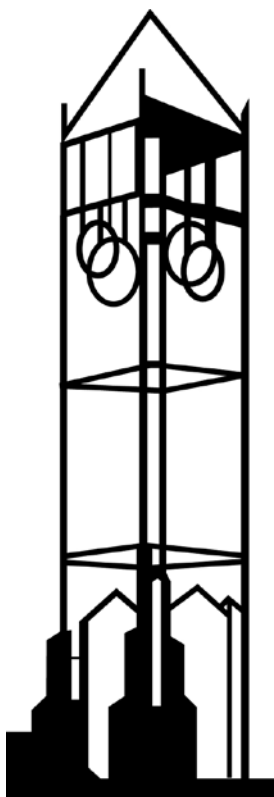


Controlling for Observed and Unobserved Site Characteristics in Rum Models of Recreation Demand

Babatunde Abidoye, Joseph A. Herriges, Justin Tobias



Working Paper No. 10011
May 2010

IOWA STATE UNIVERSITY
Department of Economics
Ames, Iowa, 50011-1070

Iowa State University does not discriminate on the basis of race, color, age, religion, national origin, sexual orientation, gender identity, sex, marital status, disability, or status as a U.S. veteran. Inquiries can be directed to the Director of Equal Opportunity and Diversity, 3680 Beardshear Hall, (515) 294-7612.

Controlling for Observed and Unobserved Site Characteristics in RUM Models of Recreation Demand¹

Babatunde O. Abidoye Joseph A. Herriges² Justin L. Tobias
Iowa State University Iowa State University Purdue University

May 25, 2010

Abstract

Random Utility Maximization (RUM) models of recreation demand are typically plagued by limited information on environmental and other attributes characterizing the available sites in the choice set. To the extent that these unobserved site attributes are correlated with the observed characteristics and/or the key travel cost variable, the resulting parameter estimates and subsequent welfare calculations are likely to be biased. In this paper we develop a Bayesian approach to estimating a RUM model that incorporates a full set of alternative specific constants, insulating the key travel cost parameter from the influence of the unobserved site attributes. In contrast to estimation procedures recently outlined in Murdock [21], the posterior simulator we propose (combining data augmentation and Gibbs sampling techniques) can be used in the more general mixed logit framework in which some parameters of the conditional utility function are random. Following a series of generated data experiments to illustrate the performance of the simulator, we apply the estimation procedures to data from the Iowa Lakes Project. In contrast to an earlier study using the same data (Egan *et al.* [7]), we find that, with the addition of a full set of alternative specific constants, water quality attributes no longer appear to influence the choice of where to recreate.

¹This research was supported in part by the U.S. Environmental Protection Agency and the Iowa Department of Natural Resources. Although the research described in this article has been funded in part by the United States Environmental Protection Agency through R82-5310-010, it has not been subject to the Agency's required peer review policy and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred. All remaining errors are, of course, our own.

²Contact author information: 260 Heady Hall, Department of Economics, Iowa State University, Ames, IA 50011. email: jaherrig@iastate.edu. Phone: 515-294-4964

1 Introduction

McFadden's Random Utility Maximization (or RUM) model provides the framework most often used to characterize recreation demand, linking the frequency of site visitation to individual attributes, the characteristics of alternatives in the choice set, and the travel cost required to reach each site. The estimated models can, in turn, be used to infer the value households place on access to sites and/or changes to site characteristics. Such information is key to policy-makers seeking to manage recreational resources. One advantage analysts have in modeling recreation demand is that, unlike most empirical demand studies, there is rich variation in the price data. The travel cost differs both across individuals and alternatives because of differences in each person's proximity to recreational sites. Unfortunately, variation in the price data is frequently offset by a paucity of information characterizing the attributes of the sites themselves. Researchers are often limited to one or two measures of site quality such as fish catch rates (Chen, Lupi and Hoehn [5] and Morey, Rowe and Watson [20]), fish toxin levels (Phaneuf, Kling, Herriges [23]) or dummy variable indicators capturing different levels of water quality (Parsons, Helm and Bondelid [22]).³ The risk in this setting is that unobserved site attributes may be correlated with the observed attributes or travel costs (or both), leading to omitted variables bias for the estimated parameters and biasing any subsequent welfare calculations.⁴

One solution to this problem is to include a full set of alternative specific constants (ASC's) when specifying the conditional utilities derived from visiting sites. These constants absorb and isolate the impact of site-specific attributes (including those unobserved by the analyst), allowing the key travel cost parameter to be consistently estimated. However, two problems emerge. First, when the available choice set is large, a full set of ASC's will greatly expand the parameter space, making the RUM model difficult to estimate. Second, the impacts that site attributes have on site selection are no longer identified, having been absorbed into the alternative specific constants. This limits the scope for policy analysis or regulators who are often interested in how changing site attributes (particularly a site's environmental conditions) will alter recreational usage patterns and the welfare of their constituent residents.

In an important recent article, Murdock [21] provides a resolution to both problems. Drawing on

³There are, of course, exceptions. Hanemann [10] highlights the importance of a large set of water quality attributes in determining site selection, including chemical oxygen demand (COD), phosphorus and fecal coliform bacteria levels. In a related study of beach usage in the Boston-Cape Cod area, Bockstael, Hanemann and Strand [3] employ a large number of water quality attributes, finding that these factors again are significant determinants of recreation demand.

⁴It is not hard to imagine possible correlations between observed and unobserved environmental attributes. For example, fish catch rates are likely to be lower in water bodies suffering from high pollution levels. The fish catch rates, in this case, might serve as a proxy for a myriad of water quality attributes affecting recreational site choices. Unobserved site attributes might also be correlated with travel costs, both because individuals might choose to locate closer to sites with higher water quality and because regulators may place a higher priority on improving the water quality of sites near population centers.

innovations in the industrial organization literature by Berry [1] and Berry, Levinsohn, and Pakes [2], Murdock suggests dividing the estimation task, employing a contraction mapping routine to estimate the alternative specific constants, while a standard maximum likelihood routine is used to estimate the model’s remaining parameters conditional on the estimated ASC’s. Whereas joint estimation of all of the RUM model’s parameters can be difficult, Murdock’s iterative approach is significantly faster and more stable, addressing the first problem noted above. To address the second issue (i.e., identification of the site attribute affects) Murdock suggests a second stage estimation in which the ASC’s are regressed on observed site attributes. As she notes, the advantage of this approach is that any concerns regarding correlation between observed and unobserved site attributes can be readily dealt with at this stage of the analysis using standard instrumental variable techniques in the context of a simple linear regression model.

There are, however, limitations to this estimation procedure. In particular, it cannot be used to obtain maximum likelihood parameter estimates if the RUM model includes random parameters (the so-called *mixed logit* model).^{5,6} Unfortunately, the mean-fitting feature of maximum likelihood estimation, which emerges from the standard logit model and underlies the insight behind Murdock’s [21] approach, does not hold once random parameters are introduced.⁷ Without this feature, the alternative specific constants obtained by the contraction mapping routine no longer solve the standard first order conditions implied by maximum likelihood estimation. In turn, this implies that the remaining parameter estimates for the RUM model, which are obtained conditional on the ASC’s, are also not maximum likelihood estimates. While forcing the alternative specific constants to insure mean fitting may be a desirable feature of an estimator, it is no longer clear what the statistical properties are of the resulting parameter estimates.⁸

The purpose of this paper is to provide an alternative approach to estimating the parameters of a RUM model including a full set of alternative specific constants, but one that does allow for the inclusion of random parameters. In particular, we propose a Bayesian approach using data augmentation and Gibbs sampling to characterize the posterior distribution of model’s parameters. Using a series of generated data experiments we demonstrate that our particular posterior simulator yields a posterior distribution for the key travel cost parameter that is insulated from the influence

⁵To our knowledge, Klaiber and von Haefen [14] were the first to note this limitation of the procedure.

⁶See Train [26] for a description of the mixed logit model.

⁷The mean-fitting nature of the logit model stems from its membership in the linear exponential family of distributions, which the mixed logit model is not a member of. In her technical appendix to [21], Murdock essentially uses a standard logit model (by conditioning on the random parameters) to prove that the first order conditions for maximum likelihood estimation imply that the estimates will be mean fitting. In an appendix to this paper, available from the authors upon request, a similar proof demonstrates that mean fitting is no longer implied by maximum likelihood estimation once random parameters are introduced.

⁸It should be noted that the limitation to Murdock’s procedure does not carry over to the earlier work of Berry [1] and Berry, Levinsohn, and Pakes [2]. In those papers, a contraction mapping is used to fit observed shares in the context of a GMM estimator and do not implicitly rely upon first order conditions derived from maximum likelihood estimation.

of unobserved site attributes, even those correlated with price or the observed site attributes.⁹

The influence of observed site characteristics on site visitation is captured using a hierarchical structure in the RUM model, allowing the distribution of the alternative specific constants to depend upon the observed site attributes. Unlike the post-estimation second stage regression used in Murdock [21], our approach proceeds jointly rather than sequentially and fully embraces the informational content provided by all stages of the hierarchy in the estimation process. That is, information provided by the hierarchical priors can be used to help “predict” the site-specific constants in addition to what is learned by the first-stage exercise of intercept estimation in the multinomial choice model. The adoption of such a hierarchical specification allows the researcher to borrow strength from what is learned about the estimation of other site-specific parameters and use it (in conjunction with site-level observables) to predict values of the given site-specific constant. As such, the hierarchical model serves to shrink the model estimates toward common means and helps mitigate concerns regarding overfitting, a common criticism of the highly-parameterized fixed effects model. While we do not focus on this in our empirical analysis, the hierarchical structure for the ASC’s can also be readily generalized to allow for possible correlation between the observed and unobserved site attributes using an instrumental variables approach along the lines described in Rossi, Allenby, and McCulloch [24] (section 7.1) and Lancaster [17] (chapter 8).

We illustrate our method using data from the Iowa Lakes Project, a large scale recreation demand study containing information on the visitation patterns of approximately 4,400 Iowa residents to the 130 primary recreational lakes in the state. One advantage of this study is that, in addition to household level usage data, detailed information is available on both site attributes and lake water quality. Moreover, this same data was recently used in Egan *et al.* [7] to estimate a RUM model of lake usage as a function site attributes, individual characteristics and travel cost, but without the use of alternative specific constants. The authors find that households significantly and substantially respond to both site characteristics and water quality attributes in deciding which lakes to visit. Using a similar specification that does not contain site-specific constants, yet estimated from a Bayesian point of view, we are able to replicate these qualitative results. Importantly, however, we find that once alternative specific constants are included in the model, the impact of water quality attributes is no longer clear, while site characteristics such as wake restrictions and boat ramps remain important factors. We also find that there is nearly a 20% drop in the coefficient of travel cost with the addition of ASC’s, which in turn leads to an increase in most welfare calculations by approximately 25%.

The outline of the paper is as follows. Section 2 provides a more detailed description of the basic

⁹By “insulated” we mean that the posterior is approximately centered around the parameter of the (known) data generation process and tends to collapse around this value as the sample size grows. This result is specific to the posterior simulator we employ - had we chosen to implement traditional blocking approaches with panel data, we would not obtain this desirable sampling performance.

RUM model, highlighting the potential for omitted variables bias and summarizing the related literature. Section 3 then describes our proposed method for estimating the parameters of the model, including a full set of alternative specific constants. A series of generated data experiments is employed in Section 4 to illustrate the performance of our posterior simulator under varying assumptions regarding unobserved site attributes. Section 5 provides a description of the Iowa Lakes Project and the data used in our empirical analysis. Estimation results and model comparison exercises are provided in Section 6 and the implications of various model specifications for welfare calculations in Section 7. The paper concludes with a summary in Section 8.

