2010

Bayesian inference in modeling recreation demand

Babatunde Oluwakayode Abidoye
Iowa State University

Bayesian inference in modeling recreation demand

by

Babatunde Abidoye

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:
Joseph A. Herriges, Co-major Professor
Justin L. Tobias, Co-major Professor
Catherine L. Kling
Brent Kreider
Michael Larsen

Iowa State University

Ames, Iowa

2010

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to dedicate this thesis to God, my wife Joanna and my son Joel without whose support I would not have been able to complete this work. I would also like to thank my family especially my parents, Daniel and Rachel, for their love and instilling in me the love of learning.

Special thanks to my major professors, Joe Herriges and Justin Tobias for excellent guidance in the writing of this dissertation. Also, the contribution of Cathy Kling and the Resource and Environmental Group at ISU to my research can not go without mention. I would also like to thank my committee members - Dr. Brent Kreider and Michael Larsen for their efforts and contributions to this work.

Finally, I thank all those who helped me with various aspects of conducting research and the writing of this thesis.

Introduction

Overview

McFaddens (1981) hypothesis of random utility maximization (RUM) has been a workhorse for researchers in different fields of economics to study household choices among competing goods. The RUM hypothesis speculates that maximization of utility is the driving force behind individual agents decision to choose among available alternatives and thus individual preference distribution is a consequence of choices made by the whole population. This conjectures makes the RUM model appealing to theorists and practitioners alike. In the recreation demand models, a number of important additions have been made to the model to reflect the nature of the problem faced by consumers. This includes the introduction of an aggregate demand incorporating the need for multiple choices over a period of time.

The focus of this dissertation is to add to the existing literature on applying RUM models to recreation demand by consistently estimating and evaluating the demand for and welfare derived from recreational sites. The models proposed in this dissertation allows the Random Utility Maximization (RUM) model to be robust to issues such as: limitations in the information available to researchers; relaxing the assumption of constant marginal utility of income and, incorporating model uncertainty in the estimation process. Another way to look at this issues is in terms of model misspecification and the models proposed in each of the chapters proposes to correct for these misspecification. For example, the first paper, looks at respecifying the RUM model to incorporate unobserved site characteristics by the introduction of alternative specific constants with the ability to recover parameters of the observed site characteristics that is of interest for policy scenario simulations. Ignoring this issues, we will argue, will lead to inconsistent estimates of the parameters that determine the demand for recreational site and consequently lead to mistaken inference and policy recommendations.

The models are then applied to study visitation patterns of Iowans to recreational sites in order to measure the value residents place in the existing lake recreation opportunities, as well as the changes in welfare that would result from proposed site improvements. The data used for all the papers in this dissertation is from the Iowa Lakes Valuation Project at Iowa State University. This is a four year panel data study, sponsored by the Iowa Department of Natural Resources and the US EPA, that seeks to elicit the visitation patterns of Iowans to major recreational lakes in the state. We make use of the first panel of the data for 2002 sent to 8,000 households to elicit their visitation patterns to 129 principal Iowa lakes with socio-demographic data. There is also a parallel data collected by Iowa State University's Limnology Laboratory that collected three times a year, over the course of a five-year project, thirteen distinct water quality measurements at each of the lakes. This provides detailed information on the characteristics of each lake.

The first paper *Controlling for Observed and Unobserved Site Characteristics in RUM Models of Recreation Demand* accounts for the fact that researchers do not always have data on all the attributes that a site has to offer individuals. The nature of many environmental goods restricts the site attributes data available to researchers and to the extent that site specific factors are omitted from the analysis and correlated with either observed site attributes or the marginal utility of income parameter, the resulting parameter estimates and subsequent welfare analysis will be biased. This is in line with the classic omitted variable bias problem. Until recently, researchers typically ignore this problem and estimate the demand for the site assuming that all the attributes are known. 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. This paper controls for unobserved site specific attributes in a RUM model of recreation demand through the use of alternative specific constants. Whereas

Murdock (2006) employs a contraction mapping algorithm to estimate site level fixed effects (drawing on earlier work by Berry, Levinsohn and Pakes (1995)), we couch the problem within a Hierarchical Bayesian framework. The model is fit via Markov Chain Monte Carlo (MCMC) methods, combining data augmentation and Gibbs sampling.

The second paper, *RUM Models Incorporating Nonlinear Income Effects* relaxes the assumption of constant marginal utility of income in RUM models. One of the main appeals of the RUM model is its consistency with the assumption of a utility maximizing agent. However, even though the model specification can in principle be generalized to allow for seemingly valid cases of varying effects of marginal utility of income, researchers typically impose constant marginal utilities. The assumption of constant marginal utility of income is a restrictive formulation of individual preferences and choice behavior, preventing the satisfaction derived from a recreational good to vary across individuals depending on their level of income. In this chapter, we propose a RUM model that allows for nonlinear income effect that combines nonparametric estimation with Taylor series approximation and Bayesian methodology of hierarchical modeling. Data collected in the Iowa lakes project and, as is typical of many research in the area of recreation demand, have income brackets data available for each agent.

The third paper is *Model Uncertainty and Recreation Demand* applies the techniques of Bayesian Model Averaging (BMA) and variable selection to incorporate model uncertainty in recreation demand models. The motivation for this is the fact that economic theory provides relatively little guidance regarding the form that the relationship between environmental goods and utility should take and which variables ought to be included in the analysis. In many applications choices must be made between, for example, level and logarithmic specifications for an environmental characteristic. While model selection criterion can be used to narrow the set of specifications, there is the risk that the analyst (even inadvertently) may engage in a “fishing” process among the available models, biasing the final outcome of the analysis. This paper proposes an alternative approach that draws on the Bayesian paradigm to integrate the variable selection process into the model and reflect the accompanying uncertainty about which is the “correct” specification into subsequent counterfactual predictions.

To conclude, in all the three papers, extending the RUM model to allow for the uniqueness of data in modeling recreation demand is the primary goal. This is important because the estimated models can be used to infer the value households place in access to sites and/or changes to site characteristics which can be a key information to policy-makers seeking to manage recreational resources. All the papers propose models to handle the issues raised showing that the model works with series of generated data experiments and application to the Iowa Lakes data. Welfare analysis algorithms are also suggested.

Dissertation organization

Each of the chapters is a separate paper, with an introduction, literature review, Model description and conclusion. The papers as listed above represent the next three chapters and the dissertation concludes with a general conclusion for all the papers. The tables and figures are presented in each of the papers.

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

A paper submitted to Journal of Environmental Economics and Management

Babatunde Abidoye

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 (2006), 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.* (2009)), we find that, with the addition of a full set of alternative specific constants, water quality attributes no longer influence the choice of where to recreate.

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 in 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 (1999) and Morey, Rowe and Watson (1993)), fish toxin levels Phaneuf, Kling, Herriges (2000)) or dummy variable indicators capturing different levels of water quality (Parsons, Helm and Bondelid (2003)).[1] 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.[2]

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

[1]There are, of course, exceptions. Hanemann (1981) 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 (1986) employ a large number of water quality attributes, finding that these factors again are significant determinants of recreation demand.

[2]It 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.

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 for 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 a recent article, Murdock (2006) provides a resolution to both problems. Drawing on innovations in the industrial organization literature by Berry (1994) and Berry, Levinsohn, and Pakes (1995), 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 the estimation procedure proposed by Murdock. In particular, and contrary to the claims in Murdock (2006), the procedure cannot be used to obtain maximum likelihood parameter estimates if the RUM model includes random parameters (the so-called *mixed logit* model).[3,4] Murdock's use of the contraction mapping is based on the claim that "...in any random utility model the inclusion of a dummy variable for a particular alternative means that the predicted number of times it is selected will be exactly equated with the actual number of times it is selected in the data" (Murdock (2006), p. 8.) Unfortunately, while this mean-fitting feature of maximum likelihood estimation does emerge for the standard

[3]See Train (2009) for a description of the mixed logit model.

[4]To our knowledge, Klaiber and von Haefen (2008) were the first to note the error in Murdock (2006).

logit model, it does not hold once random parameters are introduced.[5] 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.[6]

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 of unobserved site attributes, even those correlated with price or the observed site attributes.[7]

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 (2006), 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"

[5]The 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 Murdock (2006), 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.

[6]It should be noted that the limitation to Murdock's procedure in no way carries over to the earlier work of Berry (1994) and Berry, Levinsohn, and Pakes (1995). 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.

[7]By "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.

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 (2005) (section 7.1) and Lancaster (2004) (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.* (2009) 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 about 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 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.

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 (2006), 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 ($\boldsymbol{s}_j$), 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 $\boldsymbol{s}_j$ or p_{ij} (e.g., ϵ_{ij} is *i.i.d.* extreme value). That is

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

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, ($\boldsymbol{s}_j^o$), leading to a reduced form specification

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

where

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

In (3), $\boldsymbol{s}_j^u$ denotes the unobserved site attributes and $\boldsymbol{\alpha}_0^k$ $(k = o, u)$ denote the subset of parameters associated with $\boldsymbol{s}_j^k$ $(k = o, u)$.[8] 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 $\alpha_j^u \equiv \boldsymbol{s}_j^u \boldsymbol{\alpha}_0^u$, equation (1) becomes

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

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

$$\alpha_j = \boldsymbol{s}_j^o \boldsymbol{\alpha}_0^o + \boldsymbol{s}_j^u \boldsymbol{\alpha}_0^u \tag{5}$$

$$= \boldsymbol{s}_j^o \boldsymbol{\alpha}_0^o + \alpha_j^u \tag{6}$$

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

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

[8]The specification in equation (2) is similar to that used in Egan *et al.* (2009) and in much of the recreation demand literature.

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 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.[9] 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 (2006) 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 $\boldsymbol{\alpha}_0^o$ are informative for the ASC parameters, and we fail to capitalize upon this source of learning in the sequential approach

[9]Note that this regression will have only $J-1$ observations.

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.[10] 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 $(\alpha_1, \ldots, \alpha_{J-1},\beta)$ 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 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.

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

[10] There have been several papers using a Bayesian framework to estimate a model similar to that originally proposed in Berry, Levinsohn and Pakes (1995), including a full set of alternative specific constants to control for unobserved alternative attributes. Yang, Chen, Allenby (2003) develop a posterior simulation alternative to Berry, Levinsohn and Pakes (1995) in modeling aggregate supply and demand. However, the routine is conditional on correctly specifying the underlying supply relationships. Jiang, Manchanda and Rossi (2009) provide a Bayesian counterpart to the contraction mapping approach outlined in Berry, Levinsohn and Pakes (1995), though again the analysis is couched in the context of aggregate supply and demand data.

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 (1993), Herriges and Phaneuf (2002)). 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.[11] The particular form of the model we use is the repeated mixed logit (RXL) model employed by Egan *et al.* (2009) (allowing individual parameters of the model to be random), but with the addition of the full set of alternative specific constants advocated by Murdock (2006). 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.

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:

$$U_{ijt} = \begin{cases} \boldsymbol{z}_i \gamma + \varepsilon_{i0t} & j = 0 \\ \alpha_j + p_{ij}\beta + \varphi_i + \varepsilon_{ijt} & j = 1, \ldots, J. \end{cases} \tag{8}$$

In this form of the model, the utility from visiting any one of the recreation sites (i.e., $j = 1, \ldots, J$) 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 variation in preferences to recreate in the population. The mean of this assumed normal distribution is restricted to be zero for identification purposes,

[11]Herriges, Kling and Phaneuf (1999) provide a summary of the repeated logit model and the implications of its underlying assumptions.

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 Herriges and Phaneuf (2002)).

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 $\varepsilon_{ijt} \overset{iid}{\sim} \mathcal{N}(0,1)$. 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 $\boldsymbol{z}_i\gamma$ in equation (1), 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:

$$\tilde{U}_{ijt} = U_{ijt} - U_{i0t} = \alpha_j + p_{ij}\beta - \boldsymbol{z}_i\gamma + \varphi_i + \tilde{\varepsilon}_{ijt} \tag{9}$$

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

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

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

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

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

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

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

$$\tilde{\boldsymbol{U}}_{\boldsymbol{i\cdot t}} = \boldsymbol{M}_{\boldsymbol{i\cdot t}}\boldsymbol{\theta} + \boldsymbol{v}_{\boldsymbol{i\cdot t}}, \tag{11}$$

where

$$\boldsymbol{M}_{\boldsymbol{i\cdot t}} = \begin{bmatrix} \boldsymbol{I}_{\boldsymbol{J}} & \boldsymbol{p}_{\boldsymbol{i\cdot}} & \mathbf{1}_{\boldsymbol{J}} \otimes \boldsymbol{z}_i \end{bmatrix} \quad \text{and} \quad \boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\alpha}_{\boldsymbol{\cdot}}' & \beta & \boldsymbol{\gamma}' \end{bmatrix}'.$$

Although $\boldsymbol{M}_{\boldsymbol{i\cdot t}}$ 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, $\boldsymbol{v}_{\boldsymbol{i\cdot t}}$ is the composite error term

$$\boldsymbol{v}_{\boldsymbol{i\cdot t}} = \mathbf{1}_{\boldsymbol{J}}\varphi_i + \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{i\cdot t}} \tag{12}$$

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

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

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

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

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 $t = 1, 2, \ldots, 52$ 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.

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]).[12] The alternative specific constants also provide the sole avenue by which the observed site attributes impact the recreation demand decision. In Murdock's (2006) 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, $\boldsymbol{s}_j^o$. 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.,

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

[12] We 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.

where $\boldsymbol{q}_j$ 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 $\boldsymbol{q}_j$, although it is often assumed - at least implicitly - that these unobserved characteristics are uncorrelated with those in $\boldsymbol{q}_j$. 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 (2005)(section 7.1) and Lancaster (2004) (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, we do claim victory in the sense that the inclusion of site-specific constants, in conjunction with our particular posterior simulator, will yield accurate estimates of the first-stage parameters in (8), even when such correlation and confounding is present. 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:

$$\boldsymbol{\alpha_0} \sim \mathcal{N}(\boldsymbol{\mu_\alpha}, \boldsymbol{V_\alpha}) \tag{16}$$

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

$$\boldsymbol{\gamma} \sim \mathcal{N}(\boldsymbol{\mu_\gamma}, \boldsymbol{V_\gamma}) \tag{18}$$

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

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

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 (2007) (pp. 336) and represents the *inverse gamma* distribution. The prior means ($\boldsymbol{\mu_\alpha}$, μ_β, $\boldsymbol{\mu_\gamma}$) in our empirical work and generated data experiments are set to zero vectors of appropriate dimensions with the respective prior variance for the parameters ($\boldsymbol{V_\alpha}$, V_β, and $\boldsymbol{V_\gamma}$) 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[13]

Posterior Simulator

Let

$$\boldsymbol{\Xi} = \begin{bmatrix} \boldsymbol{\theta} & \boldsymbol{\alpha_0} & \sigma^2_\alpha & \{\varphi_i\} & \sigma^2_\varphi \end{bmatrix}$$

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

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

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 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.

[13]It 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.

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 $\tilde{\boldsymbol{U}}_{\boldsymbol{i\cdot t}}|\Xi, \boldsymbol{y}$

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 $\alpha_j, \beta, \boldsymbol{\gamma}$, 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.[14] 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^r_{ijt} for all $j \neq k$ from a truncated normal distribution with mean μ_{ij}, a variance of 1, and upper truncation point $U_{ikt} = U^{r-1}_{ikt}$.

2: Draw U^r_{ikt} from a truncated normal distribution with its mean μ_{ij}, a variance of 1, and

[14]From equation (1), $\mu_{ij} = \alpha_j + P_{ij}\beta + \varphi_i$ for $j = 1, \ldots, J$ and $\mu_{i0} = \boldsymbol{z}_i\gamma$.

a lower truncation point at the $\max(U^r_{ijt})$ for all $j \neq k$.

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

<u>**Step 2:**</u> $\boldsymbol{\theta}|\boldsymbol{\Xi}_{-\theta}, \tilde{\boldsymbol{U}}, \boldsymbol{y}$

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

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

where

$$\mathbf{D}_\theta \equiv \left[T \sum_{i=1}^{N} \mathbf{M}'_{\mathbf{i\cdot t}} \boldsymbol{\Omega}^{-1} \mathbf{M}_{\mathbf{i\cdot t}} + \boldsymbol{\Sigma}_\theta^{-1} \right]^{-1}$$
$$\mathbf{d}_\theta \equiv \sum_t \sum_i \mathbf{M}'_{\mathbf{i\cdot t}} \boldsymbol{\Omega}^{-1} \tilde{\boldsymbol{U}}_{\boldsymbol{i\cdot t}} + \boldsymbol{\Sigma}_\theta^{-1} \mu_\theta$$

and

$$\boldsymbol{\Sigma}_\theta = \begin{bmatrix} \sigma^2_\alpha I_J & 0 & \mathbf{0} \\ \mathbf{0} & V_\beta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{V}_\gamma \end{bmatrix}, \ \mu_\theta = \begin{bmatrix} \mathbf{Q}\boldsymbol{\alpha}_\mathbf{0} \\ \mu_\beta \\ \boldsymbol{\mu}_{\boldsymbol{\gamma}} \end{bmatrix}.$$

The blocking strategy adopted in this step is key to our proposed simulator, jointly drawing the parameters $\boldsymbol{\alpha}.$, β and $\boldsymbol{\gamma}$. 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 $\boldsymbol{\gamma}$ 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 $\beta, \boldsymbol{\gamma}$ 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 $\boldsymbol{\gamma}$, 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: $\{\varphi_i\}|\boldsymbol{\Xi}_{-\varphi_i}, \tilde{\boldsymbol{U}}, \boldsymbol{y}$

$$\{\varphi_i\}|\boldsymbol{\Xi}_{-\varphi_i}, \tilde{\boldsymbol{U}}, \boldsymbol{y} \sim \mathcal{N}(D_\varphi d_\varphi, D_\varphi) \tag{23}$$

where

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

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

$$\boldsymbol{M}_{i\cdot t}^{\varphi} = \begin{bmatrix} \boldsymbol{I}_J & \boldsymbol{P}_i \end{bmatrix} \quad \text{and} \quad \boldsymbol{\theta}^{\varphi} = \begin{bmatrix} \boldsymbol{\alpha}.' & \beta \end{bmatrix}'.$$

Step 4: $\boldsymbol{\alpha}_0|\boldsymbol{\Xi}_{-\boldsymbol{\alpha_0}}, \tilde{\boldsymbol{U}}, \boldsymbol{y}$

The remaining steps of our posterior simulator involve the sampling of parameters of the hierarchical priors. Once we condition on the $\boldsymbol{\alpha_j}$'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:

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

where

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

Step 5: $\sigma_\alpha^2|\boldsymbol{\Xi}_{-\sigma_\alpha^2}, \tilde{\boldsymbol{U}}, \boldsymbol{y}$

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

<u>**Step 6:**</u> $\sigma^2_\varphi|\boldsymbol{\Xi}_{-\sigma^2_\varphi}, \tilde{\boldsymbol{U}}, \boldsymbol{y}$

$$\sigma^2_\varphi|\boldsymbol{\Xi}_{-\sigma^2_\varphi}, \tilde{\boldsymbol{U}}, \boldsymbol{y} \sim IG\left[\frac{N}{2} + \alpha_\varphi, \left(b_\varphi^{-1} + .5\sum_{i=1}^{N}\varphi_i^2\right)^{-1}\right]. \tag{26}$$

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 provide information as to how many draws will be needed to achieve the same level of precision that would be obtained under independent and identical distribution (*iid*) sampling from the posterior.

