z_k is a vector of attributes of good k and the individual attributes; ε_k is a vector of random components that are assumed to be known to the individual but unknown to the analyst; and α_k,

β, and γ_k are the estimated parameters. The individual is assumed to maximize the utility function, subject to the budget constraint. The Lagrangian of this utility maximization problem is as follows:

$$\mathcal{L} = \frac{\psi_1}{\alpha_1} x_1^{\alpha_1} + \sum_{k=2}^{K} \frac{\gamma_k \psi_k}{\alpha_k} \left\{ \left(\frac{x_k}{\gamma_k} + 1 \right)^{\alpha_k} - 1 \right\} + \lambda \left(E - \sum_{k=1}^{K} p_k x_k \right), \tag{2}$$

where E is income, p_k is the price of good k, and λ is the Lagrangian multiplier. The KKT condition for this problem is given as

$$\begin{aligned} &\psi_k \left(\frac{x_k^*}{\gamma_k} + 1 \right)^{\alpha_k - 1} = \lambda p_k, \text{ if } x_k^* > 0, \\ &\psi_k \left(\frac{x_k^*}{\gamma_k} + 1 \right)^{\alpha_k - 1} < \lambda p_k, \text{ if } x_k^* = 0, \end{aligned} \tag{3}$$

where x_k^* is the solution to the optimization problem. The first good is assumed to be the outside good $x_1^* > 0$; therefore, the KKT condition can be rewritten as

$$\begin{aligned} &V_k + \varepsilon_k = V_1 + \varepsilon_1, \text{ if } x_k^* > 0, \\ &V_k + \varepsilon_k < V_1 + \varepsilon_1, \text{ if } x_k^* = 0, \end{aligned} \tag{4}$$

where $V_1 = (\alpha_1 - 1)\ln x_1 - \ln p_1$ and $V_k = \beta' z_k + (\alpha_k - 1)\ln\left(\frac{x_k^*}{\gamma_k} + 1\right) - \ln p_k$. Consider an individual who chooses to consume only the first M of K goods; thus, $x_k^* > 0$ for $k = 1, \dots, M$ and $x_k^* = 0$ for $k = M + 1, \dots, K$. If the ε_k's are independent and identically distributed draws from the type I extreme value distribution with inverse scale parameter σ for all k, Bhat (2008) shows that the probability of observing $x^* = (x_1^*, \dots, x_K^*)$ has closed form:

$$P(x^*) = \frac{1}{p_1} \frac{1}{\sigma^{M-1}} \left[\prod_{i=1}^{M} f_i \right] \left[\sum_{i=1}^{M} \frac{P_i}{f_i} \right] \left[\frac{\prod_{i=1}^{M} e^{V_i/\sigma}}{(\sum_{k=1}^{K} e^{V_k/\sigma})^M} \right] (M-1)!, \tag{5}$$

where $f_i = \left(\frac{1-\alpha_i}{x_i^* + \gamma_i} \right)$. Note that this probability can collapse to the standard conditional logit model, when only one alternative is chosen ($M = 1$), and there are no satiation effects ($\alpha_k = 1$ for all k), as shown in Bhat's study (2008).

3.2 The integer programing extreme value (IPEV) model

If the assumption of continuous consumption is not satisfied, then the consumer's problem cannot be solved using the KKT condition. For example, this assumption is violated by the recreation demand. Recreational trips can only be numbered as non-negative integers; thus, the consumer's problem should be written as a non-linear integer programing problem. In this case, consumer's utility maximization is as follows:

$$Max\ U(x) \quad s.t. \sum_{k=1}^{K} p_k x_k \leq E, x_1 > 0, \text{ and } x_k \in \mathbb{Z}_+ \text{ for all } k \geq 2, \tag{6}$$

where $\mathbb{Z}_+$ is the set of non-negative integers.

An integer programing problem is generally known to be NP-hard, and it may be difficult to find its exact solution. Approximation algorithms, such as the local search, greedy method, genetic algorithm, simulated annealing, and tabu search, have been proposed as alternatives to the above approach. We focus on the local search algorithm, which iteratively attempts to identify the local optima within a given neighborhood (Aarts and Lenstra, 2003). This algorithm proceeds as follows. Let x^1 be the starting value of consumption. The neighborhood of x^1 with radius δ is defined as $N^1(x^1, \delta) = \{s \in \mathbb{Z}_+ | d(x^1, s) \leq \delta\}$, where $d(x^1, s)$ is the Euclidean distance between x^1 and s. If there exists another consumption $x^2 \in N^1(x^1)$ that yields $U(x^2) > U(x^1)$, then consider the updated neighborhood $N^2(x^2, \delta) = \{s \in \mathbb{Z}_+ | d(x^2, s) \leq \delta\}$, and iterate this process. If there is no allocation that satisfies $x^{n+1} \in N^n(x^n, \delta)$ and $U(x^{n+1}) > U(x^n)$, then define the consumption x^n as the approximate solution of the integer programing problem.

Assume that the first good x_1 is the Hicksian composite good with $p_1 = 1$ and $x_1 > p_k$ for all k. Let $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_K)$ be the solution of (6) using the local search algorithm with radius 1. Then, the approximate local optimal conditions are as follows:

$$\begin{aligned} &U(\tilde{x}_k^+) \leq U(\tilde{x}) \text{ and } U(\tilde{x}_k^-) \leq U(\tilde{x}) \text{ if } \tilde{x}_k > 0, \\ &U(\tilde{x}_k^+) \leq U(\tilde{x}) \ \text{ if } \tilde{x}_k = 0, \end{aligned} \tag{7}$$

where $\tilde{x}_k^+ = (\tilde{x}_1 - p_k, \tilde{x}_2, \dots, \tilde{x}_{k-1}, \tilde{x}_k + 1, \tilde{x}_{k+1}, \dots, \tilde{x}_K)$ and $\tilde{x}_k^- = (\tilde{x}_1 + p_k, \tilde{x}_2, \dots, \tilde{x}_{k-1}, \tilde{x}_k - 1, \tilde{x}_{k+1}, \dots, \tilde{x}_K)$. As shown in Appendix A, for the additively separable utility function of (1), this condition can be rewritten as

$$\begin{aligned} &V_{1,k}^- - V_k^- + \varepsilon_1 \leq \varepsilon_k \leq V_{1,k}^+ - V_k^+ + \varepsilon_1 \ \text{ if } \tilde{x}_k > 0, \\ &V_k^+ + \varepsilon_k \leq V_{1,k}^+ + \varepsilon_1 \ \text{ if } \tilde{x}_k = 0, \end{aligned} \tag{8}$$

where

$$V_{1,k}^+ + \varepsilon_1 = \ln[U_1(\tilde{x}_1) - U_1(\tilde{x}_1 - p_k)] = -\ln \alpha_1 + \ln\left[\tilde{x}_1^{\alpha_1} - (\tilde{x}_1 - p_k)^{\alpha_1}\right] + \varepsilon_1,$$

$$V_{1,k}^{-} + \varepsilon_1 = \ln[U_1(\tilde{x}_1 + p_k) - U_1(\tilde{x}_1)] = -\ln\alpha_1 + \ln\left[(\tilde{x}_1 + p_k)^{\alpha_1} - \tilde{x}_1^{\alpha_1}\right] + \varepsilon_1$$

$$\begin{aligned} V_k^{+} + \varepsilon_k &= \ln[U_k(\tilde{x}_k + 1) - U_k\,(\tilde{x}_k)] \\ &= \beta' z_k + \ln\gamma_k - \ln\alpha_k + \ln\left[\left(\frac{\tilde{x}_k + 1}{\gamma_k} + 1\right)^{\alpha_k} - \left(\frac{\tilde{x}_k}{\gamma_k} + 1\right)^{\alpha_k}\right] + \varepsilon_k, \text{ and} \end{aligned}$$

$$\begin{aligned} V_k^{-} + \varepsilon_k &= \ln[U_k(\tilde{x}_k) - U_k\,(\tilde{x}_k - 1)] \\ &= \beta' z_k + \ln\gamma_k - \ln\alpha_k + \ln\left[\left(\frac{\tilde{x}_k}{\gamma_k} + 1\right)^{\alpha_k} - \left(\frac{\tilde{x}_k - 1}{\gamma_k} + 1\right)^{\alpha_k}\right] + \varepsilon_k. \end{aligned}$$

Specifically, V_k^{+} and V_k^{-} are the deterministic components in the log transformations of the utility difference obtained by increasing or decreasing one unit in k good, respectively. $V_{1,k}^{+}$ and $V_{1,k}^{-}$ are those of the first outside good obtained by increasing or decreasing one unit of k good, respectively. Note that the local optimal condition (8) of the integer programing model is similar to the KKT condition (4) of the MDCEV model.