2 Controlling for Unobserved Site Attributes

As noted above, a significant concern in the recreation demand literature is that the analyst typically has relatively few attributes characterizing the individual sites in the choice set. To the extent that unobserved site attributes are correlated with either observed site attributes or the travel cost variable (or both), the resulting parameter estimates and subsequent welfare analysis will be contaminated by this correlation. Using both a generated data experiment and an empirical application to recreational fishing in Wisconsin, Murdock demonstrates that ignoring the unobserved site characteristics can “. . . cause biased standard errors that can outrageously overstate the precision of the [parameter] estimates. . .” (Murdock [21], p. 14.) and welfare predictions that are off by up to a factor of four.

The nature of the issue can be illustrated using a simple RUM model. Suppose the utility individual i receives from visiting site j is a linear function of a vector of site attributes (\mathbf{s}_j), the travel cost required to visit the site (p_{ij}) and an idiosyncratic error component (ϵ_{ij}) that is uncorrelated across sites and individuals and uncorrelated with either \mathbf{s}_j or p_{ij} (e.g., ϵ_{ij} is *i.i.d.* extreme value). That is

$$U_{ij} = \mathbf{s}_j \boldsymbol{\alpha}_0 + p_{ij} \beta + \epsilon_{ij} \quad i = 1, 2, \dots, N; j = 1, 2, \dots, J. \quad (1)$$

Given distributional assumptions regarding the ϵ_{ij} 's, choice probabilities can be derived for each individual and alternative, providing the basis for estimating parameters of the conditional utility function in (1). Unfortunately, the analyst may only observe a subset of the site attributes, (\mathbf{s}_j^o), leading to a reduced form specification

$$U_{ij} = \mathbf{s}_j^o \boldsymbol{\alpha}_0^o + p_{ij} \beta + \tilde{\epsilon}_{ij} \quad i = 1, 2, \dots, N; j = 1, 2, \dots, J. \quad (2)$$

where

$$\tilde{\epsilon}_{ij} = \mathbf{s}_j^u \boldsymbol{\alpha}_0^u + \epsilon_{ij} \quad i = 1, 2, \dots, N; j = 1, 2, \dots, J. \quad (3)$$

In (3), \mathbf{s}_j^u denotes the unobserved site attributes and $\boldsymbol{\alpha}_0^k$ ($k = o, u$) denote the subset of parameters associated with \mathbf{s}_j^k ($k = o, u$).¹⁰ Given this specification, consistent estimation of the parameters $\boldsymbol{\alpha}_0^o$ and β will require that the observed site characteristics and travel cost variables be uncorrelated with the unobserved characteristics. However, in applications where there are numerous unobserved site attributes, this condition is unlikely to hold, resulting in correlation between the error term $\tilde{\epsilon}_{ij}$ and the included explanatory variables and leading to the classic omitted variables bias (and inconsistency) problem.

One solution to this problem is to introduce a full set of alternative specific constants to capture the unobserved site attributes. In particular, letting $\alpha_j^u \equiv \mathbf{s}_j^u \boldsymbol{\alpha}_0^u$, equation (1) becomes

$$U_{ij} = \alpha_j^u + \mathbf{s}_j^o \boldsymbol{\alpha}_0^o + p_{ij} \beta + \epsilon_{ij} \quad i = 1, 2, \dots, N; j = 1, 2, \dots, J. \quad (4)$$

Unfortunately, perfect collinearity between the alternative specific constant α_j^u and the observed site effects, $\mathbf{s}_j^o \boldsymbol{\alpha}_0^o$, will preclude identification of both α_j^u and $\boldsymbol{\alpha}_0^o$. Instead, one can only identify an overall alternative specific constant

$$\alpha_j = \mathbf{s}_j^o \boldsymbol{\alpha}_0^o + \mathbf{s}_j^u \boldsymbol{\alpha}_0^u \quad (5)$$

$$= \mathbf{s}_j^o \boldsymbol{\alpha}_0^o + \alpha_j^u \quad (6)$$

capturing the total impact of the site characteristics (observed and unobserved) on latent utility. That is, we can employ the model:

$$U_{ij} = \alpha_j + p_{ij} \beta + \epsilon_{ij} \quad i = 1, 2, \dots, N; j = 1, 2, \dots, J. \quad (7)$$

This resolves the omitted variables problem since the error term (ϵ_{ij}) is once again uncorrelated with the explanatory variable (p_{ij}). Unfortunately, in addressing the omitted variables issue, we have created two new problems. First, there are now $J-1$ alternative specific constants to estimate, which can be challenging when the choice set is large. Second, the impact of the site characteristics on consumer welfare is no longer separately identified.

If the individual utilities (i.e., the U_{ij} 's) in equation (7) were observable, we would have a classic linear regression model and the alternative specific constants could be treated as fixed effects. In this setting, familiar partitioned regression techniques could be used to ease the computational burden of estimating the many ASC's. However, given the nonlinear nature of the RUM model, these techniques are not available. Murdock's solution, however, is somewhat analogous. She

¹⁰The specification in equation (2) is similar to that used in Egan *et al.* [7] and in much of the recreation demand literature.

uses a contraction mapping routine, together with the mean-fitting nature of maximum likelihood estimation (MLE) in the logit setting (i.e., imposing that the actual and fitted shares are equal under MLE) to separate the estimation of a full set of alternative specific constants from the estimation of the remaining parameters. Once the ASC's are estimated, the relatively small number of remaining parameters are obtained using standard maximum likelihood estimation, conditioning on the ASC's. This is an elegant solution to the problem, with both steps in the estimation process being easy to implement. Murdock goes on to suggest that the role of the observed site attributes in determining recreation demand can be captured using a second stage regression that fits the linear regression model implicit in equation (6), replacing the ASC's (i.e., the α_j 's) with their fitted values from the first stage and treating α_j^u as the error term.¹¹ Murdock observes that any omitted variables bias resulting from correlation between the observed site attributes in (6) and the unobserved site attributes imbedded in α_j^u can be handled using instrumental variables techniques. As noted in the introduction, the principle drawback to the method proposed by Murdock [21] is that it does not generalize to the mixed logit setting, which allows for preference heterogeneity across individuals through the use of random parameters. Moreover, the parameters α_0^o are informative for the ASC parameters, and we fail to capitalize upon this source of learning in the sequential approach to estimation.

In the next section we propose an alternative to Murdock's two-step procedure that can be used in the mixed logit setting. Before proceeding with the technical details, some intuition as to why our approach works may help. We approach the estimation problem from a Bayesian perspective, combining data augmentation and Gibbs sampling to characterize the posterior distribution of the model's parameters, but in the process draw on results in the standard fixed effects model familiar to non-Bayesians.¹² The data augmentation aspect of the simulator involves treating the unobserved latent site utilities (i.e., the U_{ij} 's) as additional parameters of the model. At each stage a simulated value for this otherwise missing information is obtained based on the observed decisions made by each individual. The key to the approach is that, conditional on these draws of the latent utilities, the model is effectively linear, and thus the problem of characterizing the posterior distribution of the parameters in (7) (i.e., the ASC's α_j and the travel cost parameter β) proceeds in a manner very similar to the classic fixed effects model. Indeed, with a diffuse prior on the parameters, the corresponding posterior mean of $(\alpha_1, \dots, \alpha_{J-1}, \beta)$ reduces to the non-Bayesian's fixed-effects estimator. By blocking together the simulation of the conditional posterior distribution of ASC's and the travel cost parameter, we isolate the impact of the unobservables (capturing them entirely

¹¹Note that this regression will have only $J - 1$ observations.

¹²There have been several papers using a Bayesian framework to estimate a model similar to that originally proposed in Berry, Levinsohn and Pakes [2], including a full set of alternative specific constants to control for unobserved alternative attributes. Yang, Chen, Allenby [27] develop a posterior simulation alternative to Berry, Levinsohn and Pakes [2] in modeling aggregate supply and demand. However, the routine is conditional on correctly specifying the underlying supply relationships. Jiang, Manchanda and Rossi [13] provide a Bayesian counterpart to the contraction mapping approach outlined in Berry, Levinsohn and Pakes [2], though again the analysis is couched in the context of aggregate supply and demand data.

in the alternative specific constants) and insulate the travel cost parameter from their effects, much like the standard result that the fixed effects estimator is unbiased even when correlation exists between the fixed effects and other explanatory variables included in the model.

3 Model

The basic RUM model presented in the previous section considers only a single choice from among the available alternatives in the choice set. In order to capture the observed outcome that individuals often take multiple trips to one or more sites during a course of a season, it is common practice in the recreation demand literature to employ the repeated logit model (Morey, Rowe and Watson [20], Herriges and Phaneuf [11]). In this extension of the basic RUM model, individuals are assumed to repeatedly choose from among the same set of alternatives over a fixed number of choice occasions. Furthermore, each decision is assumed conditionally independent across individuals and choice occasions.¹³ The particular form of the model we use is the repeated mixed logit (RXL) model employed by Egan *et al.* [7] (allowing individual parameters of the model to be random), but with the addition of the full set of alternative specific constants advocated by Murdock [21]. This section begins by describing the structure of the RXL model and developing the necessary notation, followed by a specification of the prior distributions employed in our analysis and a description of the Gibbs sampler used to generate draws from the posterior distribution.

3.1 The Repeated Logit Model

In the repeated logit model it is assumed that, on each choice occasion t , individual i chooses from among $J + 1$ alternatives, including the option to “stay at home” ($j = 0$). We assume the conditional utility individual i derives from alternative j at time t is given by:

$$U_{ijt} = \begin{cases} z_i\gamma + \varepsilon_{i0t} & j = 0 \\ \alpha_j + p_{ij}\beta + \varphi_i + \varepsilon_{ijt} & j = 1, \dots, J. \end{cases} \quad (8)$$

In this form of the model, the utility from visiting any one of the recreation sites (i.e., $j = 1, \dots, J$) is decomposed into an overall site-specific effect (α_j), a price (or travel cost) effect (p_{ij}), an individual specific effect φ_i , and an idiosyncratic error term ε_{ijt} . The term φ_i is included in the model to allow for heterogeneity in preferences to recreate across individuals and, specifically, we assume $\varphi_i \sim \mathcal{N}(0, \sigma_\varphi^2)$. The parameter σ_φ^2 is estimated within the model and characterizes the extent of

¹³Herriges, Kling and Phaneuf [12] provide a summary of the repeated logit model and the implications of its underlying assumptions.

variation in preferences to recreate in the population. The mean of this assumed normal distribution is restricted to be zero for identification purposes, as a non-zero mean (and thus an overall intercept parameter) will be introduced through our hierarchical prior for α_j . *Ceteris paribus*, an individual with a small φ_i is more likely to stay at home on a given choice occasion than someone with a larger φ_i . The inclusion of these individual-level heterogeneity terms thus mimics the standard nested logit structure in which all of the recreational sites are included in a single nest (See Herryges and Phaneuf [11]).