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., $\boldsymbol{z_i}$'s) are included in the experiment, 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

$$\alpha_j = \alpha_{01} + \alpha_{02}\boldsymbol{s}_j^o + \alpha_0^u\boldsymbol{s}_j^u \tag{27}$$

with $\boldsymbol{s}_j^o$ and $\boldsymbol{s}_j^u$ drawn jointly from the bivariate normal distribution, with

$$\begin{bmatrix} \boldsymbol{s}_j^o \\ \boldsymbol{s}_j^u \end{bmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_o^2 & \rho\sigma_o\sigma_u \\ \rho\sigma_o\sigma_u & \sigma_u^2 \end{pmatrix}\right]. \tag{28}$$

In our experiments we use $\sigma_o^2 = 0.03$ and $\sigma_u^2 = 0.05$, with ρ denoting the degree of correlation between the observed and unobserved site attributes. This, in turn, implies that

$$\alpha_j|\boldsymbol{s}_j^o \overset{ind}{\sim} \mathcal{N}\left(\alpha_{01} + \alpha_{02}\boldsymbol{s}_j^o, (\alpha_0^u)^2\sigma_u^2\right). \quad j = 1, 2, \ldots, J. \tag{29}$$

In the notation of our hierarchical specification for the alternative specific constants (see equation (15)), $\boldsymbol{q}_j = [1 \quad \boldsymbol{s}_j^o]$, $\alpha_0 = (\alpha_{01} \quad \alpha_{02})'$, and $\sigma_\alpha = \alpha_0^u \sigma_u$. Experiment 1 assumes that $\boldsymbol{s}_j^o$ and $\boldsymbol{s}_j^u$ are uncorrelated (i.e., $\rho = 0$), while experiment 2 assumes that they are correlated with $\rho = 0.7$. Note that in these first two experiments there is no correlation between the unobserved attribute $\boldsymbol{s}_j^u$ and travel cost.

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 1a and Table 1b. The Gibbs sampler described in section 3.3 is implemented for 55,000 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 for so;e of the parameters are provided in Figure 1.

Parameter posterior means and posterior standard deviations for the first generated data experiment are reported in the second column of Table 1a and Table 1b. Even with only a small number of alternatives in the choice set, the results suggest that our algorithm performs well in recovering the parameters of the model.

One concern, however, in any MCMC algorithm is 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 so-called *inefficiency factors*:

$$\sqrt{\text{inefficiency factor}} \equiv \sqrt{f} = \sqrt{1 + 2\sum_{j=1}^{m-1}(1 - \frac{j}{m})\rho_j}, \tag{30}$$

where m is the number of draws after convergence and ρ_j represents the correlation between simulations j periods (iterations) apart. These factors help to intuitively understand the numerical accuracy of the MCMC-based posterior estimate by comparing it to the level of precision that would have been obtained under *iid* sampling. For example, if a given parameter's simulations yield $\sqrt{f} = 5$, then we will need to run the simulator for $25M$ iterations in order to obtain the same level of numerical precision that would be obtained from M *iid* posterior draws.

The inefficiency factors are reported for each parameter in column 4 of Table 1. Unfortunately, estimates of these inefficiency factors are often quite high, especially for the coefficient on travel cost with an inefficiency factor of about 21. Such high values are an expected consequence of a model of this type and this complexity, as we must sample the latent utilities from their respective conditional posterior distributions and also sample $\{\varphi_i\}$ separately from the blocked elements of $\boldsymbol{\theta}$. The large inefficiency factors indicate that we will need to run our simulations longer - in some cases more than $400M$ times as long - to obtain the same level of accuracy as would be obtained in M *iid* draws. It is also worth noting that the posterior mean of σ_α^2 is high (as indicated in Table 1), relying as it does on only $J = 30$ alternative specific constants. Consequently, the prior will be quite influential.[15] This will improve with a larger number of alternatives.

Parameter posterior means and standard deviations for the second generated data experiment are reported in the fifth column of Table 1. In this experiment there is substantial correlation ($\rho = 0.7$) between the observed and unobserved site attributes. As anticipated, we find that the posterior mean of α_{02} is far from the true value of 0.98, suffering what is akin to omitted variables bias found in a classical linear setting. 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 $\boldsymbol{s}_j^o$ and $\boldsymbol{s}_j^u$, 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:

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

[15] However, the mode of the posterior distribution (not reported) is consistently close to the true value even when $J = 30$.

As in the previous two experiments, the unobserved site characteristic ($\boldsymbol{s}_j^u$) 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.[16] 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 $Corr(p_{ij}, \boldsymbol{s}_j^u) \approx 0.22$. In experiment 4 we create a higher level of correlation by setting $\kappa = 0.7$, yielding $Corr(p_{ij}, \boldsymbol{s}_j^u) \approx 0.46$. For both experiments the simulator described in section 3.3 was used to obtain 55,000 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 2a and Table 2b, along with the true parameters used to generate the data.

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 or 1.5 standard deviations 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 $(\beta, \boldsymbol{\gamma})$. Indeed, when we implement such a simulator to this data, we obtain a posterior mean and standard deviation of β equal to -5.63 and 0.07, 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, though for some parameters mixes rather slowly, 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 $\boldsymbol{s}_j^o$.

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

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 ($\boldsymbol{Q}$) 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 Sechi transparency (a measure of the depth of water clarity), Nitrogen, and Chlorophyll.[17] In addition, the exact same data was used by Egan *et al.* (2009) 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 more than fifty-two day trips per-year (223). The latter sample exclusion follows the procedure used in Egan *et al.* (2009), wherein the authors note that individuals

[17] The 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.* (2009).

taking such frequent trips are usually local residents who are counting casual visit 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.[18] 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.[19] 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.

[18] Egan *et al.* (2009) also found that their qualitative results are not sensitive to the specific cut-off of fifty-two trips per year.

[19] The "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.

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 then used to initialize two different chains, run simultaneously on two different machines with different seeds. Each of these runs produced 30,000 simulations, leaving us with a total of 60,000 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 $\alpha_j' s$, 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 12.5 days to provide the set of simulations used here.

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 $\boldsymbol{q}_j$. 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.

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 $\alpha_j's$ are higher (less negative) for popular destination lakes in the state, such as Okoboji Lake, 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.[20] Consistent with our expectations, the posterior mean for β in Table 4 is negative (-0.0137), 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.* (2009), 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 Egan *et al.* (2009), the signs and overall magnitudes of these coefficients in our analysis are not sensitive to the construction of $\boldsymbol{q}_j$. 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.

[20] While 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.

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.* (2009), 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.* (2009), 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.[21] Only Total Phosphorus, which contributes to algae growth, is convincingly massed on one side of zero in our analysis, suggesting (as in Egan *et al.* (2009)) that high levels of Total Phosphorus reduces the appeal of a site.

Table 6 represents our attempt to reproduce results in Egan *et al.* (2009) 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 Egan *et al.* (2009) 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

[21] It 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.

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.

Model Comparison

In what follows we consider four different models, similar to those employed by Egan *et al.* (2009), and with an eye toward determining the set of factors that play 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 and non-water quality attributes.

- **Model B:** Includes Secchi transparency as the only water quality attribute, all non-water quality attributes included.

- **Model C:** No water quality attributes included, all non-water quality attributes included.

- **Model D:** Includes Secchi transparency as the sole site attribute (all other attributes dropped).

When the Bayesian is faced with a problem of model selection / comparison of this sort, he or she effectively treats the model itself as a parameter and notes, by Bayes rule:

$$p(\mathcal{M}_k|data) = \frac{p(data|\mathcal{M}_k)p(\mathcal{M}_k)}{p(data)} \tag{32}$$

where $\mathcal{M}_k$ denotes the k^{th} model under consideration, $p(data|\mathcal{M}_k)$ 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:

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

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:

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

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

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

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. The results of these Bayes factor calculations are presented in Table 7.

Though individuals might not be responsive to the individual water quality attributes as shown in our result, there is the possibility that a combination of these attributes might be important determinant of recreational demand. In order to investigate this, we compute the Bayes factor for each of the model to compare if the data favors this hypothesis. Model C is the model that includes none of the water quality attributes relative to Model A and B that has one or more water quality attribute included in the model. The result of the Bayes factor of 5.61E+17 when we compare the unrestricted model (Model A) to Model C signals that Model C is clearly preferred to Model A. This implies that the water quality measures do not matter jointly either.

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 characteristics. Specifically, we consider changes to site attributes from baseline levels ($\boldsymbol{Q}^0$) to an alternative conditions ($\boldsymbol{Q}^1$). Although our focus is on the welfare implications of a change in site quality, other types of counterfactuals (such as the loss of an entire site) can be considered using a similar algorithm with few modifications.

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

$$\boldsymbol{\Upsilon_{it}^{s}(\Xi_{-\alpha.}, Q^{s}) = \max_{j}(U_{ijt}^{s} | \Xi_{-\alpha.}, Q^{s}) \quad s = 0, 1,} \tag{36}$$

where $\boldsymbol{\alpha}. = (\alpha_1, \ldots, \alpha_J)$ denotes the vector of alternative specific constants. Changes in the site characteristics impact individual consumers by altering the overall appeal of the sites, as reflected in the α_j's. Thus, we no longer have a single set of alternative specific constants, but a set for each scenario (denoted $\boldsymbol{\alpha}_{\cdot}^{s}$). We use the hierarchical structure in equation (15) to simulate the changes to these constants resulting from a change in the site attributes.

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

$$CV = \frac{T}{-\beta} E_{\boldsymbol{\Xi}} \left[\Upsilon_{it}^{1}(\boldsymbol{\Xi}_{-\boldsymbol{\alpha}.}, Q^{s}) - \Upsilon_{it}^{0}(\boldsymbol{\Xi}_{-\boldsymbol{\alpha}.}, Q^{s}) \right] \tag{37}$$

The term in square brackets above measures the change in utility per choice occasion. Since the choice occasions are treated as independent of each other, the seasonal change in expected utility is just a simple scaled multiple of the expected change per choice occasion. 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 (35) 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 $\boldsymbol{\Xi}_{-\boldsymbol{\alpha}.}^{(r)} (r = 1, \ldots, R)$, denote a draw from the posterior distribution of $\boldsymbol{\Xi}_{-\boldsymbol{\alpha}.}$. Draw $\alpha_j^{s(r)} \quad (j = 1, \ldots, J; s = 0, 1)$ 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 $\boldsymbol{q}_j^{s} \boldsymbol{\alpha}_0^{(r)}$ and variance $(\sigma_\alpha^{(r)})^2$.

Step 2: Draws of the utility levels $U_{ijt}^{s(r)}|\boldsymbol{\Xi}_{-\boldsymbol{\alpha}.}, Q^s$ are obtained using equation (1), with

$$U_{ijt}^{s(r)} = \begin{cases} \boldsymbol{z}_i\gamma^{(r)} + \varepsilon_{i0t}^{(r)} & j = 0 \\ \alpha_j^{s(r)} + p_{ij}\beta^{(r)} + \varphi_i^{(r)} + \varepsilon_{ijt}^{(r)} & j = 1, \ldots, J. \end{cases} \tag{38}$$

where

$$\epsilon_{ijt}^{(r)} \sim \mathcal{N}(0,1) \text{ and } \quad \varphi_i^{(r)} \sim \mathcal{N}(0, \sigma^2_{\varphi(r)}). \tag{39}$$

The simulation based estimate is then computed using

$$\widehat{CV} = \frac{1}{R}\sum_{r=1}^{R} \frac{T}{-\beta^{(r)}} \left[\left(\max_j U_{ijt}^{1(r)}\right) - \left(\max_j U_{ijt}^{0(r)}\right)\right]. \tag{40}$$

We applied the algorithm for calculating the compensating variation to the unrestricted model. We look at a scenario in which the water quality attributes in 9 zonal lakes (listed in Table 4) are improved to the quality of Lake W. Okoboji which is one of the clearest lakes in the state. The results are presented in table 8. The CV estimates in dollar amount reflect the dollar amount that the individuals will loose or gain if the policy is implemented. As expected, the estimated compensated variation shows no strong evidence that improving these attributes lead to higher welfare. The probability that the density of the compensation variation density is greater than zero $[P(\widehat{CV} > 0)]$ is approximately 34%. The CV estimates are generally negative, though caution should be applied in interpreting this. The result is less definitive owing largely to greater posterior uncertainty associated with the water quality attributes.

Also reported in Table 8 is the loss of lake W. Okoboji which can be thought of as a ban on the use of the lake due to several reasons such as flooding which can be a problem in Iowa. This is an interesting computation give the nature of the lake and the number of visits to this lake. The result shows that the closure of the lake will result in a CV value of nearly \$9 per Iowa household to compensate for this change. Despite the availability of alternative sites for households to visit, this welfare loss seems reasonably high.

Summary

Controlling for unobserved site characteristics is important as researchers are typically restricted to single measure of site attributes and 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 inconsistent. In this study, we extend earlier recreational demand models by allowing for unobserved site characteristics in the model. In particular, a Bayesian posterior simulator is employed for fitting this model, which not only consistently estimates the price effect on utility, but also unifies the analysis into a single model while allowing for a more general error structure.

A number of conclusions can be drawn from the model and particular application described in this paper. Firstly, our methodology offers an improved estimate of the travel cost parameter without explicitly modeling the correlation between the unobserved site characteristics and travel cost. Our approach of isolating the travel cost parameter from the unobserved site characteristics through the use of hierarchical prior on the alternative specific constant has genuine promise for analyzing similar models in other fields. While our Bayesian solution to this problem is similar to that discussed by Yang, Chen, and Allenby (2003) and Jiang, Manchanda and Rossi (2009), our particular strategy of blocking parameters together when implementing the posterior simulator insulates us from concerns regarding omitted site-level characteristics and their impact on estimation results. Secondly, our methodology results in the shrinkage of the site-specific parameters toward a common mean through the prior which is a desirable sampling distribution property of the estimator in Murdock (2006) and can help in applications where some sites are not visited in the sample.

In terms of the application, the results indicate that unobserved site characteristics do matter in understanding factors that affect site visitation patterns in Iowa. In contrast to earlier studies, there is no clear indication that site water quality attributes substantially influence lake visitation. That is, with alternative specific constant, water quality attributes does not matter. Similar to earlier studies, we found evidence that socio-demographic variables are important factors in determining recreational demand and site attributes also encourage individuals to take trips to these lakes. However, in contrast to other studies, our results are robust to different model specification of water quality attributes.

Tables and Figures

Table 1a: Generated Data Experiments 1 and 2

	Experiment 1: Exogenous Observable Site Attributes ($\rho = 0$)			Experiment 2: Endogenous Observable Site Attributes ($\rho = 0.7$)	
Parameter	Posterior Mean	True Value	$\sqrt{\text{Inefficiency Factor}}$	Posterior Mean	True Value
$\alpha_0(1)$	-3.43	-3.52	7.94	-3.48	-3.52
	(0.06)			(0.06)	
$\alpha_0(2)$	0.93	0.98	1.00	1.59	0.98
	(0.29)			(0.28)	
β_0	-4.49	-4.50	20.30	-4.49	-4.50
	(0.02)			(0.03)	
$\gamma_0(1)$	0.92	0.96	10.49	0.89	0.96
	(0.37)			(0.36)	
$\gamma_0(2)$	0.73	0.75	12.33	0.71	0.75
	(0.21)			(0.21)	
σ_φ^2	0.46	0.40	20.12	0.46	0.40
	(0.04)			(0.04)	
σ_α^2	0.09	0.05	1.00	0.08	0.05
	(0.02)			(0.02)	

Posterior standard deviation in parentheses

Table 1b: Generated Data Experiments 1 and 2[a]

	Experiment 1: Exogenous Observable ($\rho = 0$)			Experiment 2: Endogenous Observable ($\rho = 0.7$)	
Parameter	Posterior Mean	True Value	$\sqrt{\text{Inefficiency Factor}}$	Posterior Mean	True Value
α_1	-3.15	-3.27	13	-2.94	-3.04
	(0.03)			(0.03)	
α_2	-3.91	-4.00	17	-3.96	-4.02
	(0.03)			(0.03)	
α_3	-3.77	-3.90	18	-3.95	-4.07
	(0.04)			(0.03)	
α_4	-3.21	-3.28	16	-3.42	-3.46
	(0.03)			(0.03)	
α_5	-3.81	-3.93	12	-3.93	-4.00
	(0.03)			(0.03)	
α_6	-3.07	-3.17	19	-3.25	-3.30
	(0.04)			(0.04)	
α_7	-3.85	-3.97	21	-3.98	-4.06
	(0.04)			(0.03)	
α_8	-3.18	-3.29	19	-3.04	-3.13
	(0.03)			(0.03)	
α_9	-3.22	-3.30	18	-3.02	-3.05
	(0.03)			(0.03)	
α_{10}	-3.50	-3.57	20	-3.48	-3.51
	(0.04)			(0.03)	
α_{11}	-3.58	-3.68	17	-3.44	-3.53
	(0.03)			(0.03)	
α_{12}	-3.46	-3.56	15	-3.50	-3.55
	(0.03)			(0.03)	
α_{13}	-3.45	-3.54	18	-3.62	-3.67
	(0.04)			(0.03)	
α_{14}	-3.50	-3.61	17	-3.57	-3.66
	(0.04)			(0.04)	
α_{15}	-3.63	-3.71	12	-3.64	-3.69
	(0.03)			(0.03)	

[a] Posterior standard deviation in parentheses

Table 1b (Contd): Generated Data Experiments 1 and 2[a]