Consider an individual who consumes only the first M of K goods. Assume that ε_k is the type I extreme value distribution with the inverse scale parameter σ for all k. The probability of observing $\tilde{x}^* = (\tilde{x}_1^*, \tilde{x}_2^*, \ldots, \tilde{x}_M^*, 0, \ldots, 0)$ can be calculated as

$$\begin{aligned} &P(\tilde{x}_1^*, \tilde{x}_2^*, \ldots, \tilde{x}_M^*, 0, \ldots, 0) \\ &= \int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty} \left\{ \prod_{k=2}^{M} \left[\Pr\left(V_{1,k}^{-} - V_k^{-} + \varepsilon_1 \le \varepsilon_k \le V_{1,k}^{+} - V_k^{+} + \varepsilon_1\right)\right] \prod_{k=M+1}^{K} \left[\Pr\left(\varepsilon_k \right.\right.\right. \\ &\qquad\qquad \left.\left.\left. \le V_{1,k}^{+} - V_k^{+} + \varepsilon_1\right)\right]\right\} \frac{1}{\sigma} g\left(\frac{\varepsilon_1}{\sigma}\right) d\varepsilon_1 \\ &= \int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty} \left\{ \prod_{k=2}^{M} \left[G\left(\frac{V_{1,k}^{+} - V_k^{+} + \varepsilon_1}{\sigma}\right) \right.\right. \\ &\qquad\qquad \left.\left. - G\left(\frac{V_{1,k}^{-} - V_k^{-} + \varepsilon_1}{\sigma}\right)\right] \prod_{k=M+1}^{K} G\left(\frac{V_{1,k}^{+} - V_k^{+} + \varepsilon_1}{\sigma}\right)\right\} \frac{1}{\sigma} g\left(\frac{\varepsilon_1}{\sigma}\right) d\varepsilon_1, \end{aligned} \tag{9}$$

where $G(t) = \exp(-\exp(-t))$ and $g(t) = \exp(-t) \cdot \exp(-\exp(-t))$ are the distribution and density function of the type I extreme value, respectively.

As Appendix B shows, the probability in Equation (9) has a closed-form expression:

$$P(\tilde{x}_1^*, \tilde{x}_2^*, \dots, \tilde{x}_M^*, 0, \dots, 0) = \sum_{s_1=1}^{1} \sum_{s_2=0}^{1} \cdots \sum_{s_M=0}^{1} \frac{\prod_{k=1}^{M} (-1)^{1-s_k}}{\sum_{k=1}^{K} e^{-\frac{d_k^*}{\sigma}}}, \tag{10}$$

where d_k^* is the adjusted utility difference between good 1 and k such that $d_1^* = 0$ and $d_k^* = 1[k \leq M] \cdot \left[s_k \cdot \left(V_{1,k}^+ - V_k^+\right) + (1 - s_k) \cdot \left(V_{1,k}^- - V_k^-\right)\right] + 1[k > M] \cdot (V_{1,k}^+ - V_k^+)$ for all $k \geq 2$, where $1[k \leq M]$ is an indicator function equal to one if $k \leq M$ and zero otherwise. This probability can be rewritten as

$$P(\tilde{x}_1^*, \tilde{x}_2^*, \dots, \tilde{x}_M^*, 0, \dots, 0) = \sum_{s_1=1}^{1} \sum_{s_2=0}^{1} \cdots \sum_{s_M=0}^{1} \left[\prod_{k=1}^{M} (-1)^{1-s_k} \frac{e^{\frac{v_1}{\sigma}}}{\sum_{k=1}^{K} e^{\frac{v_k^*}{\sigma}}}\right], \tag{11}$$

where $v_1 = \beta' z_1 + \ln \gamma_1 - \ln \alpha_1 - \ln p_1$ and $v_k^* = v_1 - d_k^*$.

This closed-form probability has 2^{M-1} terms. Therefore, for a large choice set, calculating probability using this equation can be difficult. Alternatively, the parameters can be estimated using a simulated probability (Train, 2009). Let ε_1^r be the *r*th draw from the type I extreme value with the inverse scale σ. The simulated probability of observing $\tilde{x}$ is

$$\begin{aligned} &\check{P}(\tilde{x}_1^*, \tilde{x}_2^*, \dots, \tilde{x}_M^*, 0, \dots, 0) \\ &\quad = \frac{1}{R} \sum_{r=1}^{R} \left\{ \prod_{k=2}^{M} \left[G\left(\frac{V_{1,k}^+ - V_k^+ + \varepsilon_1^r}{\sigma}\right) \right.\right. \\ &\quad \left.\left. - G\left(\frac{V_{1,k}^- - V_k^- + \varepsilon_1^r}{\sigma}\right) \right] \prod_{k=M+1}^{K} G\left(\frac{V_{1,k}^+ - V_k^+ + \varepsilon_1^r}{\sigma}\right) \right\}, \end{aligned} \tag{12}$$

where R is the number of draws.

3.3 Mixed IPEV model

The IPEV model assumes that the error component of ε is the iid extreme value distribution. The mixed IPEV model, which introduces a mixed distribution to the IPEV model, can incorporate more general error structures. Assume that the log transformations of the utility differences have the additional error components η: $V_k^+(\eta) + \varepsilon_k$, $V_k^-(\eta) + \varepsilon_k$, $V_{1,k}^+(\eta) + \varepsilon_1$, and $V_{1,k}^-(\eta) + \varepsilon_1$. Let $f(\eta)$ be the density of η. The probability of observing $\tilde{x}^*$ in the mixed IPEV model is as follows:

$$P(\tilde{x}_1^*, \tilde{x}_2^*, \dots, \tilde{x}_M^*, 0, \dots, 0) = \int_\eta \sum_{s_1=1}^{1} \sum_{s_2=0}^{1} \cdots \sum_{s_M=0}^{1} \frac{\prod_{k=1}^{M} (-1)^{1-s_k}}{\sum_{k=1}^{K} e^{-\frac{d_k^*(\eta)}{\sigma}}} f(\eta) d\eta, \text{ where}$$

$$d_k^*(\eta) = 1[k \le M] \left[s_k \left(V_{1,k}^+(\eta) - V_k^+(\eta) \right) + (1 - s_k) \left(V_{1,k}^-(\eta) - V_k^-(\eta) \right) \right] \tag{13}$$

$$+ 1[k > M] \left(V_{1,k}^+(\eta) - V_k^+(\eta) \right).$$

This probability does not have closed-form. However, it can be approximated through simulations. Let η^r be the *r*th draw from $f(\eta)$. The simulated probability is as follows:

$$\breve{P}(\tilde{x}_1^*, \tilde{x}_2^*, \dots, \tilde{x}_M^*, 0, \dots, 0) = \frac{1}{R} \sum_{r=1}^{R} \sum_{s_1=1}^{1} \sum_{s_2=0}^{1} \cdots \sum_{s_M=0}^{1} \frac{\prod_{k=1}^{M} (-1)^{1-s_k}}{\sum_{k=1}^{K} e^{-\frac{d_k^*(\eta^r)}{\sigma}}}, \tag{14}$$

where R is the number of draws.

3.4 Comparison with discrete choice model

Consider the special case of $M = 1$ (i.e., only one alternative is chosen). Namely, $\tilde{x}_1^* > 0$ and $\tilde{x}_k^* = 0$ for all $k \ge 2$. In this case, the probability of Equation (11) collapses to the standard logit for the discrete choice model:

$$P(\tilde{x}_1^*, 0, \dots, 0) = \frac{e^{\frac{v_1}{\sigma}}}{\sum_{k=1}^{K} e^{\frac{v_k^*}{\sigma}}}, \tag{15}$$

where

$$v_1 = \beta' z_1 + \ln \gamma_1 - \ln \alpha_1 - \ln p_1,$$
$$v_k^* = \beta' z_k + \ln \gamma_k - \ln \alpha_k + \ln \left[\left(\frac{\tilde{x}_k + 1}{\gamma_k} + 1 \right)^{\alpha_k} - \left(\frac{\tilde{x}_k}{\gamma_k} + 1 \right)^{\alpha_k} \right] - \ln[\tilde{x}_1{}^{\alpha_1} - (\tilde{x}_1 - p_k)^{\alpha_1}]. \tag{16}$$

When $\alpha_k = \gamma_k = 1$, this probability can be reduced to a conditional logit model:

$$P(\tilde{x}_1^*, 0, \dots, 0) = \frac{e^{b' z_1 - c \cdot \ln p_1}}{\sum_{k=1}^{K} e^{b' z_k - c \cdot \ln p_k}}, \tag{17}$$

where $b = \beta/\sigma$ and $c = 1/\sigma$.