The idiosyncratic error term ε_{ijt} captures any remaining unobservable aspects of conditional utility and is assumed to be independent across the $J + 1$ alternatives. We assume $\varepsilon_{ijt} \stackrel{iid}{\sim} \mathcal{N}(0, 1)$. This assumption can be relaxed, and more flexible correlation and substitution patterns permitted across the alternatives. We maintain this assumption here, however, both because it is rather common in the recreation demand literature, mimics the often-used nested logit structure in empirical practice, and the complexity of the current model leads us to consider this parsimonious variant of the model as a starting point. Finally, individual demographic characteristics (such as age and gender) are assumed to impact the individual's propensity to stay at home, through the term $\mathbf{z}_i\gamma$ in equation (8), but are assumed to not impact the relative preference for any given recreation site. Such an extension could, again, be relaxed by allowing $\gamma = \gamma_j$, although this generalization may potentially introduce many new parameters in the model.

The choice among the alternatives on any given choice occasion depends, of course, only on relative utility levels. We use the stay-at-home-option as the base alternative, defining:

$$\tilde{U}_{ijt} = U_{ijt} - U_{i0t} = \alpha_j + p_{ij}\beta - \mathbf{z}_i\gamma + \varphi_i + \tilde{\varepsilon}_{ijt} \quad (9)$$

where $\tilde{\varepsilon}_{ijt} = \varepsilon_{ijt} - \varepsilon_{i0t}$ for $j = 1, \dots, J$. Stacking the error differences over alternatives, let

$$\tilde{\boldsymbol{\varepsilon}}_{i \cdot t} = \begin{bmatrix} \varepsilon_{i1t} - \varepsilon_{i0t} \\ \varepsilon_{i2t} - \varepsilon_{i0t} \\ \vdots \\ \varepsilon_{iJt} - \varepsilon_{i0t} \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}^*)$$

where

$$\boldsymbol{\Sigma}^* = \begin{bmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ 1 & 1 & \ddots & \vdots \\ 1 & 1 & \cdots & 2 \end{bmatrix}$$

Stacking all the variables across alternatives, we then have

$$\tilde{\boldsymbol{U}}_{i \cdot t} = \boldsymbol{\alpha} + \mathbf{p}_i\beta - (\mathbf{1}_J \otimes \mathbf{z}_i)\gamma + \mathbf{1}_J\varphi_i + \tilde{\boldsymbol{\varepsilon}}_{i \cdot t}, \quad (10)$$

where $\mathbf{1}_J$ is a $J \times 1$ vector of ones,

$$\tilde{\mathbf{U}}_{i \cdot t} = \begin{bmatrix} \tilde{U}_{i1t} \\ \tilde{U}_{i2t} \\ \vdots \\ \tilde{U}_{iJt} \end{bmatrix}; \quad \boldsymbol{\alpha} \cdot = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_J \end{bmatrix}; \quad \text{and} \quad \mathbf{p}_{i \cdot} = \begin{bmatrix} p_{i1} \\ p_{i2} \\ \vdots \\ p_{iJ} \end{bmatrix}.$$

Grouping our covariates together, the vector of utility differences can be written more compactly as

$$\tilde{\mathbf{U}}_{i \cdot t} = \mathbf{M}_{i \cdot t} \boldsymbol{\theta} + \mathbf{v}_{i \cdot t}, \quad (11)$$

where

$$\mathbf{M}_{i \cdot t} = [\mathbf{I}_J \quad \mathbf{p}_{i \cdot} \quad \mathbf{1}_J \otimes \mathbf{z}_i] \quad \text{and} \quad \boldsymbol{\theta} = [\boldsymbol{\alpha} \cdot' \quad \beta \quad \boldsymbol{\gamma}'']'.$$

Although $\mathbf{M}_{i \cdot t}$ does not formally depend on t in our application, it may in other instances, and we continue to make use of this notation here to remind us of the assumed repeated nature of the decision problem. Finally, $\mathbf{v}_{i \cdot t}$ is the composite error term

$$\mathbf{v}_{i \cdot t} = \mathbf{1}_J \varphi_i + \tilde{\boldsymbol{\varepsilon}}_{i \cdot t} \quad (12)$$

where $E(\mathbf{v}_{i \cdot t}) = 0$ and

$$E(\mathbf{v}_{i \cdot t} \mathbf{v}_{i \cdot t}') \equiv \boldsymbol{\Omega} = \sigma_\varphi^2 \mathbf{1}_J \mathbf{1}_J' + \boldsymbol{\Sigma}^*. \quad (13)$$

The observed choice y_{it} is linked to the latent variable vector $\tilde{\mathbf{U}}_{i \cdot t}$ as follows:

$$y_{it}(\tilde{\mathbf{U}}_{i \cdot t}) = \begin{cases} 0 & \text{if } \max\{\tilde{U}_{ijt}\}_{j=1}^J \leq 0 \\ k & \text{if } \max\{\tilde{U}_{ijt}\}_{j=1}^J = \tilde{U}_{ikt} > 0. \end{cases} \quad (14)$$

What we observe for every individual is a count of the number of visits to the full menu of potential sites over a given period of time, which for us represents a calendar year. Within the RXL framework, we imagine that a series of decisions were made by the individual at particular choice occasions - which in our case is weekly - in a manner that is consistent with this aggregate data. For example, if we know that a person visits just a single site k once and nothing else, then in 51 of the 52 cases, y_{it} takes on the value of the stay-at-home option (0), while in the remaining case, $y_{it} = k$. The actual ordering of these occurrences is not informed by the likelihood function, as nothing in the model depends on t , and we do not have information on the specific timing of decisions within the data. Our posterior simulator, then, will sample the U_{ijt} for $t = 1, 2, \dots, 52$ in a manner that satisfies the observed information on the total number of visits (and non-visits), and the particular order in which this is done is irrelevant for estimation and inferential purposes.

3.2 Hierarchical Priors

As described in the previous section, the alternative specific constants (α_j) play a particularly important role in the model. The α_j 's capture both observed and unobserved attributes of the site

that might influence a person’s propensity to visit that site (as in equation (5)).¹⁴ The alternative specific constants also provide the sole avenue by which the observed site attributes impact the recreation demand decision. In Murdock’s [21] two-stage estimation procedure, this is accomplished in the second stage, in which the fitted $\hat{\alpha}_j$ ’s are regressed on the observed site attributes, \mathbf{s}_j^o . In our Bayesian approach, this is captured by incorporating a hierarchical structure into our model, assuming that the α_j ’s are drawn from an underlying distribution whose mean is a function of the observed site characteristics; i.e.,

$$\alpha_j \stackrel{ind}{\sim} \mathcal{N}(\mathbf{q}_j \boldsymbol{\alpha}_0, \sigma_\alpha^2). \quad j = 1, 2, \dots, J \quad (15)$$

where \mathbf{q}_j includes a constant term and the observed site characteristics that influence demand for site j . This simple hierarchical structure is mostly silent about any possible correlation between unobserved site attributes and the observed attributes included in \mathbf{q}_j , although it is often assumed - at least implicitly - that these unobserved characteristics are uncorrelated with those in \mathbf{q}_j . If this assumption does not hold, then $\boldsymbol{\alpha}_0$ will simply capture the correlation between the α_j ’s and the observed site attribute, rather than a causal relationship, suffering from a form of omitted variables bias. In these instances, the hierarchical structure for the ASC’s can be readily generalized to allow for possible correlation between the observed and unobserved site attributes using an instrumental variables approach along the lines described in Rossi, Allenby, and McCulloch [24] (section 7.1) and Lancaster [17] (chapter 8).

We do not explore this possibility in the present paper, as no compelling instruments were identified in our data set and we do have available an extensive list of site attributes. In this regard we recognize that application of our methods will not solve all problems - to the extent that relevant unobserved site characteristics are omitted, yet correlated with observed site characteristics, our posterior estimates of the $\boldsymbol{\alpha}_0$ parameters will continue to be plagued by poor sampling properties. However, even when such correlation and confounding is present, the inclusion of site-specific constants, in conjunction with our particular posterior simulator, will yield accurate estimates of the first-stage parameters in (8). We will elaborate on this issue when describing our generated data experiments in the following section.

To complete our model, we specify priors for the remaining parameters. These are enumerated below:

¹⁴We assume that these site-specific effects are constant over both time and individual. The model could readily be generalized to allow for heterogeneity of preferences towards the site attributes by allowing the α_j to vary over individuals with some common mean. Allowing the site-effects to vary over time is substantially more difficult in that most recreation demand data sets do not have diary data regarding when individuals visit specific sites, but rather simply record how many times each site is visited over the course of a season.

$$\boldsymbol{\alpha}_0 \sim \mathcal{N}(\boldsymbol{\mu}_\alpha, \mathbf{V}_\alpha) \quad (16)$$

$$\beta \sim \mathcal{N}(\mu_\beta, V_\beta) \quad (17)$$

$$\boldsymbol{\gamma} \sim \mathcal{N}(\boldsymbol{\mu}_\gamma, \mathbf{V}_\gamma) \quad (18)$$

$$\sigma_\alpha^2 \sim \mathcal{IG}(a_\alpha, b_\alpha) \quad (19)$$

$$\sigma_\varphi^2 \sim \mathcal{IG}(a_\varphi, b_\varphi) \quad (20)$$

The hyperparameters of the priors above are supplied by the researcher and are in general chosen to be relatively vague to allow dominance of the information from the data. While \mathcal{N} above obviously refers to the normal distribution, $\mathcal{IG}(\cdot, \cdot)$ follows the notation in Koop, Poirier and Tobias [16] (pp. 336) and represents the *inverse gamma* distribution. The prior means $(\boldsymbol{\mu}_\alpha, \mu_\beta, \boldsymbol{\mu}_\gamma)$ in our empirical work and generated data experiments are set to zero vectors of appropriate dimensions with the respective prior variance for the parameters $(\mathbf{V}_\alpha, V_\beta, \text{ and } \mathbf{V}_\gamma)$ set to identity matrices of the appropriate dimensions. We also select the hyperparameters of the variances by choosing $a_\alpha = 3$; $b_\alpha = 5$ and $a_\varphi = 3$; $b_\varphi = 5$. This leads to a reasonably non-informative prior for the variances with prior mean and standard deviation equal to 0.1.¹⁵

3.3 Posterior Simulator

Let

$$\boldsymbol{\Xi} = [\boldsymbol{\theta} \quad \boldsymbol{\alpha}_0 \quad \sigma_\alpha^2 \quad \{\varphi_i\} \quad \sigma_\varphi^2]$$

denote all the parameters of the model. The joint posterior distribution of $\boldsymbol{\Xi}$ and the latent utility \tilde{U} defines the augmented posterior density for the parameters in our model. By Bayes theorem this posterior density is obtained as:

$$\begin{aligned} p(\boldsymbol{\Xi}, \tilde{U} | \mathbf{y}) &\propto \left[\prod_{t=1}^T \prod_{i=1}^N \phi(\tilde{U}_{i \cdot t}; \mathbf{M}_{i \cdot t} \boldsymbol{\theta}, \boldsymbol{\Omega}) \right. \\ &\quad \times \left\langle I(y_{it} = j) I(\tilde{U}_{ijt} > \max[\tilde{U}_{i, -j, t}, 0]) + I(y_{it} \neq j) I(\tilde{U}_{ijt} < \max[\tilde{U}_{i, -j, t}, 0]) \right\rangle \left. \right] \\ &\quad \times \left[\prod_{j=1}^J p(\alpha_j | \boldsymbol{\alpha}_0, \sigma_\alpha^2) \right] \left[\prod_{i=1}^N p(\varphi_i | \sigma_\varphi^2) \right] p(\beta) p(\boldsymbol{\gamma}) p(\boldsymbol{\alpha}_0) p(\sigma_\alpha^2) p(\sigma_\varphi^2). \end{aligned} \quad (21)$$

As mentioned previously, the individual y_{it} data are not directly observed, but are constructed to be consistent with the total number of trips taken to all of the sites over a given period of

¹⁵It should be noted that in cases where the data provides little information such as “small” J , the priors can be quite influential when making posterior inferences concerning these common parameters.

time. We construct the individual y_{it} artificially, though without loss of generality, to match these aggregate counts; the timing of when these decisions are assumed to occur does not affect the augmented posterior distribution or its simulator. Therefore, our particular assignment of the y_{it} values, provided they properly reproduce the total counts, is arbitrary.