	Experiment 1: Exogenous Observable ($\rho = 0$)			Experiment 2: Endogenous Observable ($\rho = 0.7$)	
Parameter	Posterior Mean	True Value	$\sqrt{\text{Inefficiency Factor}}$	Posterior Mean	True Value
α_{16}	-2.84	-2.93	17	-2.85	-2.89
	(0.03)			(0.03)	
α_{17}	-3.72	-3.78	16	-3.90	-3.93
	(0.03)			(0.03)	
α_{18}	-3.16	-3.22	15	-3.12	-3.14
	(0.04)			(0.04)	
α_{19}	-3.01	-3.11	20	-3.00	-3.07
	(0.04)			(0.03)	
α_{20}	-3.71	-3.86	19	-3.94	-4.04
	(0.03)			(0.03)	
α_{21}	-3.39	-3.45	16	-3.43	-3.46
	(0.03)			(0.03)	
α_{22}	-3.43	-3.50	14	-3.49	-3.52
	(0.03)			(0.03)	
α_{23}	-3.54	-3.67	19	-3.67	-3.75
	(0.04)			(0.03)	
α_{24}	-3.48	-3.54	20	-3.46	-3.49
	(0.03)			(0.03)	
α_{25}	-3.70	-3.81	20	-3.99	-4.05
	(0.04)			(0.03)	
α_{26}	-3.03	-3.08	20	-3.02	-3.03
	(0.03)			(0.03)	
α_{27}	-3.30	-3.37	20	-3.28	-3.34
	(0.03)			(0.03)	
α_{28}	-3.26	-3.31	15	-3.22	-3.25
	(0.03)			(0.03)	
α_{29}	-3.33	-3.40	20	-3.38	-3.42
	(0.04)			(0.04)	
α_{30}	-3.21	-3.31	20	-3.41	-3.47
	(0.03)			(0.03)	

[a] Posterior standard deviation in parentheses

Table 2a: 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
$\alpha_0(1)$	-3.53	-3.52	-3.55	-3.52
	(0.06)		(0.06)	
$\alpha_0(2)$	0.92	0.98	0.90	0.98
	(0.29)		(0.29)	
β_0	-4.50	-4.50	-4.52	-4.50
	(0.02)		(0.02)	
$\gamma_0(1)$	.92	0.96	0.90	0.96
	(0.03)		(0.04)	
$\gamma_0(2)$	0.80	0.75	0.78	0.75
	(0.02)		(0.02)	
σ^2_φ	0.41	0.40	0.41	0.40
	(0.01)		(0.01)	
σ^2_α	0.08	0.05	0.09	0.05
	(0.02)		(0.02)	

[a]Posterior standard deviation in parentheses

Table 2b: Generated Data Experiments 3 and 4[a]

Parameter	Experiment 3: Moderate Correlation Between Price and Unobservables ($\kappa = 0.5$) Posterior Mean	True Value	Experiment 4: High Correlation Between Price and Unobservables ($\kappa = 0.7$) Posterior Mean	True Value
α_1	-3.29	-3.27	-3.26	-3.27
	(0.03)		(0.03)	
α_2	-4.01	-4.00	-4.00	-4.00
	(0.02)		(0.03)	
α_3	-3.89	-3.90	-3.92	-3.90
	(0.02)		(0.02)	
α_4	-3.27	-3.28	-3.32	-3.28
	(0.02)		(0.02)	
α_5	-3.90	-3.93	-3.93	-3.93
	(0.02)		(0.03)	
α_6	-3.17	-3.17	-3.21	-3.17
	(0.02)		(0.02)	
α_7	-3.94	-3.97	-3.98	-3.97
	(0.02)		(0.02)	
α_8	-3.29	-3.29	-3.29	-3.29
	(0.03)		(0.04)	
α_9	-3.32	-3.30	-3.38	-3.30
	(0.03)		(0.04)	
α_{10}	-3.58	-3.57	-3.59	-3.57
	(0.02)		(0.03)	
α_{11}	-3.69	-3.68	-3.68	-3.68
	(0.03)		(0.03)	
α_{12}	-3.56	-3.56	-3.58	-3.56
	(0.02)		(0.03)	
α_{13}	-3.56	-3.54	-3.58	-3.54
	(0.02)		(0.02)	
α_{14}	-3.59	-3.61	-3.61	-3.61
	(0.02)		(0.03)	
α_{15}	-3.70	-3.71	-3.73	-3.71
	(0.02)		(0.03)	

[a] Posterior standard deviation in parentheses

Table 2b (contd): Generated Data Experiments 3 and 4[a]

Parameter	Experiment 3: Moderate Correlation Between Price and Unobservables ($\kappa = 0.5$) Posterior Mean	True Value	Experiment 4: High Correlation Between Price and Unobservables ($\kappa = 0.7$) Posterior Mean	True Value
α_{16}	-2.97	-2.93	-2.94	-2.93
	(0.03)		(0.03)	
α_{17}	-3.81	-3.78	-3.83	-3.78
	(0.02)		(0.02)	
α_{18}	-3.23	-3.22	-3.25	-3.22
	(0.03)		(0.03)	
α_{19}	-3.12	-3.11	-3.15	-3.11
	(0.02)		(0.03)	
α_{20}	-3.86	-3.86	-3.89	-3.86
	(0.02)		(0.02)	
α_{21}	-3.47	-3.45	-3.50	-3.45
	(0.03)		(0.03)	
α_{22}	-3.53	-3.50	-3.55	-3.50
	(0.02)		(0.03)	
α_{23}	-3.68	-3.67	-3.73	-3.67
	(0.03)		(0.02)	
α_{24}	-3.58	-3.54	-3.56	-3.54
	(0.02)		(0.03)	
α_{25}	-3.77	-3.81	-3.79	-3.81
	(0.02)		(0.02)	
α_{26}	-3.09	-3.08	-3.10	-3.08
	(0.02)		(0.03)	
α_{27}	-3.39	-3.37	-3.41	-3.37
	(0.02)		(0.03)	
α_{28}	-3.34	-3.31	-3.34	-3.31
	(0.02)		(0.03)	
α_{29}	-3.41	-3.40	-3.44	-3.40
	(0.02)		(0.03)	
α_{30}	-3.30	-3.31	-3.38	-3.31
	(0.02)		(0.02)	

[a]Posterior standard deviation in parentheses

Table 3: Summary Statistics

Variable	Model Variable	Mean	Std. Dev.	Min	Max
Total Day Trips (2002)[a]	T_i	6.33	9.97	0	50
Travel Cost ($100's)	P_{ij}	1.37	.83	4.4E-03	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 (μg/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 (μg/l)	$Q_{j(13)}$	40.64	38.01	2.45	182.92

[a] 39% of the sample did not visit any lake that year

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

[c] Unsure = 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 along with Travel Cost and Demographic Parameters.[a]

Parameter	Model A		Model B		Model C		Model D	
	Mean	$P(\cdot > 0 \mid y)$	Mean	$P(\cdot > 0 \mid y)$	Mean	$P(\cdot > 0 \mid y)$	Mean	$P(\cdot > 0 \mid y)$
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 Lcaria	-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

[a] We only report the result for the nine major lakes in each zone and Okoboji. Estimates for the other sites are presented in Table 9.

Table 5: Posterior Means of Hierarchical Parameters

Parameter	Parameter							
	Model A[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_{01}	-4.1487	0.0000	-4.087	0.0000	-4.064	0.0000	-3.700	0.0000
	(-579.84)							
Lake Attributes								
α_{02} (Acres)	0.0001	0.9999	1.0E-04	1.0000	0.0001	1.0000		
	(5.87)							
α_{03} (Ramps)	0.2616	0.9956	0.204	0.9870	0.209	0.9888		
	(-2.18)							
α_{04} (Wake)	0.1261	0.9601	0.102	0.9870	0.098	0.9261		
	(2.61)							
α_{05} (Handicap)	0.1738	0.9914	0.166	0.9915	0.162	0.9907		
	(-0.85)							
α_{06} (State Park)	0.2118	0.9975	0.237	0.9997	0.236	0.9998		
	(-2.67)							
Water Quality								
α_{07} (Sechi)	-0.0129	0.3851	0.005	0.5672			0.026	0.7318
	(6.20)							
α_{08} (Total Nitrogen)	0.0126	0.8305						
	(7.70)							
α_{09} (Total Phosphorus)	-0.0012	0.0304						
	(11.38)							
α_{010} (Volatile SS)	-0.0002	0.4902						
	(10.02)							
α_{011} (Inorganic SS)	0.0030	0.8940						
	(10.6)							
α_{012} (Cyanobacteria)	2.11E-06	0.5203						
	(20.22)							
α_{013} (Chlorophyll)	0.0011	0.7933						
	(12.67)							
Variance								
σ^2_α	0.1283	1.0000	0.118	1.0000	0.120	1.0000	0.190	1.0000
σ^2_φ	3.0359	1.0000	1.800	1.0000	1.800	1.0000	1.770	1.0000

[a] The 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		
Sechi	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"		$H0$	$H0$	$H0$
Model A		1.92E+16	5.61E+17	3.85E+11
		(75)[a]	(81.74)	(53.35)
Model B			17.68	3.19E-15
			(5.74)	(-66.76)

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

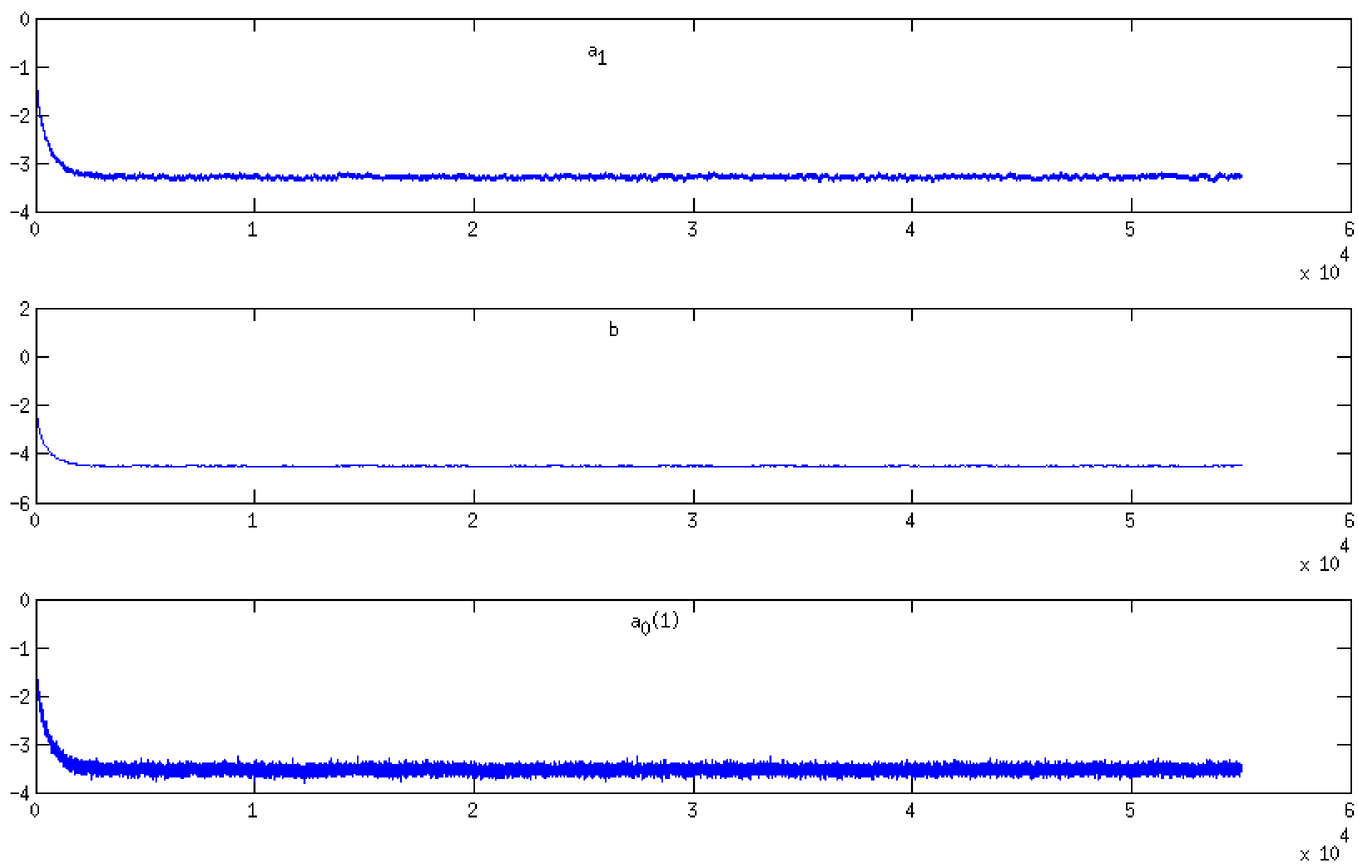
Table 8: Annual Compensating Variation Estimates

	CV ($)	$P(\cdot > 0\|y)$
Close W. Okoboji	-8.83	0
Upgrade 9 Zone lakes to W. Okoboji	-2.09	0.3445

Table 9: Posterior Means of Alternative Specific Constants

Lake	Mean	$P(\cdot > 0\|y)$	Lake	Mean	$P(\cdot > 0\|y)$	Lake	Mean	$P(\cdot > 0\|y)$
Arbor	-3.9531	0.0000	Hooper	-4.3621	0.0000	North Twin	-3.3963	0.0000
Arrowhead	-3.6007	0.0000	Indian	-3.825	0.0000	Oldham	-4.3514	0.0000
Arrowhead	-4.197	0.0000	Ingham	-3.473	0.0000	Otter Creek	-4.0709	0.0000
Ave. of the Saints	-4.2732	0.0000	Kent Park	-3.7271	0.0000	Ottumwa Lagoon (proper)	-3.4701	0.0000
Badger Creek	-3.7862	0.0000	Lacey-Keosauqua	-3.4238	0.0000	Pierce Creek	-4.0572	0.0000
Badger	-3.4403	0.0000	Ahquabi	-3.5608	0.0000	Pleasant Creek	-3.4609	0.0000
Beaver	-4.2021	0.0000	Anita	-3.5061	0.0000	Pollmiller	-3.7479	0.0000
Beed's	-3.471	0.0000	Cornelia	-3.568	0.0000	Prairie Rose	-3.5396	0.0000
Big Creek	-3.0414	0.0000	Darling	-3.495	0.0000	Rathbun	-2.8365	0.0000
Spirit Lake	-2.7191	0.0000	Geode	-3.3603	0.0000	Red Haw	-3.6508	0.0000
Black Hawk	-3.2616	0.0000	Hendricks	-3.7857	0.0000	Red Rock	-2.8546	0.0000
Blue	-3.5363	0.0000	Icaria	-3.286	0.0000	Robert's Creek	-3.9832	0.0000
Bob White	-4.0834	0.0000	Iowa	-3.8595	0.0000	Rock Creek	-3.6636	0.0000
Brigg's Woods	-3.7134	0.0000	Keomah	-3.8389	0.0000	Rogers	-4.2409	0.0000
Brown's	-3.4209	0.0000	Manawa	-2.951	0.0000	Saylorville	-2.8576	0.0000
Brushy Creek	-3.2852	0.0000	Macbride	-3.1802	0.0000	Silver	-3.3694	0.0000
Carter	-3.6447	0.0000	Miami	-3.8	0.0000	Silver	-4.2648	0.0000
Casey	-3.9555	0.0000	Minnewashata	-3.4989	0.0000	Silver	-4.0497	0.0000
Center	-3.6776	0.0000	Lake of The Hills	-3.6902	0.0000	Silver	-3.5992	0.0000
Central	-4.019	0.0000	Three Fires	-3.413	0.0000	Slip Bluff	-4.3962	0.0000
Clear	-2.6609	0.0000	Orient	-4.1171	0.0000	South Prairie	-4.2258	0.0000
Cold Springs	-3.7921	0.0000	Pahoja	-3.4659	0.0000	Spring	-3.9248	0.0000
Coralville	-2.9909	0.0000	Smith	-3.7449	0.0000	Springbrook	-3.6208	0.0000
Crawford Creek	-3.9971	0.0000	Sugema	-3.4248	0.0000	Storm Lake	-2.9066	0.0000
Crystal	-3.5002	0.0000	Wapello	-3.3908	0.0000	Swan	-3.3787	0.0000
Dale Maffit	-4.0361	0.0000	Little River	-3.7504	0.0000	Thayer	-4.2525	0.0000
DeSoto Bend	-3.3166	0.0000	Little Sioux Park	-3.6362	0.0000	Three Mile	-3.3801	0.0000
Diamond	-3.8888	0.0000	Little Spirit	-3.1446	0.0000	Trumbull	-3.8127	0.0000
Dog Creek	-3.819	0.0000	Little Wall	-3.9045	0.0000	Tuttle	-4.0171	0.0000
Don Williams	-3.4962	0.0000	Littlefield	-3.8378	0.0000	Twelve Mile	-3.4411	0.0000
East Osceola	-3.7384	0.0000	Lost Island	-3.2257	0.0000	Union Grove	-3.945	0.0000
East Okoboji	-2.5417	0.0000	Lower Gar	-3.3816	0.0000	Upper Gar	-3.4804	0.0000
Easter	-3.6856	0.0000	Lower Pine	-3.691	0.0000	Upper Pine	-3.5796	0.0000
Eldred Sherwood	-4.1119	0.0000	Manteno Pond	-4.335	0.0000	Viking	-3.4303	0.0000
Five Island	-3.4866	0.0000	Mariposa	-4.0882	0.0000	Volga	-3.4514	0.0000
Fogle	-4.0875	0.0000	Meadow	-4.5621	0.0000	West Okoboji	-2.3539	0.0000
George Wyth	-3.4002	0.0000	Meyers	-4.1341	0.0000	West Osceola	-3.6718	0.0000
Green Belt	-4.4943	0.0000	Mill Creek	-3.7956	0.0000	White Oak	-4.5478	0.0000
Green Castle	-4.233	0.0000	Mitchell Impoundment	-4.4572	0.0000	Williamson Pond	-4.7821	0.0000
Green Valley	-3.5961	0.0000	Moorehead	-3.8776	0.0000	Willow	-3.9834	0.0000
Greenfield Lake	-4.1219	0.0000	Mormon Trail	-4.1189	0.0000	Wilson	-4.2772	0.0000
Hannen	-3.9533	0.0000	Nelson Park	-4.2833	0.0000	Windmill	-3.9386	0.0000
Hawthorn	-3.8175	0.0000	Nine Eagles	-3.8103	0.0000	Yellow Smoke	-3.5456	0.0000
Hickory Grove	-3.8573	0.0000						

Figure 1 Trace Plots of Selected Parameters $(\alpha_1, \beta, \alpha_0(1))$- Generated Data Experiment 1



RUM Models Incorporating Nonlinear Income Effects.

A paper to be submitted to Journal of Environmental Economics and Management

Babatunde Abidoye

Abstract

The assumption of constant marginal utility though convenient is a restrictive formulation of individual preferences and choice behavior. In this chapter we propose a generalized formulation of the Random Utility Maximization (RUM) model that allows the data to dictate the relationship between preferences and income and prices.

Introduction

The Random Utility Maximization (RUM) model has been a major tool for estimating demand systems in economics and marketing. One of the main appeals of the model is its consistency with the assumption of a utility maximizing agent. However, even though the model specification can in principle be generalized to allow for seemingly valid cases of varying effects of marginal utility of income, researchers typically impose constant marginal utilities. The assumption of constant marginal utility of income is a restrictive formulation of individual preferences and choice behavior, preventing the satisfaction derived from a recreational good to vary across individuals depending on their level of income.