4. Demand prediction and welfare analysis with integer demand

4.1 The continuous demand model

Demand prediction requires solving the utility-maximization problem. Let $x^*(p, z, E, \varepsilon)$ be the demand function that is the solution to the problem and $V(p, z, E, \varepsilon) = U\big(x^*(p, z, E, \varepsilon)\big)$ be the indirect utility function. For continuous demand, Pinjari and Bhat (2021) show that there exists an analytical solution to $x^*(p, z, E, \varepsilon)$, when all the satiation parameters of the utility function (1) are equal ($\alpha_k = \alpha$ for all k), and if the chosen and non-chosen alternatives are known. However, in general, no closed-form solution to $x^*(p, z, E, \varepsilon)$ exists; therefore, the demand prediction requires numerical methods. Pinjari and Bhat (2021) propose demand forecasting method using the bisection algorithm over λ (the Lagrange multiplier) with the KKT condition.

For welfare analysis, the compensating variation (CV) associated with the change from (p^0, z^0) to (p^1, z^1) is the solution to

$$V(p^0, z^0, E, \varepsilon) = V(p^1, z^1, E - CV, \varepsilon). \tag{18}$$

von Haefen et al. (2004) develop a nesting bisection algorithm to solve the CV using the KKT condition. von Haefen (2007) proposes a bisection algorithm to solve the expenditure minimization problem. Lloyd-Smit (2018) demonstrated a more efficient method for solving the CV.

4.2 The integer demand model

For integer demand, Pinjari and Bhat's (2021) strategy for demand prediction cannot be applied because their algorithm uses the KKT first-order condition, which is not satisfied under the constraint of integer demand. Alternatively, we propose using the greedy method to find a solution to the utility-maximization problem. The greedy method was developed using integer programing literature (Aarts and Lenstra, 2003). This algorithm makes a locally optimal choice at each iteration. The greedy method is an approximation algorithm that does not always yield the optimal solutions. However, this approach performs well in many maximization problems in which the target function is globally concave.

Let $\tilde{x}(p, z, E, \varepsilon)$ be the integer demand function that is the solution to the integer programing problem. The greedy strategy for finding the solution $\tilde{x}$ conditional on the error components ε is as follows.

Algorithm 1: the greedy method of demand solution

1. Set iteration $t = 1$.
2. Let $\tilde{x}^t$ be the demand at iteration t. Set the initial values $\tilde{x}^1 = (E, 0, \dots, 0)$.

3. Construct the candidate $c_k^t = (\tilde{x}_1^t - p_k, \tilde{x}_2^t, \dots, \tilde{x}_{k-1}^t, \tilde{x}_k^t + 1, \tilde{x}_{k+1}^t, \dots, \tilde{x}_K^t)$ for all $k \geq 2$ and $c_1^t = \tilde{x}^t$.
4. If no locally optimal consumption c_k^t for $k \geq 2$, exists, such that $U(c_k^t|\varepsilon) > U\left(c_j^t|\varepsilon\right)$ for all $j \geq 2$, stop the iteration and keep $\tilde{x}^t$.
5. Otherwise, update the demand $\tilde{x}^{t+1} = c_k^t$. Set $t = t + 1$, and repeat steps 2–5.

This procedure will yield the approximate optimal consumption $\tilde{x}$ conditional on ε under the constraints of non-negative and integer consumption. x_1 is a continuous variable; therefore, a combination of the greedy method and numerical bisection algorithm can be used to solve the CV associated with the change from (p^0, z^0) to (p^1, z^1), which is defined by $\tilde{V}(p^0, z^0, E, \varepsilon) = \tilde{V}(p^1, z^1, E - CV, \varepsilon)$. The search algorithm for a CV with an integer demand is as follows.

Algorithm 2: the bisection method for searching CV

1. Calculate the baseline utility, $\tilde{V}^0 = \tilde{V}(p^0, z^0, E, \varepsilon)$.
2. At iteration t, set $CV_a^t = (CV_l^{t-1} + CV_u^{t-1})/2$. To initialize the algorithm, set $CV_l^0 = -E$ and $CV_u^0 = E$.
3. Conditional on $(p^1, z^1, E - CV_a^t, \varepsilon)$, solve for $\tilde{x}^t$ using the greedy method of algorithm 1.
4. Construct the utility $\tilde{V}^t = \tilde{V}(p^1, z^1, E - CV_a^t, \varepsilon)$. If $\tilde{V}^t > \tilde{V}^0$, set $CV_l^t = CV_a^t$ and $CV_u^t = CV_u^{t-1}$. Otherwise, set $CV_l^t = CV_l^{t-1}$ and $CV_u^t = CV_a^t$.
5. Iterate until convergence.

The predicted demand $\tilde{x}(\varepsilon)$ and $CV(\varepsilon)$ have the error components ε. The analyst can estimate the expectation of demand $E[\tilde{x}]$ and the compensating variation $E[CV]$ using a simulation. First, ε_1 of the first good is drawn from the type I extreme value distribution with the inverse scale σ. Let ε_1^t be the *t*th draw. Second, $(\varepsilon_2, \dots, \varepsilon_K)$ of other goods are drawn using the procedure proposed by von Haefen et al. (2004). From the approximated local optimal condition of (8), if the observed demand is the corner solution ($\tilde{x}_k^* = 0$), the error component ε_k has the truncated distribution of $\varepsilon_k^t < V_{1,k}^+ - V_k^+ + \varepsilon_1^t$. Otherwise ($\tilde{x}_k^* > 0$), it has the truncated distribution of $V_{1,k}^- - V_k^- + \varepsilon_1^t \leq \varepsilon_k^t \leq V_{1,k}^+ - V_k^+ + \varepsilon_1^t$. Therefore, ε_k^t for all $k \geq 2$ can be simulated by the following truncated type I extreme value distribution:

$$\begin{aligned}
\varepsilon_k^t &= -\ln\left\{-\ln\left[G\left(\frac{V_{1,k}^+ - V_k^+ + \varepsilon_1^t}{\sigma}\right)\mu^t\right]\right\}\sigma, \text{for } \tilde{x}_k^* = 0, \\
\varepsilon_k^t &= -\ln\left\{-\ln\left[(1-\mu^t)G\left(\frac{V_{1,k}^+ - V_k^+ + \varepsilon_1^t}{\sigma}\right) + \mu^t G\left(\frac{V_{1,k}^- - V_k^- + \varepsilon_1^t}{\sigma}\right)\right]\right\}\sigma, \text{for } \tilde{x}_k^* > 0,
\end{aligned} \tag{19}$$

where μ^t is the tth draw from the standard uniform distribution. The simulated demand $\tilde{x}(\varepsilon^t)$ and CV $CV(\varepsilon^t)$ can be calculated using the generated draws ε^t. Iterating this process and averaging the results provide $E[\tilde{x}]$ and $E[CV]$.

For the mixed IPEV model, demand prediction and welfare analysis require the error components of ε and η. The joint distribution of the error components $f(\varepsilon, \eta|\tilde{x}^*)$ given observed demand $\tilde{x}^*$ can be decomposed as $f(\varepsilon, \eta|\tilde{x}^*) = f_1(\eta|\tilde{x}^*) \cdot f_2(\varepsilon|\eta, \tilde{x}^*)$. von Haefen et al. (2004) proposes an adaptive Metropolis-Hastings algorithm to simulate from $f(\varepsilon, \eta|\tilde{x}^*)$. After drawing η, the random draws of ε can be simulated using Equation (19).

5. Simulation Experiment

5.1 Experimental setting

To confirm the performance of the proposed IPEV model, we provide a Monte Carlo experiment using the IPEV model and hypothetical generated data. The procedure for this experiment is as follows:

1. Define the utility function. We use γ-profile utility function: $\alpha_k \to 0$ for all $k \geq 2$. The number of goods is four ($K = 4$). The baseline utility in ψ_k has three individual attributes (z). Table 1 provides the assumed value of parameters.
2. Generate individual attributes and price data. The individual attributes are generated using random draws from $z \in \{1,2,3,4,5\}$. The sample size is one thousand individuals. The income is generated using random uniform distribution with an interval of $E \in [\$30{,}000 - \$100{,}000]$. The price is generated using random draws from uniform distribution of $p_k \in [\$100 - \$2{,}000]$.
3. Generate the simulated trip data. Generate random draws ε_k from a Gumbel distribution. The simulated trips are generated using greedy method given ε_k (Algorithm 1).
4. Estimate the parameters. IPEV models with the closed-form (IPEV1) or simulated probability (IPEV2) are applied to estimate the parameters using simulated trips. For the IPEV2 model, the parameters are estimated using the maximum simulated log-likelihood with 200 Halton draws.
5. Repeat steps 1–4 up to the two hundredth iteration.

Table 1. Estimation results of the simulation experiments

Table 1 shows the estimation results of the simulation experiments. For the IPEV1 (closed-form), all estimated parameters are significant. The errors between estimated parameters and assumed values are non-significant. The estimation results of the IPEV2 (simulation) are similar to the ones of IPEV1. Thus, the experimental results suggest that the proposed IPEV model exhibits consistency in both the maximum likelihood estimation—using the closed-form probability—and the maximum simulated likelihood method, using simulation probabilities.