We fit the above model using Markov Chain Monte Carlo (MCMC) methods, drawing specifically upon Gibbs sampling techniques. The Gibbs sampler makes use of the fact that, while joint posterior distributions frequently take unrecognizable forms (making them difficult to draw from), the conditional posterior distributions for individual blocks (or partitions) of the parameter space will often fall into well known distributional families that can be readily drawn from. Sequentially drawing from the posterior conditional distributions will lead to drawing from the joint posterior distribution of interest. In our particular implementation of the Gibbs sampler, the parameters α_j , β and γ are blocked together. This not only improves the mixing of the posterior simulator, but also preserves some desirable sampling properties of the posterior estimates of β . We will revisit this point when conducting our generated data experiments.

Step 1: Draw the $\tilde{U}_{i,t}|\Xi, \mathbf{y}$

Given the structure of our model, and to ease computation, we draw the latent utilities that individual i derives from visiting site j as an intermediate step in drawing the necessary utility differences. That is, we sample the U_{ijt} and then take differences from the baseline utility U_{i0t} to obtain the \tilde{U}_{ijt} . Drawing the U_{ijt} is straightforward, since conditional on α_j, β, γ , and $\{\varphi_i\}$ there is no correlation among the alternatives or correlation across individuals. The posterior conditional distributions for the U_{ijt} 's are univariate truncated normal with mean μ_{ij} and variance of 1 and a truncation point dictated by the visitation pattern of the individual.¹⁶ In particular, if an alternative is chosen, it must be the alternative that gives maximum utility. This places a lower bound on the utility of the chosen alternative and an upper truncation point for all the other alternatives.

We use the following steps to draw the \tilde{U}_{ijt} 's at a given draw r :

Assuming that individual i chooses alternative k at choice occasion t ,

- 1: Draw U_{ijt}^r for all $j \neq k$ from a truncated normal distribution with mean μ_{ij} , a variance of 1, and upper truncation point $U_{ikt} = U_{ikt}^{r-1}$.
- 2: Draw U_{ikt}^r from a truncated normal distribution with its mean μ_{ij} , a variance of 1, and a lower truncation point at the $\max(U_{ijt}^r)$ for all $j \neq k$.

¹⁶From equation (8), $\mu_{ij} = \alpha_j + P_{ij}\beta + \varphi_i$ for $j = 1, \dots, J$ and $\mu_{i0} = z_i\gamma$.

3: Calculate \tilde{U}_{ijt} by taking the difference between utilities from all sites and the stay at home option: $\tilde{U}_{ijt}^r = U_{ijt}^r - U_{i0t}^r$.

Step 2: $\theta | \Xi_{-\theta}, \tilde{U}, \mathbf{y}$

In this case and in all the steps that follow, Ξ_{-m} in the conditioning implies that we condition on all the parameters in Ξ other than m . Using the result of Lindley and Smith (1972), the posterior conditional for θ is given as:

$$\theta | \Xi_{-\theta}, \tilde{U}, \mathbf{y} \sim \mathcal{N}(\mathbf{D}_\theta \mathbf{d}_\theta, \mathbf{D}_\theta). \quad (22)$$

where

$$\begin{aligned} \mathbf{D}_\theta &\equiv \left[T \sum_{i=1}^N \mathbf{M}'_{i,t} \Omega^{-1} \mathbf{M}_{i,t} + \Sigma_\theta^{-1} \right]^{-1} \\ \mathbf{d}_\theta &\equiv \sum_t \sum_i \mathbf{M}'_{i,t} \Omega^{-1} \tilde{U}_{i,t} + \Sigma_\theta^{-1} \mu_\theta \end{aligned}$$

and

$$\Sigma_\theta = \begin{bmatrix} \sigma_\alpha^2 I_J & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & V_\beta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{V}_\gamma \end{bmatrix}, \quad \mu_\theta = \begin{bmatrix} \mathbf{Q}\alpha_0 \\ \mu_\beta \\ \mu_\gamma \end{bmatrix}.$$

The blocking strategy adopted in this step is key to our proposed simulator, jointly drawing the parameters $\alpha.$, β and γ . This blocking not only helps in improving the mixing of the sampler, but also avoids contamination of the travel cost parameter β stemming from unobserved site attributes correlated with travel cost. A natural alternative blocking strategy, performed by many in practice without recognition of its consequences, proceeds by drawing β and γ jointly from the posterior conditional marginalized over the α_j , and then drawing each α_j independently from their complete posterior conditional distributions. This simulator would not achieve the same objective. In short, the steps used to integrate α_j out of the β, γ conditional in this approach *assume* independence of the errors of (15) and other covariates in the model, even though such correlation may be present. Our simulator proceeds without having to model this correlation, and without needing to be concerned about it. The analogy here would be between the standard fixed and random effects estimators in the non-Bayesian paradigm. In the standard fixed effects estimator, the parameter estimates are robust to potential correlation between the regressors of the model and fixed effects. The same is not true of the random effects estimator. Our blocking strategy simultaneously simulates the alternative specific constants (our *site fixed effects*) jointly with the other parameters β and γ , much like the fixed effects estimator simultaneously estimates the fixed effects and the parameters in the linear regression model. We present results of a generated data experiment in the next section that supports this.

Step 3: $\{\varphi_i\}|\Xi_{-\varphi_i}, \tilde{\mathbf{U}}, \mathbf{y}$

$$\{\varphi_i\}|\Xi_{-\varphi_i}, \tilde{\mathbf{U}}, \mathbf{y} \sim \mathcal{N}(D_\varphi d_\varphi, D_\varphi) \quad (23)$$

where

$$D_\varphi^{-1} = JT + \frac{1}{\sigma_\varphi^2}; \text{ and } d_\varphi = \sum_{t=1}^T (\mathbf{U}_{i,t}^\varphi - \mathbf{M}_{i,t}^\varphi \boldsymbol{\theta}^\varphi)$$

and $\mathbf{U}_{i,t}^\varphi$, $\mathbf{M}_{i,t}^\varphi$, and $\boldsymbol{\theta}^\varphi$ are stacked over the sites j ($j = 1 \dots J$) and choice occasion for each individual without the stay at home equation. That is

$$\mathbf{M}_{i,t}^\varphi = [\mathbf{I}_J \quad \mathbf{P}_i] \quad \text{and} \quad \boldsymbol{\theta}^\varphi = [\boldsymbol{\alpha}' \quad \beta]'$$

Step 4: $\boldsymbol{\alpha}_0|\Xi_{-\boldsymbol{\alpha}_0}, \tilde{\mathbf{U}}, \mathbf{y}$

The remaining steps of our posterior simulator involve the sampling of parameters of the hierarchical priors. Once we condition on the α_j 's, the mean of the posterior conditional for $\boldsymbol{\alpha}_0$ is similar to Murdock's (2006) second stage linear regression of the fitted alternative specific constants on observed site attributes. Specifically:

$$\boldsymbol{\alpha}_0|\Xi_{-\boldsymbol{\alpha}_0}, \tilde{\mathbf{U}}, \mathbf{y} \sim \mathcal{N}(\mathbf{D}_{\alpha_0} \mathbf{d}_{\alpha_0}, \mathbf{D}_{\alpha_0}) \quad (24)$$

where

$$\mathbf{D}_{\alpha_0} = (\mathbf{Q}'\mathbf{Q}/\sigma_\alpha^2 + \mathbf{V}_\alpha^{-1})^{-1} \text{ and } \mathbf{d}_{\alpha_0} = \mathbf{Q}'\boldsymbol{\alpha}/\sigma_\alpha^2 + \mathbf{V}_\alpha^{-1}\boldsymbol{\mu}_\alpha.$$

Step 5: $\sigma_\alpha^2|\Xi_{-\sigma_\alpha^2}, \tilde{\mathbf{U}}, \mathbf{y}$

$$\sigma_\alpha^2|\Xi_{-\sigma_\alpha^2}, \tilde{\mathbf{U}}, \mathbf{y} \sim IG \left[\frac{J}{2} + a_\alpha, \left(b_\alpha^{-1} + .5 \sum_{j=1}^J (\alpha_j - \mathbf{Q}_j \boldsymbol{\alpha}_0)^2 \right)^{-1} \right]. \quad (25)$$

Step 6: $\sigma_\varphi^2|\Xi_{-\sigma_\varphi^2}, \tilde{\mathbf{U}}, \mathbf{y}$

$$\sigma_\varphi^2|\Xi_{-\sigma_\varphi^2}, \tilde{\mathbf{U}}, \mathbf{y} \sim IG \left[\frac{N}{2} + \alpha_\varphi, \left(b_\varphi^{-1} + .5 \sum_{i=1}^N \varphi_i^2 \right)^{-1} \right]. \quad (26)$$

4 Generated Data Experiment

In this section we conduct a series of four generated data experiments to illustrate the performance of our proposed methods. In particular we examine how our sampler performs given different degrees of correlation (a) between observed and unobserved site characteristics and (b) between travel cost and unobserved site characteristics. The experiments also shed light on the mixing properties of our simulations and are useful in surmising the accuracy of our code.

In the first pair of experiments we focus our attention on the impact that correlation between the observed and unobserved site characteristics has on parameter recovery. The sample size is set at $N = 5,000$, with each individual choosing from among $J = 30$ sites and the stay-at-home option over the course of $T = 20$ choice occasions. Two demographic variables (i.e., \mathbf{z}_i 's) are included in the experiments, with one drawn from a uniform distribution and the second a dummy variable drawn from a Bernoulli distribution with equal probability of success and failure. The travel cost for each individual to each site (p_{ij}) is drawn from a standard normal distribution. Finally, the alternative specific constants are generated assuming that

$$\alpha_j = \alpha_{01} + \alpha_{02}s_j^o + \alpha_0^u s_j^u \quad (27)$$

with s_j^o and s_j^u drawn jointly from the following bivariate normal distribution:

$$\begin{bmatrix} s_j^o \\ s_j^u \end{bmatrix} \sim \mathcal{N} \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_o^2 & \rho\sigma_o\sigma_u \\ \rho\sigma_o\sigma_u & \sigma_u^2 \end{pmatrix} \right]. \quad (28)$$

In our experiments we use $\sigma_o^2 = 0.03$, $\sigma_u^2 = 0.05$ and $\alpha_0^u = 1$. Since we assume only s_j^o is observed, (27) and (28) imply

$$\alpha_j | s_j^o \stackrel{ind}{\sim} \mathcal{N} \left(\alpha_{01} + \left[\alpha_{02} + \rho\alpha_0^u \left(\frac{\sigma_u}{\sigma_o} \right) \right] s_j^o, (\alpha_0^u)^2 \sigma_u^2 [1 - \rho^2] \right). \quad j = 1, 2, \dots, J. \quad (29)$$

In the notation of our hierarchical specification for the alternative specific constants [see equation (15)], $\mathbf{q}_j = [1 \ s_j^o]$, $\boldsymbol{\alpha}_0 = (\alpha_{01} \ [\alpha_{02} + \rho\alpha_0^u \frac{\sigma_u}{\sigma_o}])$, and $\sigma_\alpha = \alpha_0^u \sigma_u \sqrt{1 - \rho^2}$. Experiment 1 assumes that s_j^o and s_j^u are uncorrelated (i.e., $\rho = 0$), while experiment 2 assumes that they are correlated with $\rho = 0.7$. In this first experiment, therefore, we would expect accurate recovery of the structural parameter α_{02} as the second coefficient in $\boldsymbol{\alpha}_0$. This will not be the case when $\rho \neq 0$, as the second experiment will reveal.