A few studies in the literature have highlighted the implications on parameter and welfare estimates when nonlinearities in income are ignored (e.g. Herriges and Kling, 1999; Morey, Rowe and Watson, 1993; and Morey, Sharma and Karlstrom, 2003), yet most empirical analysis continue to assume a constant marginal utility of income. One practical reason for this is that, when the marginal utility of income is constant, the expected compensating variation (ECV) associated with a hypothetical policy scenario (e.g., changing site attributes or availability) take the familiar and convenient log-sum form for the typical logit specifications (i.e., multinomial and nested logit). However, once income enters the conditional indirect utility function nonlinearly, the expected CV no longer reduces to a convenient form (McFadden, 1995). While alternative approaches for obtaining welfare estimates do exist, they are significantly more complex (as in the case of McFadden's, 1995 Monte Carlo simulation or Dagsvik and Karlstrom's, 2005 algorithm) or serve only as an approximation (as in the case of Morey, Rowe and Watson's, 1993 representative agent approximation).

These studies typically add, for example, a quadratic term on income to the indirect utility function or assume a nonlinear functional form for utility to capture income nonlinearities. However, these formulations by themselves still restrict the effect of income and prices on individual preference to the specification defined by the researcher. In this paper, we allow income to impact preferences in a piecewise linear fashion, providing more flexibility in allowing the data to determine how income (and travel costs) alter the choice behavior of individuals. We also demonstrate how welfare analysis proceed within the framework.

This chapter is organized as follows. In section 2, we briefly describe why introducing nonlinearity into the RUM model is important and also review related literature in this area. Section 3 presents the model and how the parameters of interest are estimated. Section 4 describes a generated data experiment while section 5 presents a description of the data and application Section 6 provides the empirical results. Section 7 describes our posterior inference procedure and results. The paper concludes with a summary in section 8.

Related Literature

Before reviewing related literature, we will briefly describe the problem of imposing nonlinearity into RUM models. Suppose that the utility individual i receives from visiting site j is a function of an alternative specific constant α_j, the amount of income the individual has left to spend on other commodities if he visits site j, and an idiosyncratic error component (ϵ_{ij}).[1] The residual income is defined as the difference in income and travel cost to the site $(y_i - P_{ij})$ and (ϵ_{ij}) is uncorrelated across sites and individuals (e.g. ϵ_{ij} is *i.i.d.* extreme value). That is

$$U_{ij} = \alpha_j + f(y_i - P_{ij}) + \epsilon_{ij} \quad i = 1, 2, \ldots, N; j = 1, 2, \ldots, J. \tag{1}$$

Given the above preference representation, the probability that an individual i will choose alternative j over another alternative k (for $j \neq k$) can be defined as:

$$Pr(U_{ij} > U_{ik}) \equiv Pr(U_{ij} - U_{ik} > 0) = Pr\left[(\alpha_j - \alpha_k) + \{f(y_i - P_{ij}) - f(y_i - P_{ik})\} + (\epsilon_{ij} - \epsilon_{ik}) > 0\right]. \tag{2}$$

Typically, recreation demand studies that use the RUM model assume a linear form for the function $f(.)$ such that $f(.) = \beta(y_i - P_{ij})$. Given this representation, equation (2) can be rewritten as:

$$\begin{aligned} Pr(U_{ij} - U_{ik} > 0) &= Pr\left[(\alpha_j - \alpha_k) + \{\beta(y_i - P_{ij}) - \beta(y_i - P_{ik})\} + (\epsilon_{ij} - \epsilon_{ik}) > 0\right] && (3) \\ &= Pr\left[(\alpha_j - \alpha_k) + \{\beta(y_i) - \beta(y_i)\} - \{\beta(P_{ij} - P_{ik})\} + (\epsilon_{ij} - \epsilon_{ik}) > 0\right] && (4) \\ &= Pr\left[(\alpha_j - \alpha_k) - \{\beta(P_{ij} - P_{ik})\} + (\epsilon_{ij} - \epsilon_{ik}) > 0\right] && (5) \end{aligned}$$

reducing the differences in utility as a function of income and travel cost to a simple linear form with income eliminated from the choice probability. However, one might expect that at least for a certain category of individuals, income will be an important determinant of choice and price might not necessarily have a linear relationship with this preference.

As a way of accounting for income and price effects, Morey, Rowe and Watson (1993) specified an additional term to the linear specification to add some form of nonlinearity. Specifically they let $f(.) = \beta_0(y_i - P_{ij}) + \beta_{00}\sqrt{(y_i - P_{ij})}$ in the indirect utility function. They concluded that if the parameter β_{00} is significantly different from zero, then there exists nonlinearity in how income affects utility. Even though this specification leads to the capturing of an income

[1] Note that α_j captures site attributes affecting choices but for the purpose of this paper, we focus primarily on income

effect, there is no reason to believe that the shape of the function $f(.)$ specified will be a appropriate representation of the relationship between utility and income and price effect.

Also, Herriges and Kling (1999) compared the result assuming three different functional forms for the indirect utility function using the case of sport fishing modal choice.[2] They highlighted the implications on parameter and welfare estimates when nonlinearities in income are ignored and also extended and refined the work of McFadden (1995) and Morey, Rowe and Watson (1993) on calculating welfare estimates when marginal utilities vary with respect to both prices and income. They concluded that in their particular application, there is little evidence that the introduction of nonlinear income effects has a significant impact on parameter estimates and welfare values. They however signal caution in making general inferences based on the need for additional empirical applications and other alternative preference characterization.

Morey, Sharma and Karlstrom (2003) suggested a piece-wise linear spline function as a way of incorporating income effects for cases where exact level of income is not observed. The study employs a disaggregated approach to estimating the parameters of marginal utility of income assuming no relationship between the income groups. While this study is similar in spirit to what we propose in this chapter, the methodology is different. The Taylor series approximation and hierarchical nature of our approach accounts for the fact that the income groups are not fully independent. This places more structure and can allow for more accurate estimation with large number of income groups and also with the assumption of diminishing marginal utility of income. All these studies also suggested a method of evaluating welfare estimates for changes in scenarios. The properties of some of these methods are discussed in Dagsvik and Karlstrom (2005).

While these additions are useful in examining the role of income in the choice process, any test for nonlinear income effects are conditional on the specific functional form assumed. The contribution of this paper is to revisit the concept of diminishing marginal utility of income by proposing a RUM model that allows for nonlinear income effect. This approach

[2]They compared welfare estimates using linear, Generalized Leontief and Translog functions.

makes use of the Taylor series approximation to approximate the $f(.)$ function which allows for estimation of the marginal utility of income as piecewise linear. The proposed model makes use of nonparametric estimation with Taylor series approximation and Bayesian methodology of Hierarchical modeling. This is an advantage for applied researchers because it is not as computationally intensive as a full semiparametric estimation of the $f(.)$ function that will be described later. It also has the advantage of abstracting from the problem of what time frame or income value to use when estimating the indirect utility function. Also, since most data used in the recreational demand literature only have information on the income brackets that the household belongs too, the Taylor series approximation will be expected to be a natural fit for the data.

Model

In this section we present a model that accounts for nonlinear income effect by estimating a function that has a different effect on the level of utility depending on the income category the individual belongs to. This method makes use of the Taylor series function approximation to the function $f(.)$ in equation (1). Rewriting equation (1) to capture a nesting structure with the introduction of an individual random effect for individuals that visited any site j, we have:

$$U_{ijt} = \begin{cases} D_i\gamma + f(y_i) + \varepsilon_{ijt} & \text{if } j = \text{stay at home} \\ \alpha_j + f(y_i - P_{ij}) + \varphi_i + \varepsilon_{ijt} & \text{for } j = 1, \ldots, J. \end{cases} \tag{6}$$

The model is similar to that used in the first chapter with the integration of individuals' choice among alternatives and problem of allocating time between multiple recreation sites. While this model can be generalized, it aids comparison to models used in the recreation demand literature. The parameters and variables are also as defined in the first chapter with α_j defined as the alternative specific constant, φ_i individual specific effect, ε_{ijt} idiosyncratic error term and D_i is demographic characteristics. The choice among the J alternatives on choice occasion t depends only on relative utility levels. Thus, if we take the difference in utility with respect to the baseline utility of stay at home, equation (1) can be rewritten as:

$$\tilde{U}_{ijt} = \alpha_j - D_i\gamma + f(y_i - P_{ij}) - f(y_i) + \varphi_i + \tilde{\varepsilon}_{ijt} \quad \text{for } j = 1, ..., J \tag{7}$$

Full Semi-Parametric Model

One alternative to the estimation of the function relating income to utility is to consider a full semi-parametric estimation. This estimation procedure will follow the algorithm described in Koop and Poirier (2004) which has been shown to overcome many of the problems that plagues nonparametric estimations. This is achieved by imposing an informative prior that serves both as a smoothing parameter and a solution to the curse of dimensionality problem. This is particularly important in our application given that there is a high level of variability in the travel cost between individuals and sites. The approach of handling the nonparametric treatment of $f(.)$ in Koop and Poirier (2004) will lead to sorting the data by values of $x_{ij} = (y_i - P_{ij})$ such that $x_{11} \leq x_{12} \leq \ldots \leq x_{1J} \leq \ldots \leq x_{NJ}$. The idea behind this approach is to estimate each unique point f_{ij} of the function $f(.)$ with neighboring residual income having similar estimate for the function. This results in the choice of the informative prior that assumes that the difference between $f_{h.}$ and $f_{k.}$ for $x_{h.} \geq x_{k.}$ should be small. This prior imposes structure on the parameter space so that estimation can be tractable.

One major difference between this model and that of Koop (Gary and Poirier) is the fact that we have two related functions to estimate nonparametrically. Equation (7) has the term $f(y_i)$ which will result in the problem of identification of each value of the function separately from the parameters of φ_i. This will make the use of a full semi-parametric estimation problematic without further reparameterization.[3] There is also the issue of income to use for the semiparametric estimation. Even though we assume that people make choices on recreational sites to visit on a weekly basis, the actual income value that they use in determining such choices is unknown to the researcher. Employing a full semi-parametric estimation will entail making an assumption of the income values which may not be right. Also, given that most households only provide information on the income brackets in our application, assuming a midpoint income value for the sake of estimating a semi-parametric model also raises questions on the validity of the estimates.

Taylor Series Approximation

Suppose we consider a first order Taylor series approximation to the residual income term *i.e.*, $(f(y_i - P_{ij}))$ around the baseline income in (7). Specifically,

[3]One way to deal with this will be to recenter the variables in equation (4) such that

$$\tilde{U}_{ijt} = \alpha_j + f(y_i - P_{ij}) + \varphi_i + \tilde{\varepsilon}_{ijt} \quad \text{for } j = 1, ..., J \tag{8}$$

where $\varphi_i \sim \mathcal{N}(-[D_i\gamma + f(y_i)], \sigma_\varphi^2)$. The estimation method will have to take account of the correlation with the parameters in $f(y_i - P_{ij})$

$$
\begin{aligned}
\tilde{U}_{ijt} &= \alpha_j - D_i\gamma + f(y_i - P_{ij}) - f(y_i) + \varphi_i + \tilde{\varepsilon}_{ijt} \\
&\approx \alpha_j - D_i\gamma + f(y_i) - P_{ij}f'(y_i) - f(y_i) + \varphi_i + \tilde{\varepsilon}_{ijt} \\
&= \alpha_j - D_i\gamma - P_{ij}f'(y_i) + \varphi_i + \tilde{\varepsilon}_{ijt}
\end{aligned}
\tag{9}
$$

With this approximation, we can estimate the marginal utility of income function $f'(y_i)$ in different ways. One possibility is to assume that it is a continuous function of income and estimate it nonparametrically. In our particular application with only income bracket data available, we will be dealing with few points depending on how the income brackets data are classified. In recognition that the data on income is in the form of income brackets, we can reparametize the last line of equation (9) to reflect a piecewise linear form that depends on the income bracket. That is:

$$
\tilde{U}_{ijt} \approx \alpha_j - D_i\gamma - P_{ij}\sum_{b=1}^{B}\theta_b I(y_i \in Y_b) + \varphi_i + \tilde{\varepsilon}_{ijt} \quad \text{for } b = 1, ..., B \tag{10}
$$

where

$$
I(y_i \in Y_b) = \begin{cases} 1 & \text{if individual } i \text{ is in the b}^{th} \text{ income bracket} \\ 0 & \text{if otherwise} \end{cases}
$$

The parameter θ_b is capturing $f'(y_b)$ - the marginal utility of income. The model is a generalization of the RUM model where the model reduces to the standard linear income model if $\theta_b = \theta\ \forall b$. It is also analogous to a spline type model of Morey, Sharma and Karlstrom (2003), with the cut-points for the splines being determined by the income brackets in the model.

As highlighted earlier, this approach is not as computationally intensive as a full semi-parametric model and the discrete nature of the income data in most cases reduces the number of parameters to be estimated. This makes the approach advantageous since the exact income data is usually not known and partly abstracts away the problem of what time frame to use in computing income in the semi-parametric model (though implicitly it affects the validity of the Taylor series approximation).

Posterior Simulator

In this section we describe a posterior simulator to estimate the conditional posterior distributions of the parameters of equation (9). All the posterior distributions that will be presented are of the standard form and we will apply the Gibbs sampler to draw from the complete posterior distributions. The Gibbs sampler is an iterative algorithm that has become an indispensable tool to Bayesian econometricians and researchers undertaking simulation based inference. The idea is to draw from the posterior conditional distributions rather than the joint posterior distributions themselves that are usually difficult to draw from.[4] We outline each posterior conditional distribution below.

Before describing the posterior conditionals, we write equation (9) concisely by stacking over all J alternatives such that:

$$\tilde{\boldsymbol{U}}_{i.t} = \boldsymbol{M}_i\boldsymbol{\beta} + \boldsymbol{P}_i^* \sum_{b=1}^{B} \boldsymbol{\theta}_b I(y_i \in Y_b) + \mathbf{1}_{\boldsymbol{J}}\varphi_i + \tilde{\boldsymbol{\varepsilon}}_{i.t} \tag{11}$$

where $\mathbf{1}_{\boldsymbol{J}}$ is a $J \times 1$ vector of ones;

$$\tilde{\boldsymbol{U}}_{i.t} = \begin{pmatrix} \tilde{U}_{i1t} \\ \tilde{U}_{i2t} \\ \vdots \\ \tilde{U}_{iJt} \end{pmatrix}; \quad \boldsymbol{M}_i = \begin{pmatrix} 1 & 0 & \dots & 0 & -D_i \\ 0 & 1 & \dots & 0 & -D_i \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -D_i \end{pmatrix}; \quad \boldsymbol{P}_i^* = \begin{pmatrix} -P_{i1} \\ -P_{i2} \\ \vdots \\ -P_{iJ} \end{pmatrix}.$$

and

$$\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\alpha}' \\ \gamma' \end{bmatrix} \quad \tilde{\boldsymbol{\varepsilon}}_{i.t} \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma} \equiv \begin{bmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 2 \end{bmatrix}\right)$$

Stacking over choice occasions (t), we have:

$$\begin{aligned} \tilde{\boldsymbol{U}}_{i..} &= [\mathbf{1}_T \otimes \boldsymbol{M}_i]\boldsymbol{\beta} + [\mathbf{1}_T \otimes \boldsymbol{P}_i^*] \sum_{b=1}^{B} \boldsymbol{\theta}_b I(y_i \in Y_b) + \mathbf{1}_{\boldsymbol{TJ}}\varphi_i + \tilde{\boldsymbol{\varepsilon}}_{i..} \quad (12) \\ &= [\mathbf{1}_T \otimes \tilde{\boldsymbol{M}}_i]\boldsymbol{\Psi} + \mathbf{1}_{\boldsymbol{TJ}}\varphi_i + \tilde{\boldsymbol{\varepsilon}}_{i..} \quad (13) \end{aligned}$$

[4]Theoretical explanation and proof of how it works can be found in Casella and George (1992) and other Bayesian econometrics texts such as Koop, Poirier and Tobias (2007) provides more explanation of the Gibbs sampler.

Where

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\theta}_{.} \end{bmatrix}$$

and

$$\tilde{\boldsymbol{M}}_i = \begin{pmatrix} \boldsymbol{M}_i & \boldsymbol{I}_{b.} \otimes \boldsymbol{P}_i^* \end{pmatrix}$$

$\boldsymbol{I}_{b.}$ is a $1 \times B$ vector of zeros with one on the b column of the income bracket the individual belongs to. This regrouping of variables is for ease of computation.

The model outlined above is essentially a latent variable realizations of the indirect utility function, and is linked to the observed individual choices $y_{it}(\tilde{\boldsymbol{U}}_{\boldsymbol{i \cdot t}})$ as follows:

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

y_{it} is constructed from the observed count of the number of visits to the sites by individual i over a calendar year. This assumes that within the RXL framework, we assume that series of decisions are made by the individual at particular subintervals (choice occasions) − which in our case is weekly − in a manner that is consistent with this aggregate data.

Hierarchical Prior

Given the hierarchical nature of the parameters in (8), we outline the hierarchical priors for these parameters in this section. For the marginal utility of income parameters (i.e., the $\theta_b's$) we consider two priors. First, one possible prior for the marginal utility of income is an hierarchical prior on θ_b. This assumes that the individuals in the different income groups share the same class of distribution such that the parameter θ_b is drawn from the same normal population:

$$\theta_b \sim \mathcal{N}(\theta_0, \sigma_\theta^2) \tag{15}$$

A second possible prior on the marginal utility of income is an informative prior that is in the spirit of the *smoothing* prior of Koop and Poirier (2004). Though the prior has other advantages in the context of combating issues in nonparametric regression problems, for the purposes of this paper, it introduces the potential for smoothing the marginal utility of income function. This prior does not assume that individuals in different income brackets have a common mean as the previous prior assumes. The value of this prior is in the case where there are large number of income brackets and there are few or no observations per income bracket. The prior will induce a form of local averaging in estimating a continous function.