6. Empirical illustration

6.1 Example of small choice set

To illustrate the proposed IPEV model, we use data on recreational trips to the national parks in Japan. The first empirical study analyzes the trip data to the national parks in Hokkaido, a northern island of Japan. Hokkaido has six national parks, including Shiretoko, a World Heritage Site. To investigate the number of trips to each national park in Hokkaido, we conducted an online survey of Hokkaido residents. The respondents were asked about the number of trips they made to national parks from July to November 2008. The sample consisted of 763 respondents with information on 392 trips. Table 2 presents the trip distributions of the data. Note that while 69% respondents did not visit any national parks in Hokkaido, 31% made one or more trips.

We used six national parks as the choice set. Table 3 provides the descriptive statistics of the data. *Price* is the two-way trip cost between the respondent's residence and each park, which includes expressway toll, vehicle operation cost (US$0.155 per mile), and opportunity cost.[3]

Table 2 Trip distribution of the small choice set

Table 3 Descriptive statistics of the small choice set

As shown in Bhat (2008), identifying both α_k and γ_k parameters is difficult. In our study, following Kuriyama et al. (2020), we use γ-profile utility function:

$$U(x) = \frac{\psi_1}{\alpha_1} x_1^{\alpha_1} + \sum_{k=2}^{K} \gamma_k \psi_k \ln\left(\frac{x_k}{\gamma_k} + 1\right), \tag{20}$$

$$\psi_1 = \exp(\varepsilon_1) \text{ and } \psi_k = \exp(\beta' z_k + \varepsilon_k),$$

where x_1 is the Hicks composite good that is the expense other than the cost of the national park trips ($x_1 = E - \sum_{k=2}^{K} p_k x_k$). The baseline utility component ψ_k includes individual attributes.

Using these data, we compare the MDCEV model assuming continuous demand with the IPEV model assuming integer demand. For the IPEV model, we consider an estimation using a closed-form and simulated probabilities. Six choice sets were used for the data. Therefore, the maximum number of terms in the closed-form probability (10) is $2^6 = 64$, and it is feasible to calculate the closed-form probability for this small choice set data.

[3] The vehicle operation cost is calculated using an average of gasoline prices, gas mileage, and exchange rate at 2008 (US$ 1 = 104 yen). We calculate the opportunity cost using the method proposed by Cesario (1976): The opportunity cost = two-way trip time * wage rate / 3.

Table 4 provides the estimation results using the three models: (1) the MDCEV, (2) the IPEV using the closed-form probability, and (3) the IPEV using the simulated probability. The parameters of the MDCEV and IPEV1 (closed-form) models are estimated using the maximum log likelihood, whereas the parameters of the IPEV2 (simulation) model are estimated using the maximum simulated log likelihood with 200 Halton draws, following Sikder and Pinjari (2013). This table shows the large difference in the parameter values between the MDCEV and IPEV models. However, the estimated results of the IPEV1 model are remarkably close to those of the IPEV2 model. This suggests that the simulated maximum likelihood of the IPEV model performs well in the approximation of the standard maximum likelihood of the IPEV model using the closed-form probability. Furthermore, the log-likelihood of this table suggests that the IPEV models are better fit to the data than the MDCEV model.

Table 4 Estimation result of the small choice set

For demand prediction and welfare analysis, we consider two scenarios: (a) closing Shiretoko National Park, a World Heritage Site, and (b) setting an entrance fee of 5,000 yen (US$ 48) at Shiretoko park. In 2005, the Shiretoko area was designated a UNESCO World Heritage Site. Regarding scenario (a), Shiretoko Five Lakes, the most popular area for visitors, is closed in some seasons because it is a natural habitat of the brown bear, and the park manager needs to prevent accidents due to encounters between visitors and bears. In scenario (b), although Japanese national parks do not currently charge entrance fees, there are ongoing discussions about implementing entrance fees for some of them. Introducing entrance fees could potentially lead to a reduction in the number of visitors and have negative effects on the local economy.

Tables 5 and 6 provide the demand prediction and compensation variations of the two scenarios using the MDCEV and IPEV models with small choice set data. Table 5 shows that there are no significant differences in the demand prediction using the three models. One possible reason for this result is averaging by rounding up and down the integer demand. For example, consider two respondents who consume 1.6 and 2.4 trips in continuous demand. The average number of trips of continuous demand over two people was 2.0. In integer demand, 1.6 trip rounds up and 2.4 trip rounds down. Thus, the average number of trips for integer demand is also 2.0.

Table 5 Demand prediction of the small choice set
Table 6 Compensating variations of the small choice set

In contrast, Table 6 suggests that there are significant differences in the CVs between the MDCEV and IPEV models. The estimated CVs show that the welfare loss of the integer demand models (IPEV1 and IPEV2) is larger than that of the continuous demand model (MDCEV). The utility maximization problem of the integer demand model has an additional constraint on non-negative integer consumption. This constraint may cause greater welfare losses in the IPEV model. For example, consider an individual who visits the closest site only once. If the visited site is close, the individual can visit a distant site. However, the number of trips to distant sites may decrease because of higher costs. For continuous demand, it may be less than one, but a positive number, say 0.4 trips. In contrast, it may be zero for integer demand and the welfare loss of site closure with integer demand may be larger than that with continuous demand. Thus, for integer demand, the MDCEV model that assumes continuous demand may underestimate the welfare loss owing to site closure.

6.2 Example of large choice set

The second empirical study uses trip data to all national parks in Japan. Japan has 34 national parks. In 2019, the national parks had more than 369 million visitors. We conducted a nationwide web survey in February 2020 to analyze the recreation demand for national parks. The respondents were asked about their number of visits to each national park in 2019. The sample consisted of 1591 respondents with information on 3,285 trips. Table 7 presents the trip distribution for our data. Note that while 62% respondents did not visit any national park in 2019, 38% visited national parks one or more times.

Table 7 Trip distribution of the large choice set

Table 8 Descriptive statistics of the large choice set

We consider the 34 national parks as the choice set.[4] Table 8 demonstrates the descriptive statistics for the data. *Price* is the two-way trip cost between the respondent's residence and each park, which includes airfare, railroad fare, expressway toll, vehicle operation cost (US$ 0.219 per mile), and opportunity cost.[5] *Word Heritage* is a dummy variable for the World Heritage Sites. Vehicle access is regulated in some parks due to traffic and damage to the natural environment. *Vehicle regulation* is a dummy variable that prohibits private vehicle access to parks. *Area 1 – Area 7* represent the area dummies. The Japanese Ministry of the Environment classifies national parks into seven categories.

In this empirical analysis, as in the study on small choice set, we use γ-profile utility function. For large choice set data, we assume $\gamma_k = \gamma$ for all $k \geq 2$, because it is difficult to estimate γ_k for all parks, due to the large number of parameters. Instead, we introduce area dummies *Area 1 – Area 7* in the utility baseline. Using these data, we compare the MDCEV model assuming continuous demand with the IPEV

[4] Kuriyama et al. (2020) used 32 national parks as the choice set. In their analysis, two national parks, Yambaru and Amami-Gunto, were excluded from the choice set because they were established after their survey was conducted.

[5] The vehicle operation cost is calculated using an average of gasoline prices, gas mileage, and exchange rate in 2019 (US$ 1 = 110 yen).

model assuming integer demand. For the IPEV model, it may not be feasible to estimate using a closed-form probability because of the large choice set. Thirty-four choice sets were used for the data. The maximum number of the terms in the closed-form probability (10) is $2^{34} \approx 1.7 \times 10^{10}$. Therefore, we use the IPEV model with a simulated probability.

Table 9 presents the estimated parameters of the two models. The parameters of the MDCEV model are estimated using the maximum log-likelihood, whereas those of the IPEV (simulation) model are estimated using the maximum simulated log-likelihood with 200 Halton draws. This table shows the large difference in the parameter values between the MDCEV and IPEV models. The log-likelihood of this table suggests that the IPEV model is a better fit for the data than the MDCEV model.

Table 9 Estimation result of the large choice set

We consider two scenarios for demand prediction and welfare analysis: (a) closing visitor centers and (b) closing Fuji-Hakone-Izu National Park. In scenario (a), on April 7, 2020, the Japanese government declared a state of emergency owing to COVID-19. Following the Japanese government's declaration of emergency, most visitor centers in national parks were closed to avoid infection. In scenario (b), Fuji-Hakone-Izu National Park is the most popular national park in Japan. Mount Fuji was included on the UNESCO World Heritage List in 2013. Mount Fuji has not erupted since 1707, and at present, the Japan Meteorological Agency has assigned Mount Fuji the lowest threat level for volcanic eruptions. However, the closure of the national park owing to the eruption of Mount Fuji might have a large and negative impact on climbers, even if the risk of eruption is extremely low. Mount Hakone is also located in the park; in May 2015, Owakudani, the valley beneath Mount Hakone, was closed because of volcanic activity.