We fix β , γ and the hierarchical parameters $\{\alpha_{01}, \alpha_{02}, \sigma_\varphi, \sigma_\alpha\}$ at the true values listed in column 3 of Table 1. The Gibbs sampler described in Section 3.3 is implemented for 100,000 iterations, discarding the first 40,000 draws as *burn-in*. The chain is initialized at values that are relatively far from the true parameters with σ_α^2 and σ_φ^2 set to unity and all remaining parameters set to zero.

Trace plots suggested relatively rapid convergence, with 40,000 burn-in simulations representing a conservative choice.¹⁷

Posterior means and posterior standard deviations for β, γ and the hierarchical parameters $\{\alpha_{01}, \alpha_{02}, \sigma_\varphi, \sigma_\alpha\}$ are reported in the first column of Table 1 for the first generated data experiment.¹⁸ A comparison of the first and third columns of this table reveals that the algorithm is performing well at recovering the parameters of the model. The only parameter that appears potentially troubling is the variance parameter σ_α^2 . We caution the reader, however, that our experiment only has $J = 30$, and with this small number of alternatives, the prior can be influential and it is difficult to accurately pin down the variance parameter at this level of the hierarchy. With $J = 30$, the marginal posterior distribution of the variance parameters displayed a significant right-skew; reassuringly, however, the mode of this posterior was quite close to the true value .05.

Although our point estimates are suggestive of good performance, any MCMC-based inference can be affected by the degree of correlation among the parameter draws over sequential iterations. If the degree of correlation is high, the algorithm will be slow in exploring all areas of the posterior surface and a large number of draws will be needed to fully characterize the posterior distribution. A summary of the impact of high autocorrelation on the precision of our posterior estimates can be obtained by calculating *numerical standard errors* (NSE) associated with the mean estimates:

$$\text{NSE}(\bar{\theta}_m) \equiv \sqrt{\frac{\sigma^2}{m}} \sqrt{1 + 2 \sum_{j=1}^{m-1} \left(1 - \frac{j}{m}\right) \rho_j}, \quad (30)$$

where θ represents an arbitrary scalar parameter of interest, m denotes the number of post-convergence simulations, $\bar{\theta}_m$ represents our estimate of $E(\theta|\mathbf{y})$ as the sample average of our post-convergence draws, ρ_j represents the correlation between simulations j periods (iterations) apart and $\sigma^2 \equiv \text{Var}(\theta|\mathbf{y})$. The above reveals that the numerical standard deviation of the posterior mean estimate equals the “traditional” standard error under iid sampling ($\sqrt{\sigma^2/m}$) times an inflation factor, often termed an *inefficiency factor*. Since the ρ_j are positive in practice, we see that the NSE’s can be larger - sometimes dramatically so - than the standard errors under iid sampling. How precise our mean estimate will be depends on the number of post-convergence simulations produced (m) as well as the mixing of those simulations, as summarized by the ρ_j .

We present NSE’s in the second column of Table 1 and do so for the first experiment only, as results from the remaining experiments are similar. As the results clearly illustrate, the NSE’s are extremely small relative to the mean estimates, strongly indicating that the simulation-based estimates are stable and accurately approximate the posterior means of this selection of parameters.¹⁹

¹⁷Trace plots for a selection of parameters are contained in a Technical Appendix to the paper available from the authors upon request.

¹⁸For the sake of space, the posterior means and standard deviations for the alternative specific constants are relegated to the Technical Appendix, available from the authors upon request.

¹⁹Importantly, the NSE’s should not be confused with the posterior standard deviations, the latter of which are

Parameter posterior means and standard deviations for the second generated data experiment are reported in the fourth column of Table 1. In this experiment there is substantial correlation ($\rho = 0.7$) between the observed and unobserved site attributes. As anticipated, based upon the result in (29), we find that the posterior mean of α_{02} is now much larger than the “true” α_{02} . In this case, the second coefficient associated with the vector \mathbf{q}_j no longer reduces to this parameter, with the analysis suffering what is akin to omitted variables bias found in a classical linear setting. Furthermore, the variance parameter σ_α^2 is also reduced relative to that of experiment 1, consistent with (29) when $\rho \neq 0$. Despite this important limitation, however, estimates of parameters appearing in the latent utility stage of the model in (8) are unaffected by the induced correlation between s_j^o and s_j^u , with posterior means that are all quite close to their true values.

In the final two experiments we examine the potential consequences of correlation between travel cost and the unobserved site attributes. Specifically, we assume that the travel cost (p_{ij}) is a weighted average of an independent random variable $\omega_{ij} \stackrel{iid}{\sim} N(0, 1)$ and the unobserved site characteristics (s_j^u) with:

$$p_{ij} = (1 - \kappa)\omega_{ij} + \kappa s_j^u, \quad 0 < \kappa < 1. \quad (31)$$

As in the previous two experiments, the unobserved site characteristic (s_j^u) is generated from a normal distribution with mean zero and variance $\sigma_u^2 = 0.05$. As κ increases, the correlation between the travel cost and unobserved site characteristic increases.²⁰ We consider two levels of correlation. In experiment 3 we use $\kappa = 0.5$ to induce a moderate level of correlation between the unobserved attributes and travel cost, yielding $Corr(p_{ij}, s_j^u) \approx 0.22$. In experiment 4 we create a higher level of correlation by setting $\kappa = 0.7$, yielding $Corr(p_{ij}, s_j^u) \approx 0.46$. For both experiments the simulator described in section 3.3 was used to obtain 100,000 draws from the posterior distribution, with 40,000 draws discarded as the burn-in. Parameter posterior means and standard deviations from both experiments are reported in Table 2, along with the true parameters used to generate the data.²¹

In both instances, and despite the correlation between the unobserved site attributes and travel cost, the simulator does a good job in recovering the underlying parameters of the model. The posterior mean of the key coefficient on the travel cost variable (β) differs from the true value in both experiments by less than one percent and the true value falls within 1 standard deviation of the mean in both experiments. We again emphasize that this is a special feature of our particular posterior simulator and such results would not be obtained had we blocked the parameters in a “traditional” way by first integrating out the α_j and then sampling (β, γ) . Indeed, when we

provided in parentheses in the first column of the table. NSE’s quantify how close the estimated mean is to the true posterior mean, based upon the MCMC output. The posterior standard deviations reflect uncertainty regarding the parameter’s themselves based on the data at hand and supplied prior information. NSE’s will converge to zero as m increases, while the posterior standard deviations will not.

²⁰In particular, $Corr(p_{ij}, s_j^u) = \kappa\sigma_u^2\{\sigma_u^2[(1 - \kappa)^2 + \kappa^2\sigma_u^2]\}^{-1/2}$.

²¹As with experiments 1 and 2, the results for the alternative specific constants are relegated to the Technical Appendix.

implement such a simulator to this data, we obtain a posterior mean and standard deviation of β equal to -5.56 and 0.09, which is clearly bounded away from the actual parameters of the DGP.

As a whole, results from the four experiments reported in this section attest to the fact that our sampling scheme works well and insulates the travel cost coefficient from potential biases resulting from unobserved site characteristics. In the absence of legitimate instruments and an elaborated structure on (15), however, estimated coefficients on site-level variables are not immune to problems caused by relevant omitted site characteristics that are correlated with elements of \mathbf{s}_j^o .

5 Application and Data Description

We apply our methods using data from the Iowa Lakes Valuation Project at Iowa State University. The Iowa Lakes Project is a four year panel data study, sponsored by the Iowa Department of Natural Resources and the US EPA, eliciting the visitation patterns of Iowan residents to the primary recreational lakes in the state. One of the objectives of the project is to measure the value residents place in the existing lake recreation opportunities and to predict the changes in welfare that would result from proposed water quality improvements. Understanding how water quality attributes affect recreational activities will help policy makers allocate limited environmental resources and prioritize their efforts to comply with the Clean Water Act.

The data set is appropriate for our study for a number of reasons. The Iowa Lakes Project not only covers all the major lakes in the state but also provides information on a wide variety of site characteristics. The observed site characteristics (\mathbf{Q}) include both site attributes, such as lake acreage and indicators for paved boat ramps and handicap accessibility, and an unusually large number of water quality attributes, such as Secchi Transparency (a measure of the depth of water clarity), Nitrogen, and Chlorophyll.²² In addition, the exact same data was used by Egan *et al.* [7] to estimate a repeated mixed logit model but *without* the inclusion of a full set of alternative specific constants. Our analysis provides an indication of how the estimated impact of travel cost is impacted by not controlling for unobserved site characteristics.

Although data for the project was collected over a four year period (2002-2005), we focus on the 2002 survey. The initial survey was sent by mail to 8,000 randomly selected Iowa residents. The response rate among deliverable surveys was 62%, yielding a total of 4,423 returned surveys. We exclude from our analysis those individuals who (a) were not Iowa residents (42), (b) failed to complete the section of the survey asking for lake visitation patterns (360), or (c) reported taking

²²The water quality attributes were measured by Iowa State University's Limnology Laboratory three times a year at each lake. The values used in our analysis are simple averages of these measures, following the approach used in Egan *et al.* [7].

more than fifty-two day trips per-year (223). The latter sample exclusion follows the procedure used in Egan *et al.* [7], wherein the authors note that individuals taking such frequent trips are usually local residents who are counting casual visits to or the passing by of their local lake. Instead, our analysis, like theirs, is concerned with day-trips taken to lake sites solely for the purpose of recreation.²³ The cut-off of fifty-two trips per year allows for a day-trip each week.

Table 3 provides summary statistics for our sample, both in terms of household demographics and individual site characteristics. As the table indicates, the survey respondents in our data set are, on average, older males with some college or trade/vocational school. The average household size is 2.45, including (on average) 0.61 children. Travel cost (p_{ij}) is calculated using 25 cents per mile for the round-trip travel distance [computed using *PCMiler (Streets Version 17)*] plus one-third the respondent’s wage rate multiplied by the travel time.²⁴ Overall, round-trip travel costs average just under \$140, ranging from less than \$1 to \$1366.

One of the appealing features of the Iowa Lakes Project is that, not only is there a wealth of information available regarding the site attributes and lake water quality, but there is also considerable variation across the lakes in terms of these characteristics. The lakes in the Iowa Lakes Project are, on average, 667 acres in size, ranging from 10 acres to approximately 19,000 acres. The other site attributes are represented with dummy variables that indicate the availability of amenities of interest. The majority of the lakes in our sample have a paved boat ramp (85%) and wake restrictions (i.e., $Wake = 1$) (65%), while less than forty percent of the lakes have handicap facilities or are part of a local state park. There is also a wide range of water quality in Iowa lakes. For example, Secchi Transparency (which measures the depth into the lake that one can see) averages just over one meter, but varies from less than 0.1 meters (approximately 3.5 inches) to 5.67 meters (well over 18 feet). Similar ranges are found for the other water quality measures, including Total Nitrogen, Total Phosphorus, and Cyanobacteria. Moreover, these water quality measures are not highly correlated, as the source and nature of the water quality problems in individual lakes varies considerably across the state.