To describe this second prior in more detail, given sorted income brackets defined as y_b, let $\Delta_b = y_b - y_{b-1}$. The differencing matrix H can be defined as follows:

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & \vdots & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \vdots & 0 & 0 & 0 \\ \Delta_2^{-1} & [-\Delta_2^{-1} - \Delta_3^{-1}] & \Delta_3^{-1} & 0 & \vdots & 0 & 0 & 0 \\ 0 & \Delta_3^{-1} & [-\Delta_3^{-1} - \Delta_4^{-1}] & \Delta_4^{-1} & \vdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \vdots & \Delta_{B-1}^{-1} & [-\Delta_{B-1}^{-1} - \Delta_B^{-1}] & \Delta_B^{-1} \end{pmatrix} \tag{16}$$

We will also specify a hyperparameter η that governs the degree of smoothness of the marginal utility of income parameters. Large values of η will make the marginal utility of income function to be excessively jumpy leading to model overfit while too low values will over smooth the function. To let the appropriate amount of smoothing be revised by the data, we include it as a parameter in the model. The hierarchical prior for this parameter can be specified as:

$$\eta \sim \mathcal{IG}(a_\eta, b_\eta)$$

Thus, the prior for the marginal utility of income parameters can be deduced by first specifying a prior

$$H\theta|\eta \sim \mathcal{N}(0, \eta V_\theta)$$

which implies:

$$\theta|\eta \overset{iid}{\sim} \mathcal{N}(0, \eta H^{-1} V_\theta [H^{-1}]') \tag{17}$$

where

$$V_\theta = \begin{pmatrix} cI_2 & 0 \\ 0 & I_{B-2} \end{pmatrix}$$

The prior for α_j and φ_i are as described in chapter one except that the prior on α_j does not include the observed site attributes. It still allows for unobserved site characteristics but since the focus of this chapter is on the marginal utility of income, we only allow for a constant as the mean of the alternative specific constants. The prior for the site-specific constant is such that while the J sites share the same class of distribution in terms of their effect on utility, these effects are assumed to be independent. It is also a form of decomposing the "overall" effect of the attributes of the sites into observed and unobserved effect. With our specification, the estimate of the other parameters of the model should not be significantly affected.[5] The prior is specified as:

$$\alpha_j \sim \mathcal{N}(\alpha_0, \sigma_\alpha^2). \quad j = 1, 2, \ldots, J \tag{18}$$

Finally, we set priors for the "common" parameters. For the first prior on the marginal utility of income, the priors for the variability in the marginal utility of income across income brackets is summarized by Σ_θ. The priors for the variability in the site and individual level parameters is summarized by σ_φ and σ_α. We let the appropriate amount of smoothing η be revised by the data.

$$\alpha_0 \sim \mathcal{N}(\mu_\alpha, V_\alpha) \tag{19}$$

$$\theta_0 \sim \mathcal{N}(\mu_\theta, V_\theta) \tag{20}$$

$$\Sigma_\theta \sim \mathcal{IG}(a_\theta, b_\theta) \tag{21}$$

$$\sigma_\alpha^2 \sim \mathcal{IG}(a_\alpha, b_\alpha) \tag{22}$$

$$\boldsymbol{\gamma} \sim \mathcal{N}(\boldsymbol{\mu_\gamma}, \boldsymbol{V_\gamma}) \tag{23}$$

$$\sigma_\varphi^2 \sim \mathcal{IG}(a_\varphi, b_\varphi) \tag{24}$$

[5]Further details on the benefit of introducing an hierarchical structure on the alternative specific constant and justification is outlined in Chapter 1.

While $\mathcal{N}$ above obviously refers to the normal distribution, $\mathcal{IG}(\cdot,\cdot)$ follows the notation in Koop, Poirier and Tobias (2007) (pp. 336) and represents the *inverse gamma* distribution. One appeal of this prior is the conjugacy property that eases computation. The values of the terminal parameters above are chosen to be practically uninformative to allow for dominance of the data. For the prior for the common parameters, we set the prior means $(\mu_\alpha, \mu_\theta, \boldsymbol{\mu}_\gamma)$ to zero, and prior variances $(V_\alpha, V_\theta, \boldsymbol{V}_\gamma)$ equal to the identity matrix with the appropriate dimensions. We also set the prior for the hyperparameters of the variances and smoothing parameter as $a_\alpha = a_\varphi = a_\theta = a_\eta = 5$ and $b_\alpha = b_\varphi = b_\theta = b_\eta = 1$. The choice of the hyperparameters for the smoothing parameter seems reasonable in the generated data experiment and did not constrain the function to be necessarily linear.

Complete Posterior Conditionals

Let

$$\boldsymbol{\Xi} = \begin{bmatrix} \boldsymbol{\Psi} & \theta_0 & \Sigma_\theta & \boldsymbol{\alpha_0} & \sigma_\alpha^2 & \boldsymbol{\varphi.} & \sigma_\varphi^2 \end{bmatrix}$$

denote all the parameters of the model with variation on the parameters depending on the prior on the marginal utility income.[6]

<u>**Step 1:**</u> *Complete posterior conditional for* $\boldsymbol{\Psi}$

Using the result of **?** with blocking step, the posterior conditional for $\boldsymbol{\Psi}$ is given as:

$$\boldsymbol{\Psi}|\boldsymbol{\Xi}_{-\boldsymbol{\Psi}}, \tilde{\boldsymbol{U}}, \boldsymbol{y} \sim N(\mathbf{D_\Psi d_\Psi}, \mathbf{D_\Psi}). \tag{25}$$

where

$$\boldsymbol{D}_\Psi \equiv \left[T \sum_{i=1}^{N} \tilde{\boldsymbol{M}}_{it}' \Omega^{-1} \tilde{\boldsymbol{M}}_{it} + \boldsymbol{\Sigma}_\Psi^{-1} \right]^{-1}$$
$$\boldsymbol{d}_\Psi \equiv \sum_t \sum_i \tilde{\boldsymbol{M}}_{it}' \Omega^{-1} \tilde{\boldsymbol{U}}_{\boldsymbol{i.t}} + \boldsymbol{\Sigma}_\Psi^{-1} \boldsymbol{\mu}_\Psi$$

and depending on if we choose the prior from equation (15) or (17) $\boldsymbol{\Sigma}_\Psi$ will differ.

[6]For the smoothing prior, the full set of parameters will be: $\boldsymbol{\Xi} = \begin{bmatrix} \boldsymbol{\Psi} & \eta & \boldsymbol{\alpha_0} & \sigma_\alpha^2 & \boldsymbol{\varphi.} & \sigma_\varphi^2 \end{bmatrix}$

From equation (15), we have

$$\boldsymbol{\Sigma}_{\boldsymbol{\Psi}} = \begin{bmatrix} \sigma_\alpha^2 I_J & 0 & \mathbf{0} \\ \mathbf{0} & V_\gamma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Sigma}_\theta \end{bmatrix}, \; \mu_{\boldsymbol{\Psi}} = \begin{bmatrix} \mathbf{1}_J \alpha_0 \\ \boldsymbol{\mu}_\gamma \\ \mathbf{1}_B \theta_0 \end{bmatrix}$$

while equation (17) will give us

$$\boldsymbol{\Sigma}_{\boldsymbol{\Psi}} = \begin{bmatrix} \sigma_\alpha^2 I_J & 0 & \mathbf{0} \\ \mathbf{0} & V_\gamma & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \eta H^{-1} V_\theta [H^{-1}]' \end{bmatrix}, \; \mu_{\boldsymbol{\Psi}} = \begin{bmatrix} \mathbf{1}_J \alpha_0 \\ \boldsymbol{\mu}_{\boldsymbol{\gamma}} \\ \mathbf{0} \end{bmatrix}$$

The blocking strategy adopted here is appealing given that it insulates us from the issue of inconsistency that might arise from potential correlation between variables at the site level and the travel cost. This is because the sampler proceeds jointly in drawing the parameters differently from the simulator from the standard panel model estimator with random effect.

Step 2(a1): *Complete posterior conditional for* $\boldsymbol{\theta}_0$

Using a similar argument in Lindley and Smith (1974), the posterior condition for $\boldsymbol{\theta}_0$ can be expressed as:

$$\boldsymbol{\theta}_0 | \boldsymbol{\Xi}_{-\boldsymbol{\theta}_0}, \tilde{\boldsymbol{U}}, \boldsymbol{y} \sim \mathcal{N}(\boldsymbol{D}_{\theta_0} \boldsymbol{d}_{\theta_0}, \boldsymbol{D}_{\theta_0}). \tag{26}$$

where

$$\boldsymbol{D}_{\theta_0}^{-1} = \mathbf{1}'_B (\boldsymbol{I}_B \otimes \Sigma_\theta^{-1}) \mathbf{1}_B + V_\theta^{-1} = B \Sigma_\theta^{-1} + V_\theta^{-1}$$

and

$$\boldsymbol{d}_{\theta_0} = \mathbf{1}'_B (\boldsymbol{I}_B \otimes \Sigma_\theta^{-1}) \widetilde{\theta} + V_\theta^{-1} \mu_\theta = B \Sigma_\theta^{-1} \overline{\theta} + V_\theta^{-1} \mu_\theta$$

with $\widetilde{\theta} = \begin{bmatrix} \theta'_1 & \theta'_2 & \ldots & \theta'_B \end{bmatrix}'$, and $\overline{\theta} = (1/B) \sum_{b=1}^{B} \theta_b$.

Step 2(a2): *Complete posterior conditional for* Σ_θ^{-1}

$$\Sigma_\theta | \boldsymbol{\Xi}_{-\Sigma_\theta}, \tilde{\boldsymbol{U}}, \boldsymbol{y} \sim IG \left[\frac{B}{2} + a_\theta, \left(b_\theta^{-1} + 0.5 \sum_{b=1}^{B} (\boldsymbol{\theta}_b - \boldsymbol{\theta}_{\mathbf{0}})^2 \right)^{-1} \right] \tag{27}$$

Step 2b: *Complete posterior conditional for* η

$$\eta|\mathbf{\Xi}_{-\eta}, \tilde{\boldsymbol{U}}, \boldsymbol{y} \sim IG\left[\frac{B}{2} + a_\eta, \left(b_\eta^{-1} + (1/2)\boldsymbol{\theta}' H' [V_\theta]^{-1} H\boldsymbol{\theta}\right)^{-1}\right] \tag{28}$$

Step 3: $\alpha_0|\mathbf{\Xi}_{-\alpha_0}, \tilde{\boldsymbol{U}}, \boldsymbol{y}$

Once we condition on the $\alpha_j' s$, the posterior conditional for α_0 is similar to that of a linear regression and is of the form:

$$\alpha_0|\mathbf{\Xi}_{-\alpha_0}, \tilde{\boldsymbol{U}}, \boldsymbol{y} \sim \mathcal{N}(\boldsymbol{D}_{\alpha_0}\boldsymbol{d}_{\alpha_0}, \boldsymbol{D}_{\alpha_0}) \tag{29}$$

where

$$\mathbf{D}_{\alpha_\mathbf{0}} = (\mathbf{1}_J'\mathbf{1}_J/\sigma_\alpha^2 + V_\alpha^{-1})^{-1} \text{ and } \mathbf{d}_{\alpha_\mathbf{0}} = \mathbf{1}_J'\boldsymbol{\alpha}/\sigma_\alpha^2 + V_\alpha^{-1}\mu_\alpha$$

Step 4: $\sigma_\alpha^2|\mathbf{\Xi}_{-\sigma_\alpha^2}, \tilde{\boldsymbol{U}}, \boldsymbol{y}$

$$\sigma_\alpha^2|\mathbf{\Xi}_{-\sigma_\alpha^2}, \tilde{\boldsymbol{U}}, \boldsymbol{y} \sim IG\left[\frac{J}{2} + a_\alpha, \left(b_\alpha^{-1} + .5\sum_{j=1}^{J}(\alpha_j - \alpha_0)^2\right)^{-1}\right] \tag{30}$$

Step 5: $\{\varphi_i\}|\mathbf{\Xi}_{-\varphi_i}, \tilde{\boldsymbol{U}}, \boldsymbol{y}$ Stacking over sites and choice occasions, we have:

$$\varphi_i|\mathbf{\Xi}_{-\varphi_i}, \tilde{\boldsymbol{U}}, \boldsymbol{y} \sim \mathcal{N}(D_\varphi d_\varphi, D_\varphi) \tag{31}$$

where

$$D_\varphi^{-1} = \mathbf{1}_{JT}'(\boldsymbol{I}_T \otimes \Sigma^{-1})\mathbf{1}_{TJ} + \frac{1}{\sigma_\varphi}; \text{ and } d_\varphi = \mathbf{1}_{JT}'(\boldsymbol{I}_T \otimes \Sigma^{-1})(\boldsymbol{w}_{\boldsymbol{i}\cdot\cdot})$$

and $\boldsymbol{w}_{\boldsymbol{i}\cdot\cdot} = \tilde{\boldsymbol{U}}_{\boldsymbol{i}\cdot\cdot} - [\mathbf{1}_T \otimes \boldsymbol{M}_i]\boldsymbol{\beta} - [\mathbf{1}_T \otimes \boldsymbol{P}_i^*]\boldsymbol{\theta}_b I(y_i \in Y_b)$

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

$$\sigma_\varphi^2|\mathbf{\Xi}_{-\sigma_\varphi^2}, \tilde{\boldsymbol{U}}, \boldsymbol{y} \sim IG\left[\frac{N}{2} + a_\varphi, \left(b_\varphi^{-1} + .5\sum_{i=1}^{N}\varphi_i^2\right)^{-1}\right] \tag{32}$$

Step 7: Draw the $\tilde{\boldsymbol{U}}_{\boldsymbol{i.t}}|\mathbf{\Xi}, y_{it}.$

Given the structure of our model and to ease computation, we draw the latent utilities that individual i derives from visiting site j at the levels instead of differences. That is, we sample the U_{ijt} and then take the differences to get the $\tilde{U}_{ijt}$. This is straightforward to implement in our case given our assumption of *iid* distribution on the error term ε_{ijt}. At the structural level of the U_{ijt} in equation (3), there is no correlation among the alternatives conditional on the parameters $(\alpha_j, \beta, \gamma, \text{and } \varphi_i)$ of the model.

Each of the U_{ijt}'s are conditionally normal with mean μ and variance of 1 with truncation point that depends on the choice of the individual. That is, if an alternative is chosen, it must be the alternative that gives the maximum utility - this gives the upper truncation point for all the other alternatives.

We therefore follow 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 and variance from equation (1) and upper truncation point $U_{ikt} = U_{ikt}^{r-1}$.

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

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

The table below presents a summary result of a generated data experiment comparing the two priors.

Generated Data Experiment

Using generated data is important to illustrate the performance of the sampler and supports the ability of the model to consistently estimate the parameters we are interested in. In this section we describe the generated data experiment to illustrate how to allow for nonlinearity in the marginal utility of income in a RUM model.

To achieve this, we fix the number of alternatives (J) at 10. 5,000 individuals are generated to choose among the J alternatives over 5 choice occasions (T) with the individuals randomly assigned to 15 income categories (B). The *demographics* variable $\boldsymbol{D}_i$ is randomly generated from a uniform distribution and Bernoulli distribution with equal probability of success and failure.

We specify the residual income function as a natural log function $ln(y_i - P_{ij})$ which implies a diminishing marginal utility of Income. We assume this simple form since it is consistent with economic theory and the first derivative can be easily calculated to know the true marginal utility of income though other complicated functions can also be used. The income values (y_i) ranges from 1.075 to 3 while the travel costs for each individual to each site is set to be generated from a standard uniform distribution.

The alternative specific constant for a give site is drawn from a normal distribution with mean α_0 and variance σ_α^2. We also fix γ and the remaining *common* parameters of the model which are used to sample the individual random effects and the latent utility values. The latent variables are then mapped to the *observed* site selection by each individual.

The sampler described in section 3 is implemented for 22,000 iterations with 2,000 as burn-in. Different starting values were used to check for convergence. The summary of the distribution of the parameters also show that the algorithm performed well in recovering the parameters of the generated data experiment as reported in Table 1. We find that the posterior means are quite close to the true values. The posterior standard deviations are also reported and the posterior means are typically within a posterior standard deviation when the actual and estimated posterior means diverges. Trace plots using the first prior for selected parameters are reported in Figure 1.

In addition, we also calculated the inefficiency factors for the posterior distribution of the parameters as the ratio of the numerical standard error with correlated draws (NSE) over the standard error of the parameter. That is:

$$\sqrt{\text{inefficiency factor}} \equiv \sqrt{f} = \sqrt{1 + 2\sum_{j=1}^{m-1}(1 - \frac{j}{m})\rho_j}, \tag{33}$$

where m is the number of draws after convergence and ρ_j is the autocorrelation coefficient which is a correlation between draws as a function of j time separation between them. The estimates of these inefficiency factors are quite high especially for the coefficient on marginal utility of income for the generated data experiment are relatively high. Most of the parameters in θ have values around 5 except for that of the variance (σ_φ^2) which is the highest with a value of about 8. Once convergence is reached, the sampler does not seem to move much as shown in the trace plot hence the high autocorrelation coefficient and inefficiency factors. The high inefficiency factors shows that we will need to run our simulation longer to get the same level of accuracy as will be in m *iid* draws (for instance we will need about 65,000 draws to get 1000 *iid* sample).

In summary, the generated data experiment shows that our sampling scheme works well in recovering the parameters of the model including that of the marginal utility of income by income category. The first series Taylor approximation approach too seems to work well for the particular functional form assumed.

Data Description and Application

The data set we use to illustrate the methodology is from the Iowa Lakes Valuation Project at Iowa State University, a four year study of recreational lake usage in the state. The data is appropriate for our analysis in a number of ways. First, Income brackets data was collected for each household which allows for the estimation of the marginal utility of income by income group. For households that did not respond to the income bracket question, an imputed income of $56,000 (average income level using midpoint) was used and they were categorized as a separate income group. 375 households fall into this category. This will help to check if the behavior of the item nonrespondents is similar to any of the household that reported income values. There is also information on the round trip travel cost to all the 130 sites in the state. Detailed description of the data and how other variables were calculated is presented in chapter one.

Empirical Results

The application of the methodology discussed in section 3 to the Iowa Lakes Data seeks to address the following questions: (1) Does the assumption of constant marginal utility of income hold in the demand for lakes in Iowa across household? (2) Are there distributional differences in marginal utility of income and welfare derived from environmental policy shift between income categories in Iowa?

Estimation

For each of our prior on the marginal utility of income, we implement the algorithm described in Section 3. The posterior simulator is run for 100,000 iterations, discarding the first 25,000 as burn-in. Results from these runs suggested that the chain mixed reasonably well and appeared to converge after about 10,000 iteration for all parameters.