The estimated demand prediction and CVs in the two scenarios with large choice set data using the MDCEV and IPEV models are demonstrated in Tables 10 and 11, respectively. The results of the large choice set data are similar to those of the small choice set data. While Table 10 shows no significant differences in demand prediction between the MDCEV and IPEV models, the welfare loss estimated using the IPEV model is significantly larger than that estimated using the MDCEV model.

Table 10 Demand prediction of the small choice set

Table 11 Compensating variations of the small choice set

7. Discussion

In this study, we propose an IPEV model that applies the integer programing method to the MDCEV model and accounts for non-negative integer demand with multiple alternatives. Our proposed model is appealing as it satisfies the constraint imposed by the non-negative integer amount of consumption, in addition to having the same advantages as the MDCEV continuous demand approach. The IPEV model provides a single structural framework that simultaneously models the consumer choice situation and purchase quantity decision with the constraint of non-negative integer consumption, while allowing for the possibility of a corner solution, in addition to being consistent with the utility theory. Furthermore, the IPEV model has a closed-form expression for the probability of observing consumption. It may not be feasible to use the closed-form probability of the IPEV model for a large choice set. To address the calculation burden, we propose an IPEV model with a simulated probability for large choice set data.

In demand prediction and welfare analysis, finding exact solutions to integer programing problems is generally difficult. To overcome these difficulties, we suggested the use of alternative approximate approaches, including the local search algorithm and greedy method, both of which have been widely used in other integer programing studies. With an empirical application using the data on recreational trips to national parks in Japan with small and large choice sets, we compared our proposed IPEV model with the continuous demand MDCEV model. The empirical results suggested that the integer programing approach generally provides a better fit for our data than the continuous demand approach. The first example, using the small choice set, showed that the estimated results of the IPEV model with closed-form probability were remarkably close to those of the IPEV model with simulation probability. Moreover, the welfare loss estimated using the MDCEV model is significantly lower than that estimated using the IPEV model. This result suggests that ignoring the integer property of demand may cause an underestimation of welfare loss.

We propose two additional investigation paths for the proposed IPEV model. First, the heterogeneity of the preferences in the IPEV model should be investigated. For multiple discrete-continuous demands, several models have been proposed to implement heterogeneity, including the random parameter KT (von Haefen and Phaneuf, 2005), mixed MDCEV (Bhat, 2008), latent segmentation KT (Kuriyama et al., 2010), and latent segmentation MDCEV (Wafa et al., 2015). The random parameter and latent segmentation approaches for the IPEV model may be worth investigating. Second, an IPEV model with multiple constraints should be examined. The IPEV model considers an individual's decision regarding consumption allocation within budget constraints. However, time may also be a fundamental constraint on leisure trips. Castro et al. (2012) proposed an MDCEV model with double constraints (budget and time) for continuous demand, and Kuriyama et al. (2020) considered an MDCEV model with triple constraints (budget, weekend time, and long holiday time). The development of an econometric model using an IPEV with multiple constraints is an important issue for the future.

Appendix A. The approximated locally optimal condition

First, consider $U(\tilde{x}_k^+) \leq U(\tilde{x})$. For the additively separable utility function defined in Equation (1), $U(\tilde{x}_k^+) \leq U(\tilde{x})$ can be written as

$$\frac{\psi_1}{\alpha_1}(\tilde{x}_1 - p_k)^{\alpha_1} + \frac{\gamma_k \psi_k}{\alpha_k}\left\{\left(\frac{\tilde{x}_k + 1}{\gamma_k} + 1\right)^{\alpha_k} - 1\right\} \leq \frac{\psi_1}{\alpha_1}\tilde{x}_1{}^{\alpha_1} + \frac{\gamma_k \psi_k}{\alpha_k}\left\{\left(\frac{\tilde{x}_k}{\gamma_k} + 1\right)^{\alpha_k} - 1\right\}. \tag{A1}$$

It can be rewritten as

$$\frac{\gamma_k \psi_k}{\alpha_k}\left\{\left(\frac{\tilde{x}_k + 1}{\gamma_k} + 1\right)^{\alpha_k} - \left(\frac{\tilde{x}_k}{\gamma_k} + 1\right)^{\alpha_k}\right\} \leq \frac{\psi_1}{\alpha_1}\{\tilde{x}_1{}^{\alpha_1} - (\tilde{x}_1 - p_k)^{\alpha_1}\}. \tag{A2}$$

The left-hand side is the utility difference of $U_k(\tilde{x}_k + 1) - U_k(\tilde{x}_k)$, and the right-hand side is the difference in $U_1(\tilde{x}_1) - U_1(\tilde{x}_1 - p_k)$. Note that $\psi_1 = \exp(\varepsilon_1)$ and $\psi_k = \exp(\beta' z_k + \varepsilon_k)$. Taking the logarithms on both sides, we obtain the following inequality:

$$V_k^+ + \varepsilon_k \leq V_{1,k}^+ + \varepsilon_1, \tag{A3}$$

where

$$\begin{aligned} V_k^+ + \varepsilon_k &= \beta' z_k + \ln\gamma_k - \ln\alpha_k + \ln\left[\left(\frac{\tilde{x}_k + 1}{\gamma_k} + 1\right)^{\alpha_k} - \left(\frac{\tilde{x}_k}{\gamma_k} + 1\right)^{\alpha_k}\right] + \varepsilon_k \\ &= \ln[U_k(\tilde{x}_k + 1) - U_k(\tilde{x}_k)]\text{, and} \end{aligned}$$
$$V_{1,k}^+ + \varepsilon_1 = -\ln\alpha_1 + \ln[\tilde{x}_1{}^{\alpha_1} - (\tilde{x}_1 - p_k)^{\alpha_1}] + \varepsilon_1 = \ln[U_1(\tilde{x}_1) - U_1(\tilde{x}_1 - p_k)].$$

Then, consider $U(\tilde{x}_k^-) \leq U(\tilde{x})$. This inequality can be written as

$$\frac{\psi_1}{\alpha_1}(\tilde{x}_1 + p_k)^{\alpha_1} + \frac{\gamma_k \psi_k}{\alpha_k}\left\{\left(\frac{\tilde{x}_k - 1}{\gamma_k} + 1\right)^{\alpha_k} - 1\right\} \leq \frac{\psi_1}{\alpha_1}\tilde{x}_1{}^{\alpha_1} + \frac{\gamma_k \psi_k}{\alpha_k}\left\{\left(\frac{\tilde{x}_k}{\gamma_k} + 1\right)^{\alpha_k} - 1\right\}. \tag{A4}$$

It can be expressed as

$$\frac{\psi_1}{\alpha_1}\{(\tilde{x}_1 + p_k)^{\alpha_1} - \tilde{x}_1{}^{\alpha_1}\} \leq \frac{\gamma_k \psi_k}{\alpha_k}\left\{\left(\frac{\tilde{x}_k}{\gamma_k} + 1\right)^{\alpha_k} - \left(\frac{\tilde{x}_k - 1}{\gamma_k} + 1\right)^{\alpha_k}\right\}. \tag{A5}$$

Take the logarithm on both sides, it can be rewritten as

$$V_{1,k}^{-} + \varepsilon_1 \leq V_k^{-} + \varepsilon_k, \tag{A6}$$

where

$$V_{1,k}^{-} + \varepsilon_1 = -\ln \alpha_1 + \ln[(\tilde{x}_1 + p_k)^{\alpha_1} - \tilde{x}_1{}^{\alpha_1}] + \varepsilon_1 = \ln[U_1(\tilde{x}_1 + p_k) - U_1(\tilde{x}_1)]\text{, and}$$

$$V_k^{-} + \varepsilon_k = \beta' z_k + \ln \gamma_k - \ln \alpha_k + \ln\left[\left(\frac{\tilde{x}_k}{\gamma_k} + 1\right)^{\alpha_k} - \left(\frac{\tilde{x}_k - 1}{\gamma_k} + 1\right)^{\alpha_k}\right] + \varepsilon_k$$
$$= \ln[U_k(\tilde{x}_k) - U_k\,(\tilde{x}_k - 1)].$$