6 Empirical Results

We fit our random utility maximization model using the posterior simulator of Section 3.3. As described there, Gibbs sampling is used to generate simulations from the joint posterior distribution and the Gibbs algorithm is first run for 20,000 iterations. The last iteration from this process is

²³Egan *et al.* [7] also found that their qualitative results are not sensitive to the specific cut-off of fifty-two trips per year.

²⁴The “average wage rate” is calculated for all respondents as their household’s income divided by 2,000. This allows for a 40 hour work week with two weeks of vacation.

then used to initialize four different chains, run simultaneously on four different machines with different seeds. Each of these runs produced 47,500 simulations, leaving us with a total of 190,000 post-convergence draws to calculate parameter posterior means, standard deviations and other quantities of interest.

Each iteration of the simulator was found to take about 30 seconds to run (or about 10 hours to produce 1,000 iterations), with the most time consuming step being the simulation of latent data for each agent, for each of 130 different sites, and for $T = 52$ different choice occasions. Specifically, this step demands the simulation of nearly 21,400,000 truncated normal random variables, which cannot be completely vectorized or done within a single coding step, since the simulations just produced are used to update the conditioning information in the sampling of the remaining latent data. A large burn-in value of 20,000 iterations was used because, informed by our generated data experiments, we found the simulator for the parameters α'_j s, β and γ to mix relatively slowly. Although we would like to obtain more post-convergence simulations in light of the fact that the algorithm tends to exhibit slow mixing, generating additional draws is quite costly - our two machines worked for nearly 20 days to provide the set of simulations used here.

6.1 Application Results

We are primarily interested in applying the algorithm in Section 3.3 to data from Iowa Lakes Project to address the following questions: (1) Does ignoring unobserved site characteristics matter in understanding site visitation patterns of Iowans? (2) Which site attributes are important in influencing site visitation patterns? and (3) What are the welfare implications of water quality improvements?

We begin to address these questions by first considering our most general model (later denoted as Model “A”) which contains a complete set of water quality measures and non-water quality site attributes in \mathbf{q}_j . We report parameter posterior means and posterior probabilities of being positive [denoted $P(\cdot > 0|y)$] for key parameters of the model in Tables 4 and 5. For the sake of brevity we only report (in Table 4) the alternative specific constants results for an illustrative subset of sites.²⁵

The alternative specific constants α_j are negative for all the 130 sites in our sample with less than 0.0001 percent of the posterior mass density in the positive region of the parameter space. The large negative values for the site specific constants is consistent with our data as these constants reflect the difference in utility from site visits relative to staying at home. Given the large number of trip options, a large negative alternative specific constant is needed to reflect the fact that approximately 39% of our sample reported visiting none of the lakes in 2002. As expected, the α'_j s

²⁵The complete set of estimated alternative specific constants are provided in the Technical Appendix to the paper.

are higher (less negative) for popular destination lakes in the state, such as West Lake Okoboji, which has one of the highest quality site attributes in the U.S.

Within the context of the standard RUM framework used here, the negative of the travel cost coefficient (i.e., $-\beta$) can be interpreted as the marginal utility of income, which is assumed to be constant.²⁶ Consistent with our expectations, the posterior mean for β in Table 4 is negative (-0.0146), with the vast majority of the posterior mass falling in the negative region. Posterior distributions for parameters associated with the demographic variables are also consistent with the literature. Similar to Egan *et al.* [7], older individuals, females, and the less educated are more likely to stay at home. Households with a greater number of adults and children have a higher likelihood of visiting a site. However, in contrast to [7], the signs and overall magnitudes of these coefficients in our analysis are not sensitive to the construction of \mathbf{q}_j . This, we argue is a key benefit of our hierarchical model structure and posterior simulator; only estimates at the terminal stage of the hierarchy are susceptible to having limited information on the attributes. All of the posterior densities for the demographic variables coefficients are highly massed on either the positive or negative side of the distribution, showing a consistent impact on the decision to stay at home.

The results for the parameters of the hierarchical prior are presented in Table 5. The posterior mean of the “overall intercept” α_{01} is negative, as expected, and the posterior density also places the majority of its mass over negative values. Similar to Egan *et al.* [7], the availability of amenities such as boat ramps, handicap accessibility, wake restrictions and classification as a state park make a site more attractive. Larger lakes, *ceteris paribus*, are also preferred, as all posterior simulations associated with the site size (acres) parameter were positive.

The posterior results for the water quality attributes, however, are in sharp contrast to earlier studies. In particular, whereas Egan *et al.* [7], using the same data, conclude that water quality attributes are important determinants of recreational lake usage and site selection, we find little evidence of this effect. For most of the water quality coefficients, the mass of their posterior densities are more or less evenly divided between the positive and negative values [with $P(\cdot > 0|y)$ hovering around 0.5]. This result may call into question the individual importance of water quality attributes in determining recreation demand.²⁷ Only Total Phosphorus, which contributes to algae growth, is convincingly massed on one side of zero in our analysis, suggesting (as in [7]) that high levels of Total Phosphorus reduces the appeal of a site.

²⁶While it is standard in the recreation demand literature to assume a constant marginal utility of income, it is not required for the methods outlined in this paper.

²⁷It is important to emphasize, however, that by incorporating alternative specific constants into our model, the impact of site specific attributes on recreation demand is being captured entirely by the variation in the α_j 's. In essence, we have only $J - 1$ observations in modeling the impact of water quality on recreation demand. Thus, it is not surprising that the resulting posterior distributions for the water quality coefficients are relatively diffuse, with the data providing relatively little information by which update our diffuse priors on these parameters.

Table 6 represents our attempt to reproduce results in Egan *et al.* [7] from the Bayesian perspective. Here, we consider a simplified representation of latent utility, as in equation (2), that ignores unobserved site characteristics by dropping the site-specific constants. When performing this analysis, we obtain results that are similar to those in [7] with water quality attributes shown to have important influences on site selection. In particular, the posterior probabilities of being positive for these characteristics are, with the exceptions of Cyanobacteria and Volatile SS, virtually one or zero, indicating a clear role for these characteristics in recreation decisions. Also, comparing the coefficient on travel cost between the two results, we see a drop of about 20% in the estimated impact of travel cost when we allow for site-specific constants. This translates into an approximate 25% increase in the welfare effect.

6.2 Model Comparison

Our empirical results provide little evidence that individual water quality attributes impact the visitation patterns to Iowa lakes. Correlation among these characteristics, however, may make it difficult to isolate the role of any single attribute. In this section, we investigate the *joint* impact of site characteristics and water quality attributes in influencing recreational lake usage. Specifically, we consider four different models, similar to those employed by Egan *et al.* [7], and with an eye toward determining the set of factors that plays the largest role in explaining recreation decisions. These models are enumerated below:

- **Model A:** The unrestricted model just discussed containing a full set of water quality characteristics and site attributes.
- **Model B:** Includes Secchi Transparency as the only water quality characteristic, with all site attributes included.
- **Model C:** No water quality characteristics included, with all site attributes included.
- **Model D:** Includes Secchi transparency as the sole water quality characteristic, with all site attributes excluded.

Comparison among these competing models is based upon the posterior probability that a model is supported by the data. By Bayes rule, the posterior probability of model \mathcal{M}_k is given by:

$$p(\mathcal{M}_k|data) = \frac{p(data|\mathcal{M}_k)p(\mathcal{M}_k)}{p(data)} \quad k = 1, \dots, K, \quad (32)$$

where $p(data|\mathcal{M}_k)$ denotes the *marginal likelihood*, $p(\mathcal{M}_k)$ is the prior probability of Model k and $p(\mathcal{M}_k|data)$ is the posterior probability of \mathcal{M}_k . Therefore, models can be compared pairwise based

on their *posterior odds ratio* which is defined as:

$$PO_{kj} = \frac{p(\mathcal{M}_k|data)}{p(\mathcal{M}_j|data)} = \frac{p(data|\mathcal{M}_k)p(\mathcal{M}_k)}{p(data|\mathcal{M}_j)p(\mathcal{M}_j)}. \quad (33)$$

In practice, the prior odds ratio $p(\mathcal{M}_k)/p(\mathcal{M}_j)$ is usually set to unity for all the possible models considered so that:

$$PO_{kj} = \frac{p(\mathcal{M}_k|data)}{p(\mathcal{M}_j|data)} = \frac{p(data|\mathcal{M}_k)}{p(data|\mathcal{M}_j)} \equiv BF_{kj}, \quad (34)$$

with the ratio of marginal likelihoods denoted as the *Bayes factor* (BF).

For nested model comparison exercises, like those involved in deciding among Models A-D above, the *Savage-Dickey (S-D) density ratio* offers a useful computational expedient for the calculation of (34). Specifically, suppose we wish to “test” $\pi = 0$ for some subvector of coefficients π . Provided the restricted model’s prior for parameters other than π is the same as the unrestricted model’s prior for these parameters *given that* $\pi = 0$, we can write

$$BF_{12} = \frac{p(\pi = 0|data, \mathcal{M}_2)}{p(\pi = 0|\mathcal{M}_2)}, \quad (35)$$

where Model 1 in the above represents the restricted version of Model 2, imposing $\pi = 0$. The two expressions $p(\pi = 0|data, \mathcal{M}_2)$ and $p(\pi = 0|\mathcal{M}_2)$ are recognized as the posterior and prior ordinates at zero under the unrestricted model 2, respectively, and the former of these can be readily calculated given output from the posterior simulator.

Table 7 provides the resulting Bayes factors (BF_{ij}). The first row of Bayes factors uses Model A as the “unrestricted” model (j), comparing it sequentially to models B, C and D as restricted alternatives (i). In each of these cases, the Bayes factors exceed 1.0E+11, providing little support for a model including the full range of water quality attributes. Indeed, Model C, which includes none of the water quality characteristics, clearly dominates Model A. The second row of Bayes factors in Table 7 uses Model B as the baseline model (with only Secchi Transparency as a measure of lake water quality). Again, Model C dominates, with its posterior probability being almost 30 times higher than Model B. Interestingly, Model B clearly dominates Model D (in which site attributes such as wake restrictions and lake size are excluded from the model). These findings confirm the general conclusions drawn from examining the individual parameter estimates; i.e., that site attributes do significantly influence the pattern of recreational lake usage, but that there is little evidence indicating that water quality characteristics play a direct role in lake visitation.

7 Welfare Analysis

Policy and counterfactual analysis is an important part of recreation demand research. In this section we briefly describe how our model can be used to evaluate policies that affect site charac-

teristics. Specifically, we consider changes to site attributes from baseline levels (\mathbf{Q}^0) to alternative conditions (\mathbf{Q}^1). Limiting versions of this exercise can also be used to examine welfare consequences associated with changes to the available choice set (e.g., the loss of an entire site).

Let Υ_{it}^s denote the maximal utility achieved by agent i on choice occasion t under scenario s ($s = 0, 1$). That is,

$$\Upsilon_{it}^s(\Xi, \mathbf{Q}^s) = \max_j (U_{ijt}^s | \Xi, \mathbf{Q}^s) \quad s = 0, 1, \quad (36)$$

where Ξ denotes all parameters of the model.