Presented in Table 2 and 3 are posterior means, posterior standard deviation and posterior probabilities of being positive [denoted $P(\cdot > 0|y)$] associated with the parameters of the marginal utility of income, demographic variables and selected site-specific parameters of our model for each type of prior. Before discussing the results for the marginal utility of income, we discuss results for the other parameters. These results are generally consistent with the previous chapter and will not be discussed at length. The estimates for the alternative specific constant is as expected with West Okoboji Lake, one of the best lakes in Iowa, having the highest posterior mean. In general, the α_j's are higher for popular destination lakes in the state as is the case in chapter one with the values closer higher though negative. The scale of the parameters in the two chapters are relatively the same. The posterior distribution for the demographic variables are also consistent with previous work on this topic: individuals with less education, females and older are more likely to stay at home. Somewhat surprisingly, households with greater number of adults and children are more likely to stay at home (though the evidence for the effect for adults is not strong and significant).[7]

The results from using the two priors are generally consistent with our prior expectations. Specifically, there is no evidence that the assumption of constant marginal utility of income holds across the income categories. We present a plot of the values for the posterior mean using results from the smoothing prior in Figure 2. For clarity of presentation, standard error bands are not included within the figure but we present probability that the difference between marginal utility of income between two income categories is greater than zero in Table 4. The figure suggests several important results. First, there is evidence of diminishing marginal utility

[7] Demographic variables in this chapter was recoded differently from that in chapter one which made the estimates on the alternative specific constant and demographic variables different. For example, schooling was recoded from the zero to five categorization to a zero-one variable (1 been individuals with at least some college) and midpoint values were used for age instead of the zero to seven categorization.

of income across households in our sample. Except for households with income of $12,500 (group 2) and $67,500 (group 11), we find a declining pattern in the values of the marginal utility of income as income increases. Second, there is no sufficient evidence against constant marginal utility of income across some income groups. Table 4 reports $P((\theta_d - \theta_c) > 0|y)$ where a negative sign implies that the mass of the distribution for income category d is at least higher than that of income category c at least 90% of the time (Assuming Income level for group c is higher than group d). Blank sign in the upper triangular matrix implies the difference in the marginal utility of income is not significantly different (difference in the distribution overlap between the two distribution). To explain this in the context of our application, households in the income group with $37,500 have a marginal utility of income which is not different from those with income levels of $45,000, $55,000 and $56,000, but higher 90% of the time than other categories with higher income levels (>$102,500). Thus, depending on the household income categorization, marginal utility of income is constant for a range of income level and diminishing across other groups. In general, marginal utility of income is nonincreasing except for about two outliers.

Looking at the smoothing parameter in the second prior, we have a posterior mean of 0.0869 and posterior standard deviation of 0.0269. Comparing this to our choice of prior with $a_\eta = 5$ and $b_\eta = 1$, this choice of the parameters of the prior sets the prior mean and prior stand deviation equal to 0.25. The result shows that the data has moved our prior towards smaller values (posterior mean is about one-third of the prior mean and posterior standard deviation that is about one-tenth the magnitude of the prior standard deviation) hence linearity despite the posterior mean suggesting nonlinearity. A useful exercise in this case will be to change the prior hyperparameters for the smoothing parameter to keep track of when the data move our prior toward larger values of η which will signal nonlinearity.

We also estimated the model assuming constant marginal utility of income. This is equivalent to estimating the model in chapter one with modifications to the data. The posterior mean for the marginal utility of income is estimated to be 0.019 which is close to the posterior mean in Table 2 assuming Hierarchical prior on $f'(y)$. Inference made using this estimate of marginal utility on income will be misleading since income level varies by household and policy targeted to all households may not have the desired results based on misrepresentation. This will be especially important for welfare analysis in evaluating different policy changes. We will explore this in detail in the next section.

Welfare Analysis

Bayesian methodology has an advantage over other methods in conducting posterior inference and prediction. Estimating the welfare impact of changes in policy scenarios will be particularly informative to examine if introducing nonlinearity into RUM models matter. The methodology presented follows closely what is presented in chapter 1 though our focus is primarily on the loss of a site.

Using similar notation as in chapter two, we let Υ^s_{it} denote the maximum utility achieved by agent i on choice occasion t under scenario s $(s = 0, 1)$. That is,

$$\Upsilon^s_{it}(\boldsymbol{\Xi}_{-\boldsymbol{\alpha}.}) = \max_{j \in J^s}(U^s_{ijt} | \boldsymbol{\Xi}_{-\boldsymbol{\alpha}.}) \quad s = 0, 1, \tag{34}$$

where $\boldsymbol{\alpha}. = (\alpha_1, \ldots, \alpha_J)$ denotes the vector of alternative specific constants where J^s denotes the index of valuable sites under scenario s. However, since we assume no correlation between the sites in terms of attributes, closure of a site is primarily simulating the alternative choice of the household among the remaining $J - 1$ alternatives. We use this information to simulate the observed site choice in the new scenario and compare this to the base scenario which is the *status quo*, with all the sites included in the choice set.

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

$$CV = -\frac{T}{\theta_b I(y_i \in Y_b)} E\left[\Upsilon^0_{it}(\boldsymbol{\Xi}_{-\boldsymbol{\alpha}.}) - \Upsilon^1_{it}(\boldsymbol{\Xi}_{-\boldsymbol{\alpha}.})\right] \tag{35}$$

Thus, welfare analysis following this model specification is similar to what exists in the literature for constant marginal utility of income except that the marginal utility of income differs by income bracket.[8]

[8]One thing to note is the case where the change in policy makes agents move from one income bracket to

Thus Step 2, as described in chapter 2, can be re-written as:

Step 2: Draws of the utility levels $U_{ijt}^{s(r)}|\boldsymbol{\Xi}_{-\boldsymbol{\alpha}}$. are obtained using equation above, with

$$U_{ijt}^{s(r)} = \begin{cases} \boldsymbol{D}_i\gamma^{(r)} + \varepsilon_{i0t}^{(r)} & j = 0 \\ \alpha_j^s + p_{ij}\sum_{b=1}^{B}\theta_b^{(r)}I(y_i \in Y_b) + \varphi_i^{(r)} + \varepsilon_{ijt}^{(r)} & j = 1, \ldots, J. \end{cases} \tag{36}$$

where

$$\epsilon_{ijt}^{(r)} \sim \mathcal{N}(0,1) \text{ and } \quad \varphi_i^{(r)} \sim \mathcal{N}(0, \sigma^2_{\varphi(r)}). \tag{37}$$

The result using this scenario is presented in Table V for closure of the West Okoboji lake. We present the result by different income groups showing that the impact of closure of the site has varied distributional impacts depending on the income category. This can be informative to policy makers considering how a particular policy will affect different income groups. Closure of Okoboji has higher impact on the high income groups than the low income. This is in contrast to the model using constant marginal utility of income. Though the average CV estimates for the whole sample is not different from the model incorporating nonlinear income effect, there is a form of reversal in the CV estimates across income groups. When we assume constant marginal utility of income, the closure of West Okoboji has a higher CV values for low income groups than high income groups which is the opposite for the other two models incorporating nonlinear income effects.

Future Directions

One future direction in terms of estimating the marginal utility of income function will be to actually impose diminishing marginal utility of income on the parameters. This procedure will follow Geweke (1996) on Gibbs sampling in a regression model with inequality constraints. This will involve reparameterizing the marginal utility of income parameters and the difference between the parameters for high and low income to be truncated between zero and infinity.

Also, checking the sensitivity of our result to the number of brackets chosen might also be worth exploring. This can be achieved by extending the generated data experiment. Also,

the other. If this occurs the CV estimate will have to be adjusted to account for this. The algorithm presented here however ignores that possibility.

generating the agents in each bracket by first creating a continuous function of the relationship between utility and income (continuous) before grouping the agents by income brackets might be insightful. This way we do not impose on the data *a priori* that only income bracket data is available to the researcher.

Another future direction is related to deriving welfare estimate for the Taylor series approximation model. The nature of the income bracket data offers the possibility of households switching groups depending on the scenario and how large the income group is. This possibility can cause problem on which marginal utility of income parameter to use to calculate the CV estimate. The full semi-parametric estimation model might have an advantage over the Taylor series approach in this regards since we will have an estimate of the function relating utility and income rather than just the slope. Implementing the full semi-parametric model will also be instructive to compare the results with the Taylor series proposed in this paper.

Summary and Conclusion

In this chapter, an empirical methodology to incorporate nonlinear income effect in RUM models is introduced. RUM models is a major tool in estimating site visitation patterns of households, and out of convenience, researchers typically assume constant marginal utility of income in the analysis. The methodology introduced in this chapter is a case of hierarchical RUM model which allowed for nonlinear income effect following a Taylor series approximation of the functional relationship between utility and residual income.

We used this methodology to investigate the visitation patterns of Iowans to 130 major lakes in the state and estimate the demand for the sites. We also estimated the marginal utility of income across households by income groups showing that marginal utility of income is not constant across income group as is assumed in practice. Other results are in general consistent with previous studies in the literature.

Our methodology also enabled us to test if the differences in the estimates of marginal utility of income is significant and to make posterior inferences on the impact of scenario changes. Interestingly, the impact on the household of exogenous changes in the site availability differs by income groups and may affect the results of the policy. We find counter intuitive result for the welfare estimates when constant marginal utility of income is assumed. As expected the mean CV is not significantly different in all the models.

Our hope is that this methodology can be extended and applied in other applications of the RUM model. The fact that specific amount of income is not required and the simplicity of the model should make it appealing to practitioners. One relatively easy extension will be to impose diminishing marginal utility of income assumption instead of the two priors we used here. Another extension will be the use of a semi parametric approach in estimating the function instead of a Taylor series approximation. One problem with this approach is the fact that in most surveys, actual level of income is not known and the methodology can be highly intensive especially in the case of the repeated choice model.

Tables and Figures

Table 1: Generated Data Experiment Result for Taylor series Model

	Hierachical Prior on $f'(y)$		Smoothing Prior on $f'(y)$		
	Posterior mean	Posterior SD	posterior mean	posterior SD	True value
θ_1	0.8931	0.0563	0.8919	0.0428	0.9302
θ_2	0.7932	0.0571	0.8633	0.0310	0.8889
θ_3	0.9040	0.0590	0.8358	0.0276	0.8511
θ_4	0.8418	0.0506	0.8005	0.0271	0.8163
θ_5	0.7400	0.0530	0.7601	0.0271	0.7843
θ_6	0.6663	0.0538	0.7230	0.0277	0.7547
θ_7	0.7124	0.0505	0.6956	0.0284	0.7273
θ_8	0.6327	0.0532	0.6652	0.0288	0.6897
θ_9	0.6116	0.0561	0.6412	0.0296	0.6452
θ_{10}	0.6803	0.0561	0.6407	0.0296	0.6442
θ_{11}	0.6558	0.0506	0.6130	0.0334	0.5970
θ_{12}	0.5128	0.0546	0.5386	0.0400	0.5333
θ_{13}	0.4925	0.0590	0.4729	0.0413	0.4706
θ_{14}	0.3976	0.0550	0.4031	0.0448	0.4211
θ_{15}	0.3035	0.0590	0.3000	0.0537	0.3333
η			0.1581	0.0660	
σ_θ^2	0.1120	0.0358			
θ_0	0.6517	0.0871			0.6725

Table 2 : Posterior Means of Site-Specific Parameters and the rest of the model parameters[9]

Parameter	Parameter					
	Hierarchical Prior on $f'(y)$		Smoothing Prior on $f'(y)$		Constant MUI	
	Mean	$P(\cdot > 0 \mid y)$	Mean	$P(\cdot > 0 \mid y)$	Mean	$P(\cdot > 0 \mid y)$
α_j	Site-Specific Parameters					
Storm Lake	2.142	0.9658	2.0769	0.9411	2.2529	0.9792
Briggs Woods Lake	0.9226	0.8330	0.7833	0.7847	1.1325	0.9119
Prairie Rose Lake	1.1042	0.863	0.9404	0.8063	1.3492	0.9412
Big Creek Lake	1.8769	0.9478	1.7894	0.9126	2.0686	0.9756
West Okoboji Lake	2.9982	0.985	3.0265	0.9862	2.9874	0.9913
Saylorville	2.1396	0.9657	2.0711	0.9382	2.3231	0.9824
	Other Parameters					
α_0	1.0273	0.8521	0.8594	0.7898	1.2514	0.9290
γ_{01} (Age)	0.2788	1.000	0.3940	1.000	0.2519	1.0000
γ_{02} (Gender)	-0.7209	0.1586	-0.4257	0.2893	-2.2316	0.0009
γ_{03} (Schooling)	-0.4224	0.2765	-1.7626	0.0219	4.6239	1.0000
γ_{04} (Adults)	0.3570	0.7362	0.6533	0.8189	0.8345	0.9193
γ_{05} (Children)	1.1543	0.9751	1.4448	0.9779	1.5497	0.9914

We only report the result for selected zonal lakes including W. Okoboji for brevity.

Table 3: Result for Taylor series Model (Iowa Lakes Data)

	Hierachical Prior on $f'(y)$		Smoothing Prior on $f'(y)$	
	Posterior mean	Posterior SD	posterior mean	posterior SD
θ_1	0.0358	0.0060	0.0396	0.0070
θ_2	0.0281	0.0059	0.0296	0.0062
θ_3	0.0434	0.0088	0.0496	0.0113
θ_4	0.0438	0.0081	0.0498	0.0099
θ_5	0.0356	0.0079	0.0395	0.0089
θ_6	0.0301	0.0072	0.0345	0.0080
θ_7	0.0248	0.0058	0.0266	0.0062
θ_8	0.0259	0.0062	0.0288	0.0065
θ_9	0.0239	0.0061	0.0277	0.0071
θ_{10}	0.0229	0.0044	0.0245	0.0047
θ_{11}	0.0141	0.0031	0.0146	0.0032
θ_{12}	0.0169	0.0044	0.0184	0.0047
θ_{13}	0.0195	0.0045	0.0212	0.0048
θ_{14}	0.0130	0.0030	0.0140	0.0034
θ_{15}	0.0107	0.0030	0.0119	0.0034
η			0.0869	0.0269
Σ_θ	0.0456	0.0145		
θ_0	0.0223	0.0396		

Table 4: Probability that the difference between marginal utility of income between group c and d is greater than zero using model with smoothing prior

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1		-				-	-	-	-	-	-	-	-	-	-
2			+	+	+	+									-
3							-	-	-		-	-	-	-	-
4					-	-	-	-	-	-	-	-	-	-	-
5						-	-	-	-		-	-	-	-	-
6							-	-	-		-	-	-	-	-
7											-		-	-	-
8											-		-	-	-
9											-	-	-	-	-
10											-	-	-	-	-
11												+			-
12														-	-
13														-	-
14															-

Table 5: CV estimates for closure of Okoboji Lake by income group

Income Group	CV (Hierarchical Prior)	CV (Smoothing Prior)	CV (Constant MUI)
5000	-12.12	-11.64	-32.64
12,500	-19.18	-20.19	-32.75
17,500	-9.700	-8.90	-33.72
22,500	-7.410	-6.77	-28.13
27,500	-12.76	-12.11	-32.26
32,500	-15.05	-13.22	-28.86
37,500	-13.65	-13.63	-18.74
45,000	-22.47	-21.26	-30.68
55,000	-20.14	-17.93	-24.43
56,000 (inputed)	-21.73	-22.34	-24.85
67,500	-40.44	-44.50	-18.23
87,500	-22.91	-22.70	-14.66
112,500	-16.20	-15.95	-14.17
137,500	-30.82	-32.02	-11.18
>150,000	-32.12	-30.92	-8.74
meanCV	-20.90	-20.81	-23.89

Figure 1 Trace Plots of Selected Parameters $(\alpha_1, \gamma_1, \theta_1)$- Generated Data Experiment with prior 1

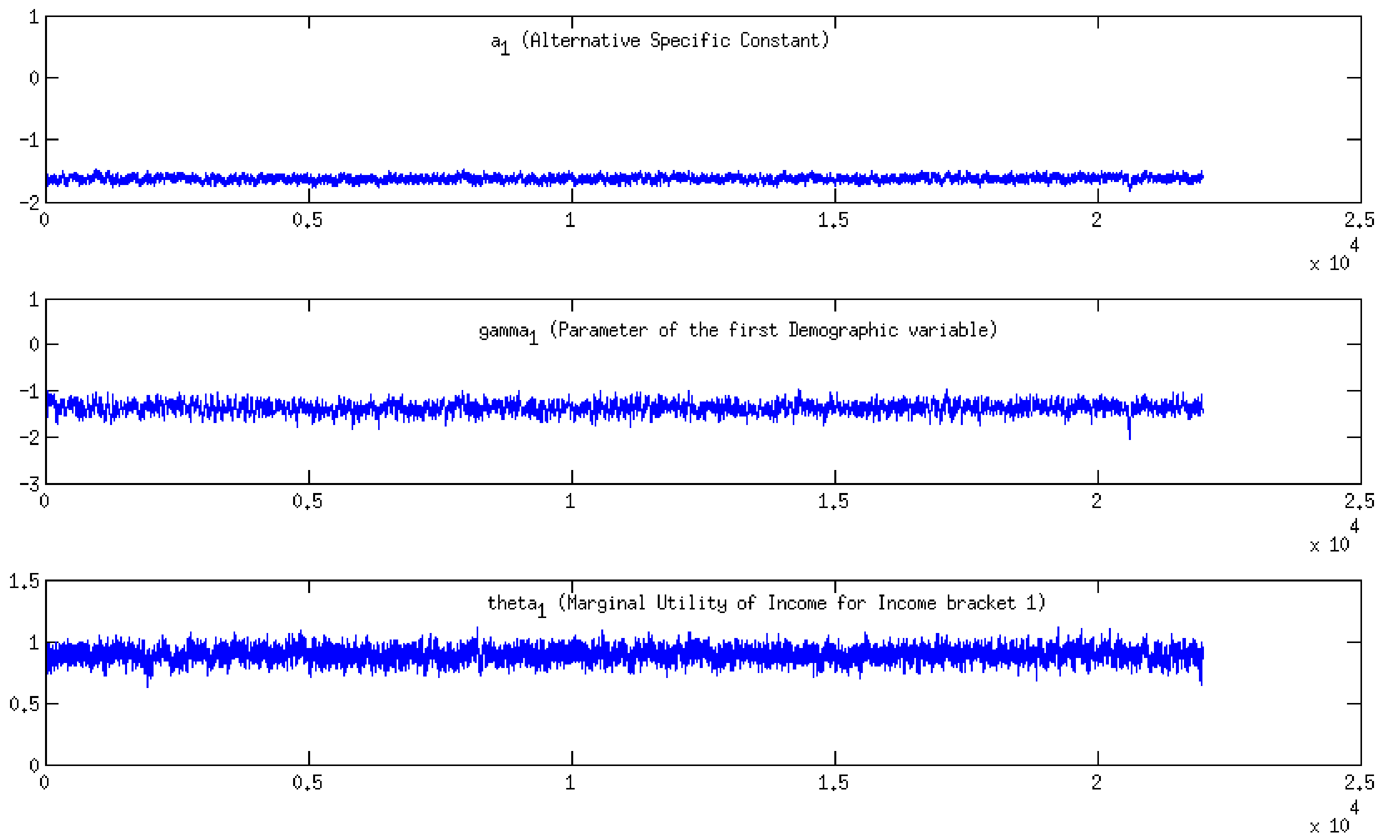
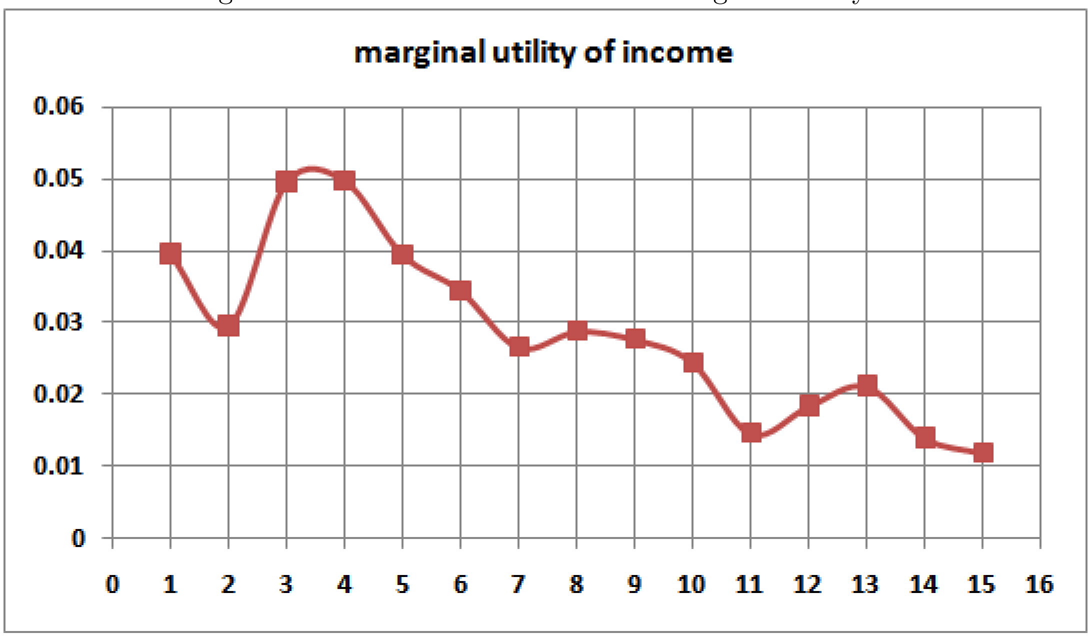


Figure 2 Plot of Posterior Mean of Marginal Utility of Income



Model Uncertainty and Recreation Demand

A paper to be submitted to American Journal of Agricultural Economics

Babatunde Abidoye

Abstract

Bayesian variable selection procedure is used to model uncertainty in recreational demand's model specification. In contrast to comparing models based on the likelihood values as in Egan *et al* with unknown sampling properties, we propose a model that draws on the Bayesian paradigm to integrate the variable selection process into the model and reflect the accompanying uncertainty about which is the "best" specification used for counterfactual predictions.