Appendix B. Derivation of the closed-form probability for the IPEV model

Let $d_k^{+} = V_{1,k}^{+} - V_k^{+}$ and $d_k^{-} = V_{1,k}^{-} - V_k^{-}$. From (9), the probability of observing $\tilde{x}^* = (\tilde{x}_1^*, \tilde{x}_2^*, \dots, \tilde{x}_M^*, 0, \dots, 0)$ is

$$\begin{aligned}
P(\tilde{x}^*) &= \int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty} \left\{ \prod_{k=2}^{M} \left[G\left(\frac{d_k^{+} + \varepsilon_1}{\sigma}\right) - G\left(\frac{d_k^{-} + \varepsilon_1}{\sigma}\right) \right] \prod_{k=M+1}^{K} G\left(\frac{d_k^{+} + \varepsilon_1}{\sigma}\right) \right\} \frac{1}{\sigma} g\left(\frac{\varepsilon_1}{\sigma}\right) d\varepsilon_1, \\
&= \sum_{s_2=0}^{1} \cdots \sum_{s_M=0}^{1} \left[\int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty} \left\{ \prod_{k=2}^{M} \left[G\left(\frac{d_k^{+} + \varepsilon_1}{\sigma}\right) \right]^{s_k} \left[-G\left(\frac{d_k^{-} + \varepsilon_1}{\sigma}\right) \right]^{1-s_k} \prod_{k=M+1}^{K} G\left(\frac{d_k^{+} + \varepsilon_1}{\sigma}\right) \right\} \frac{1}{\sigma} g\left(\frac{\varepsilon_1}{\sigma}\right) d\varepsilon_1 \right] \\
&= \sum_{s_2=0}^{1} \cdots \sum_{s_M=0}^{1} \left[\int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty} \left\{ \prod_{k=2}^{M} \left[e^{-e^{-\left(\frac{d_k^{+}+\varepsilon_1}{\sigma}\right)}} \right]^{s_k} \left[-e^{-e^{-\left(\frac{d_k^{-}+\varepsilon_1}{\sigma}\right)}} \right]^{1-s_k} \prod_{k=M+1}^{K} e^{-e^{-\left(\frac{d_k^{+}+\varepsilon_1}{\sigma}\right)}} \right\} \frac{1}{\sigma} e^{-\frac{\varepsilon_1}{\sigma}} e^{-e^{-\frac{\varepsilon_1}{\sigma}}} d\varepsilon_1 \right] \\
&= \sum_{s_1=1}^{1} \sum_{s_2=0}^{1} \cdots \sum_{s_M=0}^{1} \left[\int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty} \left\{ \prod_{k=1}^{M} \left[e^{-e^{-\left(\frac{d_k^{+}+\varepsilon_1}{\sigma}\right)}} \right]^{s_k} \left[-e^{-e^{-\left(\frac{d_k^{-}+\varepsilon_1}{\sigma}\right)}} \right]^{1-s_k} \prod_{k=M+1}^{K} e^{-e^{-\left(\frac{d_k^{+}+\varepsilon_1}{\sigma}\right)}} \right\} \frac{1}{\sigma} e^{-\frac{\varepsilon_1}{\sigma}} d\varepsilon_1 \right],
\end{aligned} \tag{B1}$$

where $d_1^{+} = V_{1,1}^{+} - V_1^{+} = 0$. This probability can be written as

$$P(\tilde{x}^*) =$$

$$\sum_{s_1=1}^{1}\sum_{s_2=0}^{1}\cdots\sum_{s_M=0}^{1}\left[\left(\prod_{k=1}^{M}(-1)^{1-s_k}\right)\right.$$

$$\left.\cdot\int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty}\left\{\prod_{k=1}^{M} e^{-s_k\cdot e^{-\left(\frac{d_k^+ + \varepsilon_1}{\sigma}\right)}} e^{-(1-s_k)\cdot e^{-\left(\frac{d_k^- + \varepsilon_1}{\sigma}\right)}} \prod_{k=M+1}^{K} e^{-e^{-\left(\frac{d_k^+ + \varepsilon_1}{\sigma}\right)}}\right\}\frac{1}{\sigma}e^{-\frac{\varepsilon_1}{\sigma}}d\varepsilon_1\right] \tag{B2}$$

$$= \sum_{s_1=1}^{1}\sum_{s_2=0}^{1}\cdots\sum_{s_M=0}^{1}\left[\left(\prod_{k=1}^{M}(-1)^{1-s_k}\right)\cdot\int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty}\left\{\prod_{k=1}^{K} e^{-e^{-\left(\frac{d_k^* + \varepsilon_1}{\sigma}\right)}}\right\}\frac{1}{\sigma}e^{-\frac{\varepsilon_1}{\sigma}}d\varepsilon_1\right]$$

$$= \sum_{s_1=1}^{1}\sum_{s_2=0}^{1}\cdots\sum_{s_M=0}^{1}\left[\left(\prod_{k=1}^{M}(-1)^{1-s_k}\right)\cdot\int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty}\exp\left(-e^{-\frac{\varepsilon_1}{\sigma}}\sum_{k=1}^{K} e^{-\left(\frac{d_k^*}{\sigma}\right)}\right)\frac{1}{\sigma}e^{-\frac{\varepsilon_1}{\sigma}}d\varepsilon_1\right],$$

where $d_k^* = 1[k \le M]\cdot[s_k\cdot d_k^+ + (1-s_k)\cdot d_k^-] + 1[k > M]\cdot d_k^+$. Let $t = e^{-\frac{\varepsilon_1}{\sigma}}$. Then, $dt = -\frac{1}{\sigma}e^{-\frac{\varepsilon_1}{\sigma}}d\varepsilon_1$, and the integral of this expression can be written as

$$\int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty}\exp\left(-e^{-\frac{\varepsilon_1}{\sigma}}\sum_{k=1}^{K} e^{-\left(\frac{d_k^*}{\sigma}\right)}\right)\frac{1}{\sigma}e^{-\frac{\varepsilon_1}{\sigma}}d\varepsilon_1 = \int_0^{\infty}\exp\left(-t\sum_{k=1}^{K} e^{-\frac{d_k^*}{\sigma}}\right)dt = \left.\frac{\exp\left(-t\sum_{k=1}^{K} e^{-\frac{d_k^*}{\sigma}}\right)}{-\sum_{k=1}^{K} e^{-\frac{d_k^*}{\sigma}}}\right|_0^{\infty} \tag{B3}$$

$$= \frac{1}{\sum_{k=1}^{K} e^{-\frac{d_k^*}{\sigma}}}.$$

Thus, we can get the closed-form probability of the IPEV model:

$$P(\tilde{x}^*) = \sum_{s_1=1}^{1}\sum_{s_2=0}^{1}\cdots\sum_{s_M=0}^{1}\frac{\prod_{k=1}^{M}(-1)^{1-s_k}}{\sum_{k=1}^{K} e^{-\frac{d_k^*}{\sigma}}}. \tag{B4}$$

Appendix C. An example of three goods case

The closed-form expression of the probability of the IPEV model for the three goods is as follows:

(a) $\tilde{x}_1^*, \tilde{x}_2^*,$ and $\tilde{x}_3^* > 0$

$$\begin{aligned}P(\tilde{x}_1^*, \tilde{x}_2^*, \tilde{x}_3^*) &= \int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty} \left\{\left[G\left(\frac{d_2^+ + \varepsilon_1}{\sigma}\right) - G\left(\frac{d_2^- + \varepsilon_1}{\sigma}\right)\right]\left[G\left(\frac{d_3^+ + \varepsilon_1}{\sigma}\right) - G\left(\frac{d_3^- + \varepsilon_1}{\sigma}\right)\right]\right\}\frac{1}{\sigma}g\left(\frac{\varepsilon_1}{\sigma}\right)d\varepsilon_1 \\ &= \frac{1}{e^{-\frac{d_1^+}{\sigma}} + e^{-\frac{d_2^+}{\sigma}} + e^{-\frac{d_3^+}{\sigma}}} - \frac{1}{e^{-\frac{d_1^+}{\sigma}} + e^{-\frac{d_2^+}{\sigma}} + e^{-\frac{d_3^-}{\sigma}}} - \frac{1}{e^{-\frac{d_1^+}{\sigma}} + e^{-\frac{d_2^-}{\sigma}} + e^{-\frac{d_3^+}{\sigma}}} \\ &+ \frac{1}{e^{-\frac{d_1^+}{\sigma}} + e^{-\frac{d_2^-}{\sigma}} + e^{-\frac{d_3^-}{\sigma}}}.\end{aligned} \tag{C1}$$

(b) $\tilde{x}_1^*$ and $\tilde{x}_2^* > 0, \tilde{x}_3^* = 0$

$$\begin{aligned}P(\tilde{x}_1^*, \tilde{x}_2^*, 0) &= \int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty} \left\{\left[G\left(\frac{d_2^+ + \varepsilon_1}{\sigma}\right) - G\left(\frac{d_2^- + \varepsilon_1}{\sigma}\right)\right]G\left(\frac{d_3^+ + \varepsilon_1}{\sigma}\right)\right\}\frac{1}{\sigma}g\left(\frac{\varepsilon_1}{\sigma}\right)d\varepsilon_1 \\ &= \frac{1}{e^{-\frac{d_1^+}{\sigma}} + e^{-\frac{d_2^+}{\sigma}} + e^{-\frac{d_3^+}{\sigma}}} - \frac{1}{e^{-\frac{d_1^+}{\sigma}} + e^{-\frac{d_2^-}{\sigma}} + e^{-\frac{d_3^+}{\sigma}}}.\end{aligned} \tag{C2}$$

(c) $\tilde{x}_1^* > 0, \tilde{x}_2^*$ and $\tilde{x}_3^* = 0$

$$P(\tilde{x}_1^*, 0,0) = \int_{\varepsilon_1=-\infty}^{\varepsilon_1=+\infty} \left\{G\left(\frac{d_2^+ + \varepsilon_1}{\sigma}\right)G\left(\frac{d_3^+ + \varepsilon_1}{\sigma}\right)\right\}\frac{1}{\sigma}g\left(\frac{\varepsilon_1}{\sigma}\right)d\varepsilon_1 = \frac{1}{e^{-\frac{d_1^+}{\sigma}} + e^{-\frac{d_2^+}{\sigma}} + e^{-\frac{d_3^+}{\sigma}}}. \tag{C3}$$

Acknowledgments

The authors gratefully acknowledge financial support from Japan Society for the Promotion of Science under KAKENHI Grant Number 22H03812 and the Environment Research and Technology Development Fund (JPMEERF20224003) of the Environmental Restoration and Conservation Agency Provided by the Ministry of Environment of Japan.