The compensating variation (CV) is then defined as the monetized change in expected maximum utility due to changes in site attributes over the course of the season. That is

$$CV_i(\mathbf{Q}^1, \mathbf{Q}^0) = T \left(E_{\Xi|y} \left(-\frac{1}{\beta} [\Upsilon_i^1(\Xi, \mathbf{Q}^1) - \Upsilon_i^0(\Xi, \mathbf{Q}^0)] \right) \right). \quad (37)$$

The term in square brackets above measures the change in maximal utility per choice occasion which, for us, does not depend on t and is identical in expectation for all t . As such, we drop t in our notation for Υ_{it}^s in (37) and will do so throughout the remainder of this discussion. The seasonal change in expected utility is then just a simple scaled multiple of the expected change per choice occasion [hence the appearance of “ T ” in (37)]. Finally, dividing by the marginal utility of income ($-\beta$) creates a monetized measure of the change in the consumer’s utility from a change in site characteristics.

We estimate the compensating variation in (37) by simulating utility values conditional on the posterior distribution of the parameters for the two scenarios. The algorithm for welfare analysis can be described in the following steps:

Step 1: Let $\Xi_{-\alpha}^{(r)}$ ($r = 1, \dots, R$), denote a draw from the posterior distribution of $\Xi_{-\alpha}$. Draw $\alpha_j^{s(r)}$ ($j = 1, \dots, J; s = 0, 1$) using (15). That is, draw the alternative specific constant for site j under scenario s ($\alpha_j^{s(r)}$) from a normal distribution with mean $\mathbf{q}_j^s \alpha_0^{(r)}$ and variance $(\sigma_\alpha^{(r)})^2$.

Step 2: Draws of the utility levels $U_{ij}^{s(r)} | \Xi, \mathbf{Q}^s$ are obtained using equation (8), with

$$U_{ij}^{s(r)} = \begin{cases} \mathbf{z}_i \gamma^{(r)} + \varepsilon_{i0}^{(r)} & j = 0 \\ \alpha_j^{s(r)} + p_{ij} \beta^{(r)} + \varphi_i^{(r)} + \varepsilon_{ij}^{(r)} & j = 1, \dots, J. \end{cases} \quad (38)$$

where

$$\varepsilon_{ij}^{(r)} \sim \mathcal{N}(0, 1) \quad \text{and} \quad \varphi_i^{(r)} \sim \mathcal{N}(0, \sigma_{\varphi}^2(r)). \quad (39)$$

Step 3: Define

$$\Upsilon_i^{s(r)} = \max\{U_{ij}^{s(r)}\}, \quad s = 0, 1, \quad r = 1, 2, \dots, R. \quad (40)$$

The simulation based estimate of the welfare effect is then computed using

$$\widehat{CV}(\mathbf{Q}^1, \mathbf{Q}^0) = \frac{1}{N} \sum_{i=1}^N \widehat{CV}_i(\mathbf{Q}^1, \mathbf{Q}^0) = \frac{1}{N} \sum_{i=1}^N \left\{ \frac{1}{R} \sum_{r=1}^R \frac{T}{-\beta^{(r)}} [\Upsilon_i^{1(r)} - \Upsilon_i^{0(r)}] \right\}, \quad (41)$$

an average of the CV's for every individual in the sample. In performing these calculations, we are treating the Iowa Lakes sample as representative of the general Iowa population in terms of individual characteristics and their proximity to the lakes in the state (which in turn determines their travel cost). If true, then the parameter in (41) estimates the average annual dollar amount that an Iowa resident receives as compensation / would pay as income for the proposed change in site characteristics. Note that this is very different than an effect for an ‘‘average’’ individual, and our approach also appropriately accounts for parameter uncertainty in the process.

To illustrate these welfare calculations, we consider two competing policy scenarios:

- Scenario 1: The closure of West Lake Okoboji, a popular destination lake in Iowa and one of the cleanest lakes in the US.
- Scenario 2: Improving nine *zonal* lakes (listed in Table 4) to have the same water quality as West Lake Okoboji. These nine lakes, chosen as potential *target* lakes by the Iowa Department of Natural Resources, are spaced evenly throughout the state.

For the first of these policy scenarios, we find the posterior mean compensating variation to be negative, as expected, with an annual loss per household of $\widehat{CV} = -\$9.52$. Moreover, the sign of the associated compensating variation is clear, with $CV^{(r)} < 0 \quad \forall r$ [i.e., $P(CV > 0|y) = 0$]. In contrast, the welfare implications of scenario 2 are less clear. Indeed, while one might expect the water quality improvements to be welfare enhancing, the posterior mean compensating variation is negative, indicating an annual loss per household of $\widehat{CV} = -\$2.20$. However, the sign of the welfare change itself is unclear, with the individual $CV^{(r)}$'s relatively evenly distributed between positive and negative values, with $P(CV > 0|y) = 0.35$. These findings are consistent with the lack of clear evidence linking water quality attributes to recreational lake usage.

8 Summary

Controlling for unobserved site characteristics is important in modeling recreation demand as researchers are typically restricted to only a small subset of the site attributes that influence site selection. To the extent that site specific factors are omitted from the analysis and correlated with either observed site attributes or the travel cost variable, the resulting parameter estimates and

subsequent welfare analysis will be contaminated. In a recent article in this journal, Murdock [21] provided a resolution to this problem in the context of the RUM model, introducing a full set of alternative specific constants to absorb the impact of the unobserved site attributes and proposing a two-stage estimation procedure, combining a contraction mapping algorithm and maximum likelihood estimation. While this solution is an elegant one, it has its limitations. Perhaps most importantly, it cannot be used in the context of mixed logit models that are frequently employed in the recreation demand literature to capture preference heterogeneity. The purpose of this paper has been to introduce an alternative approach to estimation of a RUM model with a full set of alternative specific constants, one that does allow for the inclusion of random parameters. Specifically, we employ a Bayesian framework, combining Gibbs sampling and data augmentation techniques to characterize the posterior distribution of the parameters of interest. A hierarchical model is used to allow observed site attributes to influence the alternative specific constants, while the blocking structure used in the Gibbs sampler ensures that the key travel cost parameter is insulated from effects of correlation between unobserved site attributes and either the observed site attributes or the travel cost variable itself. A series of generated data experiments illustrate this result. While we do not pursue the option within the current paper, the model can be further extended to explicitly control for correlation between the observed and unobserved site attributes using a form of instrumental variables.

Our application, using data from the Iowa Lakes Project, illustrates both the methodology itself and the importance of controlling for unobserved site attributes when modeling recreation demand. Whereas Egan *et al.* [7], using the same database, find that water quality significantly impacts the pattern of recreation demand, little evidence to support this conclusion emerges once we allow for the potential impact of unobserved site attributes.²⁸ We do, however, continue to find that site attributes (e.g., wake restrictions, boat ramps, lake size, etc.) significantly impact where households choose to recreate.

References

- [1] Berry, S. (1994). Estimating Discrete-Choice Models of Product Differentiation, *RAND Journal of Economics*, 25 (2), 242-262.

²⁸This does not, of course, prove that water quality characteristics do not influence site selection. First of all, by relegating the site characteristics to the hierarchical portion of the model, we essentially have only 129 observations (i.e., the α_j 's) to use in identifying the effects of water quality on site selection. Second, to the extent that unobserved site attributes are correlated with the observed site attributes, the posterior distributions for the hierarchical parameters will be contaminated by this correlation, reflecting the impact of both the observed and unobserved site characteristics on the alternative specific constants. Given the extensive nature of our water quality data, this is less likely to be a problem in the current application, but certainly cannot be ruled out.

- [2] Berry, S.,J. Levinsohn and A. Pakes (1995). Automobile Prices in market Equilibrium, *Econometrica*, 63 (4), 841-890.
- [3] Bockstael, N.E., W.M. Hanemann, and I.E. Strand (1986) *Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models*, report presented to the Environmental Protection Agency under cooperative agreement CR-811043-01-0, Washington, D.C.
- [4] Casella, George and Edward I. George (1992). Explaining the Gibbs Sampler, *The American Statistician*, 46 (3) (August), 167 - 174.
- [5] Chen, H.Z., F. Lupi, and J.P. Hoehn (1999). An Empirical Assessment of Multinomial Probit and Logit Models for Recreation Demand, in: *Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice* (ed. Herriges, J.A., and C.L. Kling), Cheltenham, UK: Edward Elgar, pp. 141-161.
- [6] Chib, S and B.P. Carlin (1999). On MCMC sampling in hierarchical longitudinal models, *Statistics and Computing*, (9), 17-26.
- [7] Egan, K.J., J.A. Herriges, C.L. Kling and J.A Downing (2009). Valuing Water Quality as a Function of Water Quality Measures, *American Journal of Agricultural Economics.*, 91 (1) (Feb.), 106-123.
- [8] Freeman, A.M (1979). *The Benefits of Air and Water Pollution Control: A Review and Synthesis of Recent Estimates*, Report Prepared for the Council on Environmental Quality, Bowdoin College, Brunswick, maine, December.
- [9] Geweke, J. (1991). Efficient simulation from the multivariate Normal and Student-t distributions subject to linear constraints, in: *Computer Science and Statistics: Proceedings of the Twenty-Third Symposium on the Interface* (ed. E. Keramidas), Interface Foundation of North America, Inc., Fairfax, 571-578.
- [10] Hanemann, W.M (1981). *Water Quality and the Demand for Recreation*, Department of Agricultural & Resource Economics, UCB. CUDARE Working Paper 164, March.
- [11] Herriges, J., and D. Phaneuf (2002). Inducing Patterns Correlation and Substitution in Repeated Logit Model of Recreation Demand, *American Journal of Agricultural Economics*, 84, 1076-1090.
- [12] Herriges, J.A, C.L. Kling and D. 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. Herriges, J.A., and C.L. Kling), Cheltenham, UK: Edward Elgar, pp. 141-161. 163-197.
- [13] Jiang, R., P. Manchanda and P.E. Rossi (2009). Bayesian analysis of random coefficient logit models using aggregate data, *Journal of Econometrics*, 149, 136-148.

- [14] Klaiber, H.A. and R. H. von Haefen (2008). Incorporating Random Coefficients and Alternative Specific Constants into Discrete Choice Models: Implications for In-Sample Fit and Welfare Estimates. Working paper, March.
- [15] Koop, Gary (2003). *Bayesian Econometrics*, Inc NetLibrary, NetLibrary, Inc, Published by J. Wiley.
- [16] Koop, G., D.J. Poirier and J.L. Tobias (2007). *Bayesian Econometric Methods*, Cambridge, Cambridge University Press.
- [17] Lancaster, Tony (2004). *An Introduction to Modern Bayesian Econometrics*, Published by Blackwell Publishing Ltd, pp 294-295.
- [18] Lindley, D. and A. F. M. Smith (1972). Bayes estimates for the linear model, *Journal of the Royal Statistical Society*, Series B, 34, 1-41.
- [19] Madigan, D.M. and Raftery, A.E. (1994). Model selection and accounting for model uncertainty in graphical models using Occam's Window. *Journal of the American Statistical Association*, 89, 1335-1346.
- [20] Morey, E.R., R.D. Rowe, and M. Watson (1993). A Repeated Nested-Logit Model of Atlantic Salmon Fishing, *American Journal of Agricultural Economics*, 75 (August), 578-592.
- [21] Murdock, Jennifer (2006). Handling unobserved site characteristics in random utility models of recreation demand, *Journal of Environmental Economics and Management*, 51, 1-25.
- [22] Parsons, G.R., E.C. Helm, and T. Bondelid (2003). *Measuring the Economic Benefits of Water Quality Improvements to Recreational Users in Six Northeastern States: An Application to the Random Utility Maximization Model*. Working Paper, University of Delaware, July.
- [23] Phaneuf, D.J., C.L. Kling, and J.A. Herriges (2000). Estimation and Welfare Calculations in a Generalized Corner Solution Model with an Application to Recreation Demand, *The Review of Economics and Statistics*, 82 (1) (February), 83-92.
- [24] Rossi, P.E., G. M. Allenby, and R. McCulloch (2005). *Bayesian Statistics and Marketing*, John Wiley & Sons Ltd.
- [25] Tihansky, D.P. (1975). A Survey of Empirical Studies, in: *Cost Benefit Analysis and Water Pollution Policy* (ed. H.R. Peskin and E.P. Sesking), 127-144.
- [26] Train, K. (2009). *Discrete Choice Methods with Simulation*, Cambridge University Press, 2nd edition, 2009.
- [27] Yang, S., Y. Chen, and G. Allenby (2003). Bayesian Analysis of Simultaneous Supply and Demand, *Quantitative Marketing and Economics*, (1), 251-275.