Introduction

Analysts and policymakers are typically interested in understanding the impact that changing environmental conditions has on the demand for recreational alternatives and quantifying the welfare implications of these changes. Policy makers, for example, use this information to appropriately direct resources aimed at maintaining and restoring environmental quality. However, this requires the specification of a functional relationship between individual demand and observable individual and site characteristics. Unfortunately, economic theory provides relatively little guidance regarding the form that this relationship should take and which variables ought to be included in the analysis. In many applications, limitations in the available data (e.g., describing the water quality conditions at a lake site) narrow the range of possibilities, but choices must still be made between, for example, level and logarithmic specifications for an environmental characteristic. This can lead to different posterior inference and spurious parameter estimates depending on the variable choice by the researcher.

While model selection criterion can be used to narrow the set of specifications, there is the risk that the analyst (even inadvertently) may engage in a "fishing" process among the available models, biasing the final outcome of the analysis. In a recent paper, Egan *et al.* (2009) attempt to ameliorate this problem by employing a split sample approach, using separate portions of the available data for model specification, estimation and evaluation. They isolated one third of their sample in order to consider alternative models and functional forms, using a likelihood dominance criteria to pick their final model, which is in turn estimated using a separate sample. Though this can arguably reduce the impact of the specification search process on the final parameter estimates, it does not eliminate the problem. More importantly, the procedure inevitably requires the selection of a single model and does not account for the uncertainty in this process. Indeed, the selection of the final model is not based on a test among competing models (as the alternatives are non-nested), but on a log-likelihood based ranking.

In this paper, we consider an alternative approach that draws on the Bayesian paradigm to integrate the variable selection process into the model and to reflect the accompanying uncertainty about which is the "correct" specification into subsequent counterfactual predictions. Specifically, we describe a Bayesian posterior simulator that combines the literature on hierarchical modeling, Bayesian variable selection and data augmentation. Our underlying modeling framework is the class of repeated random utility models (See Herriges and Phaneuf (2002) and Herriges, Kling and Phaneuf (2002) for a review) and follows closely the model proposed in the first paper. We then employ the stochastic search variable selection (SSVS) method described in George and McCulloch (1993) to determine the posterior probability that individual site characteristics influence the site selection decision. The model can be used to identify a preferred model specification. Alternatively, and we would argue preferably, the model can be used as part of the process of employing a Bayesian model averaging, integrating competing models into a single structure that can be used for welfare analysis and counterfactual predictions.

The model is applied using data from the 2002 survey of Iowa Lakes Project, the same data underlying the Egan *et al.* (2009) analysis. We use our model to contrast our findings with those obtained in Egan *et al.* (2009), highlighting the benefits of integrating model uncertainty into a unified framework. One advantage of this study is that we have large number of sites (130) and detailed information on both site attributes and lake water quality.

The outline of the chapter is as follows. Section 2 touches on the issue of model uncertainty in econometric analysis and also frames our approach in the context of other methods in the literature. Section 3 presents the model and how the parameters of interest are estimated. Section 4 describes a generated data experiment as a check for the performance of the sampler. Section 5 describes the data and application and section 6 provides posterior simulation and welfare analysis. The chapter concludes with a summary in section 7.

Related Literature

Model uncertainty

Researchers are often faced with the dilemma of which model specification or subset of explanatory variables will best fit their data. This problem is more pronounced in situations where economic theory does not dictate *a priori* the specific functional form or distributional assumption to be used. The inability to lay claim to a "best" model makes inference on the chosen model less certain and potentially inaccurate. This has led to widespread criticism of estimates presented for a "best" model Leamer (1983)) - for example, changing from linear to nonlinear specification or changing the functional form of some variables can lead to different estimates. A number of studies including Regal and Hook (1991) and Draper (1995) have shown the impact of ignoring uncertainty of the model on inference.

Various techniques has been proposed in the literature to account for this problem. The paper by Raftery (1995) among others have argued that the use of *p-values*, R^2 and other statistical tests based on them to search for the "best" model can lead to misleading inference and prediction. Poirier (1995) also has a discussion on problems with using hypothesis testing to select a specific model especially given that the procedure of pretesting introduces a level of uncertainty into the pretest estimator. Aside from the problem of choosing the significance level and balancing it with the power of the alternative hypothesis, most studies involves comparing more than two models. The sampling properties of the popular stepwise regression are usually unknown and making inference based on a model selected in this way can be misleading.

A solution to the hypothesis testing problem that has gained popularity among researchers is the use of Bayesian model selection and/or averaging. Bayesian model selection methods are used to select a model(s) with maximum posterior probabilities conditional on the data. Bayesian Model averaging (BMA) on the other hand employs the rules of conditional probability to estimate posterior probabilities of possible models that are used as weights in averaging over all possible models. The enormous number of possible explanatory variables and non-linearity makes the use of model selection important for reducing the size of possible models before averaging among the most probable models. Variable selection methods can also be used to select a specific model Raftery (1995). There are a number of papers in the literature that have applied BMA in economics.[1] In the environmental and resource literature, some of the papers include Clyde (2000), Clyde, Guttorp and Sullivan (2000), Koop and Tole (2004), Layton and Lee (2006), Fernandez, Ley and Steel (2002), and Leon and Leon (2003). The message of all these papers is that model uncertainty can have a big impact on parameter estimates and should be accounted for explicitly.

For problems related to uncertainty regarding predictors in a model (which is our focus in this chapter), the stochastic search variable selection (SSVS) method proposed by George and McCulloch (1993) provides an insightful and easy to implement approach to solving the problem. The model works by capturing the entire possible model setup in a hierarchical Bayes mixture model with the use of latent variables to identify the models supported by the data. The latent variable is used to nest all of the possible models. The number of visits to a particular model through the iterative sampling (Gibbs) process determines how promising the model is. SSVS makes use of both practical and statistical relevance of the model to select the "best" possible models. One major advantage of the SSVS approach especially in the Bayesian

[1] There are number of websites that are devoted to posting developments and research in this area. See http://www.research.att.com/~volinsky/bma.html for some of the papers and software.

framework is that the researcher does not have to calculate the marginal likelihoods for each of the possible models.[2]

Model uncertainty in recreation demand

One major reason in estimating the demand for recreational sites is to quantify how site attributes (especially water quality attributes) influence demand for these sites. This is essential for policy analysis and budget purposes.[3] However, economic theory provides little or no guidance as to which characteristics should be in the model and subsequent welfare analysis. Few studies have been done that incorporate this uncertainty in the model. Layton and Lee (2006) using a stated preference (SP) survey of saltwater angling in Alaska applied the procedure suggested by Buckland, Burnham and Augustin (1997) to control for model uncertainty. They estimated weights for different model specifications and used those weights to calculate the expected willingness to pay. One problem with this procedure is that model uncertainty is incorporated *ex post* and it does not account for uncertainty in the estimates of the parameters of the model. In a recent paper, Egan *et al.* (2009), in an attempt to search for the best model specifications estimate different combinations of variable specification and parameter distribution assumptions. They compared different models based on the likelihood values from different combinations of water quality measures applied to one third sample split of their data. The "best" model, chosen based on the likelihood dominance criteria, was then re-estimated using the second third of the sample.[4] While this does reduce the "fishing" problem, it still does not incorporate uncertainty in the final model estimated. The model selection process is not based on a test among competing model but on a log-likelihood based ranking which with "tight" rankings given that the log-likelihood values were very close. To limit the number of models to be compared, Egan *et al.* (2009) assumed that all the water quality variables are to be included in the model and the choice is if they are suppose to enter linearly or logarithmic fashion. This limits the number of models they compare to 32.[5] Their preferred model

[2]The marginal likelihood defined as $P(Y|m=j) = \int \mathbf{P}(Y|\theta_j)\mathbf{P}(\theta_j)d\theta_j$ are often difficult to estimate

[3]These estimates are used to justify important environmental policies such as pollution abatement programs.

[4]The final third was used to assess out-of-sample predictions

[5]In addition, they also grouped the water quality attributes to five to limit the number of models to be compared

has Secchi depth and suspended solids entering the model linearly, with the remaining water quality variables entering in a logarithmic fashion.

In this chapter, we present a Random Utility Maximization (RUM) model that incorporates model uncertainty on the specification of site attributes in recreation demand. Specifically we apply the SSVS algorithm to identify the probability that a model is supported by the data.

Model

As described in the previous sections, the model we present in this study will incorporate model uncertainty in the site characteristics attributes in recreation demand. In addition, we want our model to be relatively flexible for posterior inference including welfare analysis. For the purpose of our model, we index individuals by $i = 1, 2, \ldots, N$, choice occasions by $t = 1, 2, \ldots, T$ and sites by $j = 1, 2, \ldots, J$.

Basic Structure

The model is similar to the repeated nested logit model (Morey, Rowe and Watson (1993)) and repeated mixed logit model (Herriges and Phaneuf (2002)). These models integrate individuals' choice among alternatives and the problem of allocating time between multiple recreation sites. The model of Morey, Rowe and Watson (1993) assumes that individuals face the decision to participate in recreation activities over fixed discrete occasions and at most one trip is taken at such an occasion. Furthermore, each decision is assumed conditionally independent across individuals and choice occasions. A summary of this framework and implications of the assumptions is presented in Herriges , Kling and Phaneuf (1999). Our model is similar to that in chapter two with the incorporation of uncertainty at the level specifying the functional form of the site characteristics.

Formally, we assume that an individual i at choice occasion t has to choose among J sites and also inactivity, or "staying at home." We represent the utility that an individual derives from making a particular choice at a given time as:

$$U_{ijt} = \begin{cases} D_i\gamma + \varepsilon_{ijt} & \text{if } j = \text{stay at home } (0) \\ \alpha_j + P_{ij}\beta + \varphi_i + \varepsilon_{ijt} & \text{for } j = 1, \ldots, J. \end{cases} \tag{1}$$

where α_j is the overall site-specific effect; β is the marginal utility of income; φ_i captures the individual specific effect and ε_{ijt} represents an idiosyncratic error that is assumed to be independent across the $J+1$ alternatives with variance normalized such that $\varepsilon_{ijt} \sim N(0,1)$. We also assume that the demographic characteristics of an agent (D_i) has an effect on the likelihood of choosing the "stay at home" option and not visiting any of the recreation sites.

Given that it is the difference in utility that matters, we have the base case to be the stay at home option and take the difference in utilities. Thus,

$$\tilde{U}_{ijt} = \alpha_j + P_{ij}\beta - D_i\gamma + \varphi_i + \tilde{\varepsilon}_{ijt} \tag{2}$$

where $\tilde{U}_{ijt} = U_{ijt} - U_{i0t}$; $\tilde{\varepsilon}_{ijt} = \varepsilon_{ijt} - \varepsilon_{i0t}$; for $j = 1, ..., J$. So that

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

where

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

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

$$y_{it}(\tilde{\boldsymbol{U}}_{\boldsymbol{i\cdot t}}) = \begin{cases} 0 \text{ if } \max\{\tilde{U}_{ijt}\}_{j=1}^{J} \leq 0 \\ k \text{ if } \max\{\tilde{U}_{ijt}\}_{j=1}^{J} = \tilde{U}_{ikt} > 0. \end{cases} \tag{3}$$

Stacking over the alternatives, we have:

$$\tilde{\boldsymbol{U}}_{\boldsymbol{i.t}} = \boldsymbol{\alpha} + \boldsymbol{P}_{\boldsymbol{i.}}\boldsymbol{\beta} - (\mathbf{1}_{\boldsymbol{J}} \otimes D_i)\gamma + \mathbf{1}_{\boldsymbol{J}}\varphi_i + \tilde{\boldsymbol{\varepsilon}}_{\boldsymbol{i.t}}. \tag{4}$$

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

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

We can then re-write the above equation concisely as

$$\boldsymbol{\tilde{U}_{i.t}} = \boldsymbol{M_{i.t}\theta} + \mathbf{1_J}\varphi_i + \boldsymbol{\tilde{\varepsilon}_{i.t}} \tag{5}$$

where

$$\boldsymbol{M_{i.t}} = \begin{bmatrix} \boldsymbol{I_J} & \boldsymbol{P_{i.}} & \mathbf{1_J} \otimes D_i \end{bmatrix}; \boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\alpha}' & \boldsymbol{\beta}' & \gamma' \end{bmatrix}'.$$

Another way to write equation (5) is in terms of the error component. That is:

$$\boldsymbol{\tilde{U}_{i.t}} = \boldsymbol{M_{i.t}\theta} + \boldsymbol{v_{i.t}}$$

where

$$\boldsymbol{v_{i.t}} = \mathbf{1_J}\varphi_i + \boldsymbol{\tilde{\varepsilon}_{i.t}}$$

$$E(\boldsymbol{v_{i.t}v_{i.t}}') \equiv \boldsymbol{\Omega} = \sigma^2_\varphi \mathbf{1_J 1'_J} + \boldsymbol{\Sigma}^*.$$

Hierarchical Priors

As described earlier, the α_j's captures the overall site-specific effect. Given that these depend on the characteristics of the site, we specify an hierarchical prior on α_j with the assumption that its mean is the aggregate effect of the observed attributes and the unobserved site characteristics signals deviation from the mean.

The priors for the site-specific parameters is specified as:

$$\alpha_j \sim N(\boldsymbol{Q_j \alpha_0}, \sigma^2_\alpha). \quad j = 1, 2, ...J \tag{6}$$

where $\boldsymbol{Q_j}$ includes a constant term and the observed site characteristics that influence demand for site j.

To control for model uncertainty at the site characteristics level, we will focus on the parameters of all K possible model specification of combinations of observed site attributes ($\alpha_{0,k}$) for $k = 1, 2, \ldots, K$. Thus, we will seek to calculate the probability that a model specification or subset of variables belong to the model using the SSVS approach. If a variable k is not supported by the data, we will expect that the true value of the parameter ($\alpha_{0,k}$) be zero. To capture this we introduce an additional level to the hierarchical structure described above. Specifically, we specify a prior for each regression coefficients ($\alpha_{0,k}$) as a mixture of two normal distributions with different variances and zero mean. That is conditional on a binary latent variable $\lambda_k = 0$ or 1, each k element of α_0 can be defined as:

$$\alpha_{0,k}|\lambda_k \sim (1 - \lambda_k)N(0, \tau_k^2) + \lambda_k N(0, c_k^2 \tau_k^2)) \tag{7}$$

and

$$P(\lambda_k = 1) = 1 - P(\lambda_k = 0) = p_k; \quad 0 \le p_k \le 1. \tag{8}$$

λ_k is a binary latent variable that indicates if the observed site characteristics is supported by the data or not. With the above representation, when $\lambda_k = 0$, $\alpha_{0,k} \sim N(0, \tau_k^2)$, whereas $\alpha_{0,k} \sim N(0, c_k^2 \tau_k^2)$ when $\lambda_k = 1$. The variance term for the first normal distribution (τ_k^2) is assumed to be very small such that the distribution of the $\alpha_{0,k}$ is massed around zero and little evidence for its inclusion in the model. The second variance ($c_k^2 \tau_k^2$) on the other hand is large and signals evidence that the variable should be included in the model. p_k can be thought of as the prior probability that $\alpha_{0,k}$ should be included in the model. Thus, the prior on α_0 is represented as multivariate normal:

$$\boldsymbol{\alpha_0}|\boldsymbol{\lambda} \sim N_k(\mathbf{0}, \boldsymbol{D_\lambda V_\alpha D_\lambda}) \tag{9}$$

where $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_K)$, V_α is the prior correlation matrix and $\boldsymbol{D_\lambda} \equiv diag[L_1\tau_1, ..., L_K\tau_K]$, with $L_k = 1$ if $\lambda_k = 0$ and $L_k = c_k$ if $\lambda_k = 1$. $\boldsymbol{D_\lambda}$ is like a tuning parameter that ensures that the prior on $\alpha_{0,k}$ holds.

Finally, we set priors for the other parameters as

$$p(\boldsymbol{\lambda}) = \prod_{k=1}^{K} p_k^{\lambda_k}(1-p_k)^{1-\lambda_k} \tag{10}$$

$$\sigma_\alpha^2 \sim IG(a_\alpha, b_\alpha) \tag{11}$$

$$\sigma_\varphi^2 \sim IG(a_\varphi, b_\varphi) \tag{12}$$

$$\boldsymbol{\gamma} \sim N(\boldsymbol{\mu_\gamma}, \boldsymbol{V_\gamma}). \tag{13}$$

Posterior Simulator

The posterior simulator uses the Gibbs sampler to generate draws from the posterior conditional distribution. We will derive the posterior conditionals and describe how to make draws from the distributions.