References

Aarts, E., Lenstra, J. K. (Eds.), 2003. Local search in combinatorial optimization. Princeton University Press.

Bhat, C.R., 2005. A multiple discrete–continuous extreme value model: formulation and application to discretionary time-use decisions. Transp. Res. Part B 39(8), 679-707.

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

Bhat, C.R., 2022. A closed-form multiple discrete-count extreme value (MDCNTEV) model. Transp. Res. Part B 164, 65-86.

Bhat, C.R., Astroza, S., Sidharthan, R., Bhat, P. C., 2014. A multivariate hurdle count data model with an endogenous multiple discrete–continuous selection system. Transp. Res. Part B, 63, 77–97.

Bhat, C.R., Paleti, R., Castro, M., 2015. A new utility-consistent econometric approach to multivariate count data modeling. J. Appl. Econom. 30(5), 806–825.

Bockstael, N., Hanemann, W.M., Strand, I.E., 1986. Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models, vol. 2. Washington D.C.: U.S. Environmental Protection Agency, Office of Policy Analysis.

Cameron, A.C., Trivedi, P.K., 2013. Regression analysis of count data. Cambridge university press.

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

Cesario, F.J., 1976. Value of time in recreation benefit studies. Land Econ. 52 (1), 32–41.

Hanemann, W.M., 1978. A theoretical and empirical study of the recreation benefits from improving water quality in the boston area. Ph.D. Dissertation, Harvard University Department of Economics, Harvard University.

Hausman, J.A., Leonard, G.K., McFadden, D., 1995. A utility-consistent, combined discrete choice and count data model assessing recreational use losses due to natural resource damage. J. Public Econ. 56(1), 1-30.

Hellerstein, D., Mendelsohn, R., 1993. A theoretical foundation for count data models. Am. J. Agric. Econ. 75(3), 604-611

Herriges, J.A., Kling, C.L., and Phaneuf, D.J., 1999. Corner solution models of recreation demand: A comparison of competing frameworks. In: Herriges, J.A., Kling, C.L. (Eds.), Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice. Edward Elgar, 163-197.

Jara-Díaz, S., 2007. Transport economic theory. Emerald Group Publishing Limited.

Kuriyama, K. and Hanemann, W.M., 2006. The Integer Programming Approach to a Generalized Corner-Solution Model: an Application to Recreation Demand. Waseda University, Tokyo. Working paper.

Kuriyama, K., Hanemann, W.M., Hilger, J. R., 2010. A latent segmentation approach to a Kuhn–Tucker model: An application to recreation demand. J. Environ. Econ. Manage. 60(3), 209-220.

Kuriyama, K., Shoji, Y., Tsuge, T., 2020. The value of leisure time of weekends and long holidays: the multiple discrete–continuous extreme value (MDCEV) choice model with triple constraints. J. Choice Model. 37, 100238.

Lee, S., Allenby, G.M., 2014. Modeling indivisible demand. Market. Sci. 33 (3), 364–381.

Lloyd-Smith, P., 2018. A new approach to calculating welfare measures in Kuhn-Tucker demand models. J. Choice Model. 26, 19-27.

Phaneuf, D.J., Kling, C.L., Herriges, J.A., 2000. Estimation and welfare calculations in a generalized corner solution model with an application to recreation demand. Rev. Econ. Stat. 82(1), 83-92.

Pinjari, A.R., Bhat, C., 2021. Computationally efficient forecasting procedures for Kuhn-Tucker consumer demand model systems: application to residential energy consumption analysis. J. Choice Model. 39, 100283.

Sait, S.M., Youssef, H., 2000. Iterative Computer Algorithms with Applications in Engineering: Solving Combinatorial Optimization Problems. Wiley-IEEE Computer Society Press.

Sikder, S., Pinjari, A.R., 2013. The benefits of allowing heteroscedastic stochastic distributions in multiple discrete-continuous choice models. J. Choice Model. 9, 39-56.

Smith, V.K., 1997. Combining discrete choice and count data models: A comment. Duke University.

Train, K., 2009. Discrete Choice Methods with Simulation, second ed. Cambridge University Press, New York.

Vazirani, V.V., 2001. Approximation algorithms (Vol. 1). Berlin: Springer.

von Haefen, R.H., 2007. Empirical strategies for incorporating weak complementarity into consumer demand models. J. Environ. Econ. Manage. 54(1), 15-31.

von Haefen, R.H., Phaneuf, D.J., 2005. Kuhn–Tucker demand system approaches to non-market valuation. In: Scarpa, R., Alberini, A. (Eds.), Applications of Simulation Methods in Environmental and Resource Economics. Springer, Netherlands, pp. 135–157.

von Haefen, R.H., Phaneuf, D.J., Parsons, G.R., 2004. Estimation and welfare analysis with large demand systems. J. Bus. Econ. Stat. 22(2), 194-205.

Wafa, Z., Bhat, C.R., Pendyala, R.M., Garikapati, V.M., 2015. Latent-segmentation-based approach to investigating spatial transferability of activity-travel models. Transp. Res. Rec. 2493(1), 136-144.

Wales, T.J., Woodland, A., 1983. Estimation of consumer demand systems with binding non-negativity constraints. J. Econom. 21, 263-285.

Wolsey, L.A., Nemhauser, G.L., 1999. Integer and combinatorial optimization. John Wiley & Sons.

Wolsey, L.A., 1998. Integer Programming. Wiley-Interscience.

Zhang, Y., Zhou, H., Zhou, J., Sun, W., 2017. Regression models for multivariate count data. J. Comput. Graph. Statist. 26(1), 1-13.

Table 1

Estimation results of the simulation experiments [a]

	Assumed value	IPEV1 (closed-form)			IPEV2 (simulation)[b]		
Baseline Utility (β)		Estimates		Difference	Estimates		Difference
z_1	-1.0	-1.003	***	-0.003	-1.002	***	-0.002
		(0.024)		(0.024)	(0.024)		(0.024)
z_2	1.5	1.498	***	-0.002	1.498	***	-0.002
		(0.030)		(0.030)	(0.030)		(0.030)
z_3	-0.5	-0.504	***	-0.004	-0.503	***	-0.003
		(0.018)		(0.018)	(0.018)		(0.018)
Satiation parameters							
α_1	0.5	0.497	***	-0.003	0.497	***	-0.003
		(0.010)		(0.010)	(0.010)		(0.010)
γ_2	0.4	0.400	***	0.000	0.400	***	0.000
		(0.037)		(0.037)	(0.038)		(0.038)
γ_3	0.5	0.497	***	-0.003	0.498	***	-0.002
		(0.047)		(0.047)	(0.047)		(0.047)
γ_4	0.6	0.598	***	-0.003	0.598	***	-0.002
		(0.054)		(0.054)	(0.054)		(0.054)
Scale parameter (σ)	0.3	0.299	***	-0.002	0.300	***	0.000
		(0.010)		(0.010)	(0.010)		(0.010)

[a] *** $p < .01$. Estimates are the average of a Monte Carlo experiment with 200 iterations. The values in parentheses are standard errors.

[b] Parameters are estimated using simulation with 200 Halton draws.