Table 1: Generated Data Experiments 1 and 2^a

Parameter	Experiment 1: Exogenous Observable Site Attributes ($\rho = 0$)			Experiment 2: Endogenous Observable Site Attributes ($\rho = 0.7$)	
	Posterior Mean	NSE	True Value	Posterior Mean	True Value
$\alpha_0(1)$	-3.4318 (0.0720)	0.0142	-3.52	-3.4777 (0.0714)	-3.52
$\alpha_0(2)$	0.9301 (0.2864)	0.0012	0.98	1.5870 (0.2768)	0.98
β_0	-4.4803 (0.0214)	0.0030	-4.50	-4.4838 (0.0212)	-4.50
$\gamma_0(1)$	0.9159 (0.3150)	0.0162	0.96	0.8920 (0.3136)	0.96
$\gamma_0(2)$	0.7680 (0.2039)	0.0162	0.75	0.7525 (0.2016)	0.75
σ_φ^2	0.4374 (0.039)	0.0084	0.40	0.4323 (0.0412)	0.40
σ_α^2	0.0898 (0.0219)	0.0001	0.05	0.0768 (0.0188)	0.05

^aPosterior standard deviation in parenthesesTable 2: Generated Data Experiments 3 and 4^a

Parameter	Experiment 3: Moderate Correlation Between Price and Unobservables ($\kappa = 0.5$)		Experiment 4: High Correlation Between Price and Unobservables ($\kappa = 0.7$)	
	Posterior Mean	True Value	Posterior Mean	True Value
$\alpha_0(1)$	-3.5347 (0.0573)	-3.52	-3.5512 (0.0573)	-3.52
$\alpha_0(2)$	0.9225 (0.2866)	0.98	0.8938 (0.2909)	0.98
β_0	-4.5027 (0.0164)	-4.50	-4.5262 (0.0227)	-4.50
$\gamma_0(1)$.9104 (0.0339)	0.96	0.8941 (0.0364)	0.96
$\gamma_0(2)$	0.7961 (0.0194)	0.75	0.7857 (0.0207)	0.75
σ_φ^2	0.4094 (0.0130)	0.40	0.41 (0.01)	0.40
σ_α^2	0.0890 (0.0218)	0.05	0.0920 (0.0226)	0.05

^aPosterior standard deviation in parentheses

Table 3: Summary Statistics

Variable	Model Variable	Mean	Std. Dev.	Min	Max
Total Day Trips (2002) ^a	T_i	6.3343	9.9717	0	50
Travel Cost (\$100's)	P_{ij}	1.37	.83	0.0044	13.66
Age ^b	$D_{i(1)}$	4.85	1.42	0	7
Gender (Male=1, Female=2)	$D_{i(2)}$	1.28	0.50	0	2
Education ^c	$D_{i(3)}$	3.00	1.18	0	5
Adults (No. of adults in household)	$D_{i(4)}$	1.84	0.71	0	6
Child (No. of children in household)	$D_{i(4)}$	0.61	1.04	0	7
Lake Attributes					
Acres	$Q_{j(2)}$	667.20	2112.83	10	19000
Ramps	$Q_{j(3)}$	0.85	0.36	0	1
Wake	$Q_{j(4)}$	0.65	0.48	0	1
Handicap	$Q_{j(5)}$	0.38	0.49	0	1
State Park	$Q_{j(6)}$	0.39	0.49	0	1
Water Quality					
Secchi Transparency (m)	$Q_{j(7)}$	1.17	0.92	0.09	5.67
Total Nitrogen (mg/l)	$Q_{j(8)}$	2.19	2.53	0.55	13.37
Total Phosphorus ($\mu\text{g/l}$)	$Q_{j(9)}$	105.45	80.33	17.10	452.55
Volatile SS (mg/l)	$Q_{j(10)}$	9.30	7.98	0.25	49.87
Inorganic SS (mg/l)	$Q_{j(11)}$	10.12	17.79	0.57	177.60
Cyanobacteria (mg/l)	$Q_{j(12)}$	298.08	831.51	0.02	7178.13
Chlorophyll ($\mu\text{g/l}$)	$Q_{j(13)}$	40.64	38.01	2.45	182.92

^a39% of the sample did not visit any lake that year

^bUnsure = 0; Under 18=1, 18-25=2, 26-34=3, 35-49=4, 50-59=5, 60-75=6, 76+=7

^cUnsure = 0; Some high school or less=1, High school graduate=2, Some college or trade/vocational school=3, College graduate=4, Advanced degree=5

Table 4: Posterior Means of Selected Alternative Specific Constants, ²⁹along with Travel Cost and Demographic Parameters

Parameter	Parameter							
	Model A		Model B		Model C		Model D	
	Mean	$P(\cdot > 0 y)$	Mean	$P(\cdot > 0 y)$	Mean	$P(\cdot > 0 y)$	Mean	$P(\cdot > 0 y)$
α_j	Selected Alternative Specific Constants							
Storm Lake	-2.8665	0.0000	-3.3376	0.0000	-3.3412	0.0000	-3.3308	0.0000
Briggs Woods Lake	-3.6978	0.0000	-3.6683	0.0000	-3.6654	0.0000	-3.6571	0.0000
Silver Lake	-4.0309	0.0000	-4.0154	0.0000	-4.0115	0.0000	-3.9830	0.0000
Prairie Rose Lake	-3.5275	0.0000	-3.5012	0.0000	-3.5127	0.0000	-3.5014	0.0000
Big Creek Lake	-3.0382	0.0000	-2.9928	0.0000	-2.9963	0.0000	-2.9845	0.0000
Lake McBride	-3.1943	0.0000	-3.1318	0.0000	-3.1304	0.0000	-3.1197	0.0000
Lake Icaria	-3.2517	0.0000	-3.2506	0.0000	-3.2562	0.0000	-3.2455	0.0000
Lake Darling	-3.4999	0.0000	-3.4478	0.0000	-3.4572	0.0000	-3.4431	0.0000
Rathbun Roservoir Lake	-2.8110	0.0000	-2.8024	0.0000	-2.8095	0.0000	-2.7988	0.0000
W. Okoboji Lake	-2.2968	0.0000	-2.335	0.0000	-2.3473	0.0000	-2.3253	0.0000
	Other Parameters							
β (Travel cost)	-0.0146	0.0000	-0.0132	0.0000	-0.0131	0.0000	-0.0132	0.0000
γ_{01} (Age)	0.1219	1.0000	0.1085	1.0000	0.1078	1.0000	0.1115	1.0000
γ_{02} (Gender)	0.1262	0.9826	0.1227	0.9978	0.1155	0.9926	0.1168	0.9971
γ_{03} (Education)	-0.1813	0.0000	-0.1627	0.0000	-0.1592	0.0000	-0.1616	0.0000
γ_{04} (Adults)	-0.1025	0.0000	-0.0929	0.0000	-0.0921	0.0000	-0.0935	0.0000
γ_{05} (Child)	-0.0046	0.0453	-0.0042	0.0409	-0.0041	0.0414	-0.0042	0.0489

²⁹We only report the result for the nine major lakes in each zone and Okoboji. Estimates for the other sites are presented in the technical appendix.

Table 5: Posterior Means of Hierarchical Parameters

Parameter	Parameter							
	Model A ³⁰		Model B		Model C		Model D	
	Mean	$P(\cdot > 0 y)$	Mean	$P(\cdot > 0 y)$	Mean	$P(\cdot > 0 y)$	Mean	$P(\cdot > 0 y)$
α_{01}	-4.1487 (2.1E-112)	0.0000	-4.1090	0.0000	-4.1151	0.0000	-3.7124	0.0000
Lake Attributes								
α_{02} (Acres)	0.0001 (18.2914)	0.9999	0.0001	1.0000	0.0001	1.0000		
α_{03} (Ramps)	0.2616 (0.3309)	0.9956	0.2085	0.9885	0.209	0.9888		
α_{04} (Wake)	0.1261 (2.9286)	0.9601	0.1026	0.9331	0.098	0.9261		
α_{05} (Handicap)	0.1738 (0.7630)	0.9914	0.1660	0.9920	0.162	0.9907		
α_{06} (State Park)	0.2118 (0.2586)	0.9975	0.2382	0.9996	0.236	0.9998		
Water Quality								
α_{07} (Sechi)	-0.0129 (21.8717)	0.3851	0.0057	0.5675			0.0266	0.7365
α_{08} (Total Nitrogen)	0.0126 (47.9296)	0.8305						
α_{09} (Total Phosphorus)	-0.0012 (259.1956)	0.0304						
α_{010} (Volatile SS)	-0.0002 (148.8793)	0.4902						
α_{011} (Inorganic SS)	0.0030 (191.5077)	0.8940						
α_{012} (Cyanobacteria)	2.11E-06 (24273.222)	0.5203						
α_{013} (Chlorophyll)	0.0011 (542.8128)	0.7933						
Variance								
σ_{α}^2	0.1283	1.0000	0.1198	1.0000	0.1183	1.0000	0.1935	1.0000
σ_{φ}^2	3.0359	1.0000	1.8895	1.0000	1.8850	1.0000	1.9699	1.0000

³⁰The values in parentheses are an approximation of the difference in Schwarz criterion (or BIC) between an unrestricted model k and a restricted model j using derivations from the S-D density ratio. Higher positive values represents grades of evidence for model j .

Table 6: Posterior Means without controlling for unobserved site characteristics

Parameter	Mean	$P(\cdot > 0 y)$
Lake Attributes		
Acres	4.07E-05	1.000
Ramp	0.264	1.000
Wake	0.160	1.000
Handicap	0.149	1.000
State Park	0.381	1.000
Water Quality		
Secchi	0.031	1.000
Total Nitrogen	0.002	0.922
Total Phosphorus	-0.001	0.000
Volatile SS	-0.001	0.250
Inorganic SS	0.002	1.000
Cyanobacteria	-4.32E-06	0.207
Chlorophyll	-0.001	0.000
Other Parameters		
Travel cost	-0.017	0.000
Age	0.668	1.000
Gender	-1.211	0.001
Education	1.515	1.000
Adults	0.102	0.697
Child	0.435	0.997
σ_φ^2	1560.70	1.000

Table 7: Model Comparison

	Bayes Factor (BF)			
	Model A	Model B	Model C	Model D
“Unrestricted Model”		H_0	H_0	H_0
Model A		1.63E+16 (74.67) ^a	4.71E+17 (81.39)	2.03E+11 (52.07)
Model B			29.21 (6.75)	1.80E-15 (-21.85)

^a The values in parentheses are $2 \log_e BF$ using S-D density ratio.

Table 8: Annual Compensating Variation Estimates

	CV (\$)	$P(\cdot > 0 y)$
Close W. Okoboji	-9.5207	0
Upgrade 9 Zone lakes to W. Okoboji	-2.1988	0.3454