Table 2
Trip distribution of the small choice set

Number of trips made per respondent	The number of observations	
0	522	69%
1	150	20%
2	37	5%
3	25	3%
4	12	2%
5 and over	6	1%
Total	752	100%

Table 3

Descriptive statistics of the small choice set

Variables	Description	Mean	Std. dev.	Min	Max
x	The number of trips per individual	0.521	1.143	0.000	18.000
Price	Two-way travel cost (1,000 yen)	42.503	20.229	1.817	106.362
Male	1/0, 1 = male	0.560	0.497	0.000	1.000
Age	Age of respondent	44.654	12.020	25.000	65.000
Income	Household income (1,000 yen)	5,910	3,231	2,000	20,000

Table 4

Estimation result of the small choice set[a]

Variables	MDCEV			IPEV1 (closed-form)			IPEV2 (simulation)[b]		
	Est.		t-stat.	Est.		t-stat.	Est.		t-stat.
Baseline utility (β)									
Constant	-2.889	***	-5.12	-1.497	**	-2.19	-1.460	**	-2.16
Male	0.294	***	3.12	0.355	***	3.26	0.359	***	3.12
Age	0.006	*	1.74	0.007		1.59	0.007		1.62
Satiation parameters									
α_1	0.728	***	8.36	0.696	***	7.27	0.702	***	6.83
γ_2	2.099		0.96	0.032	**	2.11	0.031		1.64
γ_3	1.631	*	1.79	0.048	***	2.83	0.047	**	1.98
γ_4	1.437	***	2.82	0.053	***	2.76	0.052	**	2.20
Scale parameter (σ)	0.650	***	21.59	0.783	***	17.17	0.785	***	13.69
Log-likelihood	-1405.28			-1151.74			-1151.82		

[a] * $p < .10$, ** $p < .05$, *** $p < .01$.

[b] Parameters are estimated using simulation with 200 Halton draws.

Table 5

Demand prediction of the small choice set[a]

Mean trips

Scenarios	MDCEV	IPEV1 (closed-form)	Difference 1[b]	IPEV2 (simulation)	Difference 2[c]
A: Closing Shiretoko	0.489	0.488	0.001	0.488	0.001
	(0.003)	(0.000)	(0.002)	(0.000)	(0.002)
B: Setting entrance fee of 5,000 yen in Shiretoko	0.443	0.439	0.004	0.439	0.004
	(0.049)	(0.020)	(0.030)	(0.028)	(0.021)

Trip distribution

Number of trips taken per respondent	Baseline		A: Closing Shirtoko						B: Setting entrance fee					
			MDCEV		IPEV1		IPEV2		MDCEV		IPEV1		IPEV2	
0	522	69%	533	71%	533	71%	533	71%	525	70%	523	70%	523	70%
1	150	20%	142	19%	142	19%	142	19%	147	20%	149	20%	149	20%
2	37	5%	38	5%	38	5%	38	5%	38	5%	37	5%	37	5%
3	25	3%	25	3%	25	3%	25	3%	24	3%	25	3%	25	3%
4	12	2%	10	1%	10	1%	10	1%	12	2%	12	2%	12	2%
5 and over	6	1%	4	1%	4	1%	4	1%	6	1%	6	1%	6	1%

[a] Average number of trips per person. The values in parentheses are standard errors calculated using the procedure reported by Krinsky and Robb (1986) based on 200 iterations.

[b] Difference 1 = MDCEV – IPEV1.

[c] Difference 2 = MDCEV – IPEV2.

Table 6

Compensating variations of the small choice set[a]

Scenarios	MDCEV	IPEV1 (closed-form)	Difference 1[b]	IPEV2 (simulation)	Difference 2[c]
A: Closing Shiretoko	-4.90	-24.10	-24.20	19.20	19.30
	(6.40)	(3.69)	(3.69)	(2.74)	(2.65)
B: Setting entrance fee of 5,000 yen in Shiretoko	-1.34	-1.46	-1.46	0.12	0.12
	(0.28)	(0.01)	(0.01)	(0.03)	(0.03)

[a] US$ per person. The values in parentheses are standard errors calculated using the procedure reported by Krinsky and Robb (1986) and based on 200 iterations.

[b] Difference 1 = MDCEV – IPEV1.

[c] Difference 2 = MDCEV – IPEV2.

Table 7

Trip distribution of the large choice set

Number of trips made per respondent	The number of observations	
0	983	62%
1	214	13%
2	108	7%
3	63	4%
4	50	3%
5	43	3%
6	16	1%
7	14	1%
8	14	1%
9	6	0%
10 and over	80	5%
Total	1591	100.0%

Table 8
Descriptive statistics of the large choice set

Variables	Description	Mean	Std. dev.	Min	Max
x	The number of trips per individual	2.065	6.027	0.000	72.000
Price	Two-way trip cost to each site (1000 yen)	78.381	41.460	1.780	222.587
Male	1/0, 1 = Male	0.503	0.500	0.000	1.000
Income	Household income (1000 yen)	6,261	3,833	2,000	20,000
World heritage	1/0, 1 = World Heritage Site	0.206	0.404	0.000	1.000
Visitor center	The number of visitor centers	1.118	1.795	0.000	8.000
Hot spring	1/0, 1 = Hot spring	0.353	0.478	0.000	1.000
Vehicle regulation	1/0, 1 = Vehicle regulation	0.559	0.497	0.000	1.000
Area 1	1/0, 1 = The park is located in the Hokkaido area	0.177	0.381	0.000	1.000
Area 2	1/0, 1 = The park is located in the Tohoku area	0.088	0.284	0.000	1.000
Area 3	1/0, 1 = The park is located in the Kanto area	0.177	0.381	0.000	1.000
Area 4	1/0, 1 = The park is located in the Chubu area	0.147	0.354	0.000	1.000
Area 5	1/0, 1 = The park is located in the Kinki area	0.088	0.284	0.000	1.000
Area 6	1/0, 1 = The park is located in the Chugoku-Shikoku area	0.088	0.284	0.000	1.000
Area 7	1/0, 1 = The park is located in the Kyushu area	0.294	0.208	0.000	1.000

Table 9
Estimation result of the large choice set[a]

Variables	MDCEV			IPEV (simulation)[b]		
	Est.		t-stat.	Est.		t-stat.
Baseline utility (β)						
Constant	-1.115	***	-11.85	0.586	***	6.24
Male	0.498	***	22.35	0.806	***	33.06
World Heritage	0.903	***	12.47	0.943	***	10.80
Visitor center	0.094	***	5.18	0.103	***	4.72
Hot spring	0.300	***	5.00	0.396	***	5.54
Vehicle regulation	-0.665	***	-10.50	-0.745	***	-9.66
Area 1	0.466	***	8.70	0.515	***	7.83
Area 2	0.151		1.47	0.209	*	1.68
Area 3	-0.722	***	-12.21	-0.534	***	-6.81
Area 4	-0.095		-1.46	0.051		0.65
Area 5	-0.264	***	-3.37	-0.067		-0.70
Area 6	0.261	**	3.13	0.296	**	2.85
Satiation parameters						
α_1	0.993	***	95.24	0.989	***	95.77
γ	1.225	***	22.46	0.032	***	8.34
Scale parameter (σ)	0.785	***	53.82	0.924	***	44.68
Log-likelihood	-9903.05			-8303.92		

[a] * $p < .10$, ** $p < .05$, *** $p < .01$.

[b] Parameters are estimated using simulation with 200 Halton draws.

Table 10
Demand prediction of the large choice set[a]

Mean trips

Scenarios	MDCEV	IPEV (simulation)	Difference[b]
A: Closing visitor centers	1.812	1.867	-0.054
	(0.150)	(0.029)	(0.144)
B: Closing Fuji-Hakone-Izu National Park	1.865	1.837	0.027
	(0.169)	(0.001)	(0.168)

Trip distribution

Number of trips taken per respondent	Baseline		A: Closing visitor centers				B: Closing Fuji-Hakone-Izu National Park			
			MDCEV		IPEV (Simulation)		MDCEV		IPEV (Simulation)	
0	983	62%	1004	63%	990	62%	1030	65%	1031	65%
1	214	13%	217	14%	220	14%	198	12%	198	12%
2	108	7%	102	6%	105	7%	105	7%	105	7%
3	63	4%	67	4%	70	4%	64	4%	64	4%
4	50	3%	49	3%	49	3%	43	3%	43	3%
5	43	3%	37	2%	36	2%	35	2%	35	2%
6	16	1%	14	1%	18	1%	14	1%	14	1%
7	14	1%	13	1%	16	1%	14	1%	14	1%
8	14	1%	8	1%	6	0%	10	1%	10	1%
9	6	0%	7	0%	5	0%	4	0%	4	0%
10 and over	80	5%	73	5%	76	5%	74	5%	73	5%

[a] Average number of trips per person. The values in parentheses are standard errors calculated using the procedure reported by Krinsky and Robb (1986) and based on 200 iterations.
[b] Difference = MDCEV – IPEV.

Table 11

Compensating variations of the large choice set[a]

Scenarios	MDCEV	IPEV (simulation)	Difference[b]
A: Closing visitor centers	-91.37	-209.38	118.00
	(12.48)	(33.25)	(24.06)
B: Closing Fuji-Hakone-Izu National Park	-32.76	-103.00	70.24
	(10.38)	(2.85)	(11.38)

[a] US$ per person. The values in parentheses are standard errors calculated using the procedure reported by Krinsky and Robb (1986) and based on 200 iterations.

[b] Difference = MDCEV – IPEV.