Let

$$\boldsymbol{\Xi} = \begin{bmatrix} \boldsymbol{\theta} & \boldsymbol{\alpha_0} & \boldsymbol{\lambda} & \sigma_\alpha^2 & \boldsymbol{\varphi_.} & \sigma_\varphi^2 \end{bmatrix}$$

denote all the parameters of the model with $\boldsymbol{\varphi_.}$ denoting φ_i stacked over individuals.

The joint posterior distribution of $\boldsymbol{\Xi}$ and the latent utility data $\tilde{\boldsymbol{U}}$ gives us the posterior density for the parameters in our model. We use blocking step (e.g., Chib and Carlin (1999)) to obtain draws from the joint posterior conditional of the individual random effects and the site specific effects to improve the mixing of the sampler. Using Bayes theorem, we can write the posterior density as:

$$\begin{aligned} p(\Xi, \tilde{\boldsymbol{U}}|\boldsymbol{y}) \propto & \prod_{t=1}^{T}\prod_{i=1}^{N} \phi(\tilde{\boldsymbol{U}}_{\boldsymbol{i.t}}, \boldsymbol{M_{i.t}\theta}, \boldsymbol{\Omega}) \\ & \times \left\langle I(y_{i.t}=j)I(\tilde{U}_{ijt} > max[\tilde{U}_{i,-j,t},0]) + I(y_{i.t} \neq j)I(\tilde{U}_{ijt} < max[\tilde{U}_{i,-j,t},0])\right\rangle \\ & \times \left[\prod_{j=1}^{J} p(\alpha_j|\boldsymbol{\alpha}_0, \boldsymbol{\lambda}, \sigma_\alpha^2)\right]\left[\prod_{i=1}^{N} p(\varphi_i|\sigma_\varphi^2)\right] p(\boldsymbol{\alpha_0}|\boldsymbol{\lambda})p(\beta)p(\boldsymbol{\gamma})p(\boldsymbol{\alpha}_0)p(\sigma_\alpha^2)p(\sigma_\varphi^2)p(\boldsymbol{\lambda}). \end{aligned} \tag{14}$$

We outline each posterior conditional distribution below.

Step 1: Draw the hierarchical parameter conditional on the latent utility and the hierarchical prior $(\boldsymbol{\theta}|\boldsymbol{\Xi_{-\theta}}, \tilde{\boldsymbol{U}}, \boldsymbol{y})$ using the results of Lindley and Smith (1972) with blocking step.

The posterior conditional for $\boldsymbol{\theta}$ is given as:

$$\boldsymbol{\theta}|\boldsymbol{\Xi}_{-\boldsymbol{\theta}}, \tilde{\boldsymbol{U}}, \boldsymbol{y} \sim N(\mathbf{D}_\theta \mathbf{d}_\theta, \mathbf{D}_\theta). \tag{15}$$

where

$$\mathbf{D}_\theta \equiv \left[T \sum_{i=1}^{N} \mathbf{M}'_{\mathbf{it}} \boldsymbol{\Omega}^{-\mathbf{1}} \mathbf{M}_{\mathbf{it}} + \boldsymbol{\Sigma}_\theta^{-\mathbf{1}} \right]^{-1}$$
$$\mathbf{d}_\theta \equiv \sum_t \sum_i \mathbf{M}'_{\mathbf{it}} \boldsymbol{\Omega}^{-\mathbf{1}} \mathbf{w}_{\mathbf{it}} + \boldsymbol{\Sigma}_\theta^{-\mathbf{1}} \mu_\theta$$

and

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

Step 2: $\boldsymbol{\alpha}_0|\boldsymbol{\Xi}_{-\boldsymbol{\alpha}_\mathbf{0}}, \tilde{\boldsymbol{U}}, \boldsymbol{y}$

Once we condition on the $\boldsymbol{\alpha}$, the posterior conditional for $\boldsymbol{\alpha}_0$ is similar to that of a linear regression parameter. However, the introduction of the latent variable λ into the model specification helps account for model uncertainty such that less weights are put on specifications not supported by the data.

$$\boldsymbol{\alpha}_0|\boldsymbol{\Xi}_{-\boldsymbol{\alpha}_\mathbf{0}}, \tilde{\boldsymbol{U}}, \boldsymbol{y} \sim N(\mathbf{D}_{\alpha_\mathbf{0}} \mathbf{d}_{\alpha_\mathbf{0}}, \mathbf{D}_{\alpha_\mathbf{0}}) \tag{16}$$

where

$$\mathbf{D}_{\alpha_\mathbf{0}} = (\mathbf{Q}'\mathbf{Q}/\sigma_\alpha^2 + (\boldsymbol{D}_{\boldsymbol{\lambda}} \boldsymbol{V}_{\boldsymbol{\alpha}} \boldsymbol{D}_{\boldsymbol{\lambda}})^{-1})^{-1} \text{ and } \mathbf{d}_{\alpha_\mathbf{0}} = \mathbf{Q}'\boldsymbol{\alpha}/\sigma_\alpha^2 + (\boldsymbol{D}_{\boldsymbol{\lambda}} \boldsymbol{V}_{\boldsymbol{\alpha}} \boldsymbol{D}_{\boldsymbol{\lambda}})^{-1} \boldsymbol{\mu}_{\boldsymbol{\alpha}}.$$

Step 3: $\sigma_\alpha^2|\boldsymbol{\Xi}_{-\sigma_\alpha^2}, \tilde{\boldsymbol{U}}, \boldsymbol{y}$

$$\sigma_\alpha^2|\boldsymbol{\Xi}_{-\sigma_\alpha^2}, \tilde{\boldsymbol{U}}, \boldsymbol{y} \sim IG\left[\frac{J}{2} + a_\alpha, \left(b_\alpha^{-1} + .5 \sum_{j=1}^{J} (\alpha_j - \mathbf{Q}_\mathbf{j} \boldsymbol{\alpha}_\mathbf{0})^2 \right)^{-1} \right]. \tag{17}$$

Step 4: Draw the λ_k

As described earlier, the marginal posterior distribution $p(\boldsymbol{\lambda}|\alpha)$ carries information on the relevance of each model and variable specification. However, since the only link between λ and the alternative specific constants (α) is through the mean parameters α_0, the distribution of λ_k simplifies to a Bernoulli distribution with probability

$$P(\lambda_k|\alpha_0, \boldsymbol{\lambda_{-k}}) = \frac{p(\boldsymbol{\alpha_0}|\boldsymbol{\lambda_{-k}}, \lambda_k = 1)p_k}{p(\boldsymbol{\alpha_0}|\boldsymbol{\lambda_{-k}}, \lambda_k = 0)(1 - p_k)} \tag{18}$$

where $\boldsymbol{\lambda_{-k}}$ represents all $\boldsymbol{\lambda}$ except λ_k.

Step 5: $\boldsymbol{\varphi_.}|\boldsymbol{\Xi}_{-\varphi_i}, \tilde{\boldsymbol{U}}, \boldsymbol{y}$

$$\boldsymbol{\varphi_.}|\boldsymbol{\Xi}_{-\varphi_i}, \tilde{\boldsymbol{U}}, \boldsymbol{y} \sim N(D_\varphi d_\varphi, D_\varphi) \tag{19}$$

where

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

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

$$\mathbf{M}^{\varphi}_{\mathbf{i.t}} = \begin{bmatrix}\mathbf{I_J} & \mathbf{P_i}\end{bmatrix}; \boldsymbol{\theta^\varphi} = \begin{bmatrix}\boldsymbol{\alpha_.}' & \beta\end{bmatrix}'.$$

Step 6: $\sigma^2_\varphi|\boldsymbol{\Xi}_{-\sigma^2_\varphi}, \tilde{\boldsymbol{U}}_{\boldsymbol{i.t}}$

$$\sigma^2_\varphi|\boldsymbol{\Xi}_{-\sigma^2_\varphi}, \tilde{\boldsymbol{U}}_{\boldsymbol{i.t}} \sim IG\left[\frac{N}{2} + \alpha_\varphi, \left(b_\varphi^{-1} + .5\sum_{i=1}^{N}\varphi_i^2\right)^{-1}\right]. \tag{20}$$

Step 7: Draw the $\tilde{\boldsymbol{U}}_{\boldsymbol{i\cdot t}}|\boldsymbol{\Xi}, \boldsymbol{y}$

Given the structure of our model and to ease computation. We draw the latent utilities that individual i derives from visiting site j at the levels instead of differences. That is, we sample the U_{ijt} and then take the differences to get the $\tilde{U}_{ijt}$. This works in our case given our assumption of *iid* distribution on the error term ε_{ijt}. At the structural level of the U_{ijt} in equation (1), there is no correlation among the alternatives conditional on α_j, β, γ, and $\boldsymbol{\varphi_.}$.

Each of the U_{ijt}'s are conditionally normal with mean μ and variance of 1 with truncation point that depends on the choice of the individual. That is, if an alternative is chosen, it must be the alternative that gives the maximum utility - this gives the upper truncation point for all the other alternatives.

We therefore follow 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 and variance from equation (1) and upper truncation point $U_{ikt} = U_{ikt}^{r-1}$.

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

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

Generated Data experiment

In this section we illustrate the performance of the algorithm described above in accounting for model uncertainty in a RUM model with hierarchical structure. We will also use the generated data as a guide to know how many draws will be needed for our application to achieve the same level of precision under independent and identical distribution (*iid*) sampling.

To implement this, we generated data assuming that an individual has to choose between visiting $J = 10$ sites and staying at home over $T = 52$ choice occasions. We also set number of individuals $N = 3,000$. The variable in $\boldsymbol{D_i}$ consists of a uniformly generated random variable that signify the age of the individual and a vector of gender dummy variable generated from a Bernoulli distribution with equal probability of success and failure.

The alternative specific constant for site j (α_j) is drawn from a normal distribution with mean $\boldsymbol{Q_j \alpha_0}$ and variance $\sigma_{\alpha}^{2} = 0.25$ where $\boldsymbol{Q_j}$ includes an intercept term and a uniformly generated random variable $\boldsymbol{Q_{j,1}}$ which can be thought of as water clarity. However, in our estimation, we experiment by naively estimating a model with an added predictor to $\boldsymbol{Q_j}$. In our first experiment, the added predictor is generated from a standard normal distribution that is independent of $\boldsymbol{Q_{j,1}}$. In a second experiment, the added predictor is generated such that it is equal to $\boldsymbol{Q_{j,1}}$ plus a randomly generated uniformly distributed variable. That is $\boldsymbol{Q_{j,2}} = \boldsymbol{Q_{j,1}} + U(0,1)$. This is to test the performance of our model when high correlation exists the two observed site characteristics which can be the case with some water quality measures. Also, to initiate correlation between the unobserved site characteristics and the price coefficient, the travel costs for each individual to each site is set to be the sum of a normal random variable and 0.1 times the unobserved site characteristics. We also set the value of $c = 10$ and $\tau = 0.1$.[6] The remaining parameters of the model were also fixed (values reported in Table 1). These parameters and the variables in equation (1) are then used to generate the latent utility values U_{ijt} for $j = 0, .., J$ which are mapped into the observed choice of the individuals.

The Gibbs sampler described in section 3.3 is implemented for 50000 iterations with 5,000 as burn-in. The results are presented in table 1. The table report parameter posterior means and posterior probabilities of being positive [denoted $P(\cdot > 0|y)$]. We first look at the result of the model selection parameter $\boldsymbol{\lambda}$ since it is the focus of our study. The model specification that appear with the highest frequency signals that they have the largest probability in the distribution of $\boldsymbol{\lambda}$. The result for the first experiment shows that the site attribute $\boldsymbol{Q_{\cdot,1}}$ used to generate the true model appeared with the highest frequency as expected. The intercept term appeared in almost all iterations (49,976 out of the 50,000 times) while $\boldsymbol{Q_{\cdot,1}}$ appeared 36,038 times. The added predictor however appeared only 9400 times signaling that the variable is not a promising part of the model. Figure 1 and 2 presents the graphs of the distribution of the two parameters. The distribution for $\alpha_0(2)$, as expected, looks like a mixture of two normal distribution with majority of the mass in the distribution that includes the variable in the model. However, the distribution for $\alpha_0(3)$ is largely massed around zero as expected.[7]

For the second experiment where the added variable can actually act as a proxy for $\boldsymbol{Q_{\cdot,1}}$ given the high level of correlation (0.92) between the two variables, we find that the frequency is distributed relatively evenly between the two variables. Specifically, either one of the two variables is almost always visited as expected with 47% and 65% relative frequency respectively. The distributions are presented in figure 3 and 4.

The benefit of controlling for model uncertainty can be seen if we naively estimate the model assuming that the site attributes are the two variables described above. The result from this model shows that even the posterior mean for the additional variable is massed away from zero

[6] Other values of c were also used to test the robustness of our results and the conclusions were not qualitatively different

[7] Since the variable was included in the model 9,400 times, the distribution is not fully centered on zero.

(Figure 5) which imply that we would have made posterior inference assuming that the model is an important determinant of demand. $P(. > 0|y) = 0.09$ signaling that the distribution is negative more than 90% of the time. Policies directed to improve such an attribute would be a waste of resources.

The results for the other parameters also converged to their true values relatively quickly. We found that the posterior mean of the parameters are close to the true values with all of them within two standard deviation of the true values. From the posterior distributions, we see that even with the small number of alternatives used, our algorithm performed well in recovering the parameters of the model. We present the Posterior means for the parameters of the generated data experiment in Table 1.

We also calculated the inefficiency factors for the posterior distribution of the parameters as the ratio of the numerical standard error with correlated draws (NSE) over the standard error of the parameter. That is:

$$\sqrt{\text{inefficiency factor}} = \sqrt{1 + 2\sum_{j=1}^{m-1}(1 - \frac{j}{m})\rho_j} \tag{21}$$

where m is the number of draws after convergence and ρ_j is the autocorrelation coefficient which is a correlation between draws as a function of j time separation between them. The estimates of these inefficiency factors are quite high especially for the variance of the individual random effect. The high inefficiency factors shows that we will need to run our simulation longer to get the same level of accuracy as will be in m *iid* draws.

Application

The methods described above is applied to data from the Iowa lakes Valuation project at Iowa State University. This is the same data described in Egan *et al.* (2009) and in the earlier chapters of this dissertation. The wide array of observed water quality attributes and site attributes available to researchers in this data makes it appealing for our study. The water quality attributes were measured by Iowa State University's Limnology Laboratory and include attributes such as Secchi transparency (a measure of the depth of water clarity), Nitrogen and

Phosphorus.[8] The use of this study is also appealing in that there is considerable amount of variation regarding the site characteristics. 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.

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.

For the purpose of this application, the observed site characteristics ($\boldsymbol{Q}$) include the levels and appropriate natural *log* form of both site and water quality attributes. In contrast to Egan *et al.* (2009), we estimate a single model allowing the data to dictate the model with high posterior density that incorporates model uncertainty.

Empirical Results

Using the model and posterior simulator detailed in the previous sections, we fit the site choice model using the Iowa Lakes data. Similar to the first paper, Gibbs sampling is used to generate simulations from the joint posterior distribution and the Gibbs algorithm is first run for 50,000 iterations. The last iteration from this process is then used to initiate two different chains, run simultaneously on two different machines with different seeds. Each of these runs produced 50,000 draws, leaving us a total of 130,000 post-convergence draws to calculate posterior means, standard deviations and to make posterior inference. This is with a burn-in value of 20,000 iterations which was informed by our generated data experiment.

[8] The availability of a large set of water quality attribute of this nature is atypical in recreational demand data but provides for a good pool of attributes used in the literature.

Each iteration of the simulator is also similar to that of the first paper with the additional step of adding an hierarchical structure to the observed site characteristics. As with the first paper, the step of the sampler that takes longer to implement is the simulation of latent utility data for each agent, over 52 choice occasions for each of the 130 alternatives.

Estimation Results

We are primarily interested in applying the algorithm in earlier sections to data from Iowa Lakes Project to address the issue of uncertainty in model specification. Following the results of the first paper where we found no evidence that the water quality attributes included in the model influence site visitation patterns in Iowa, we seek to evaluate how incorporating model uncertainty will change this conclusion.

We addressed this question by considering a general model which contains all available set of water quality measures and non-water quality site attributes. We report parameter posterior means and posterior probabilities of being positive [denoted $P(. > 0|y)$] for key parameters of the model in Tables 2 and 4.

In general, the results are not significantly different from the first paper given that we used similar data and posterior simulator. The alternative specific constants for each site, α_j are all negative with over 99.9% of the posterior mass lying below zero. This is consistent with the data as majority of the households did not visit any given site and the conditional utility from each site reflects this. The marginal utility of income (negative of the coefficient on travel cost i.e., $-\beta$) has a posterior mean of 0.0134. This assumes constant marginal utility of income which is standard in the recreational demand literature. Turning to the individual households characteristics, similar to Egan *et al.* (2009) and the first paper, the size of the household influences the decision to visit a site with older individuals, females, and the less educated more likely to stay at home.

Table 3 provides the results for the observed site characteristics and the frequencies with which the variables are visited in the simulator. The frequency table contains information relevant to variable selection and provides ranking that can be used to select the more promising submodels for further investigation.

The results indicate that variables such as natural log of Total Phosphorus (TP) and acres including wake restrictions are identified as the submodels supported most by the data and the prior chosen. Phytoplankton growth and nutrient conditions in freshwater systems are determined by the level of total phosphorus and is consistent with the fact that high levels of Total Phosphorus reduces the appeal of a site. Other variables that are also ranked relatively high include state park classification, natural log of Total Nitrogen (TN) and the availability of handicap facility. However, there seems to be no substantial difference in the frequency of including the log and levels of total nitrogen. The volatile suspended solids (VSS) variable that decreases water clarity, though with posterior distribution massed away from zero was visited only 9% of the time.

The variables that were included in the model for chapter two were primarily informed by Egan *et al* and was not based on any particular economic theory. We compare the results in this chapter to that of chapter two. In terms of the water quality variables, the only water quality attributes that possessed the expected sign and posterior distribution massed away from zero is the Total Phosphorus variable. This variable also showed up in the model in this chapter about 60% of the time and is also consistent with the first chapter. The other water quality variables in chapter two either came out with the wrong sign or have posteriors equally massed around the positive and negative region. Incorporating model uncertainty in this chapter, though helpful in terms of the sign of some of the parameters, further supports the fact that we do not find a convincing evidence that all the water quality variables included in chapter one are important determinant of recreational site decision. Secchi depth for instance only appeared in the model 10% of the time. The non-water quality variables however are as expected and consistent with the results of chapter two. ln(Acres), Wake, State Park and handicap facilities are among the highest ranked variables in the table.

If the goal is to pick a single "best" model, we can use the rankings to select a subset of the variables by choosing a cutoff for the frequency.[9] For example, choosing variables

[9] George and McCulloch (1997) suggested using a -5 on the log posterior scale to set a factor for selecting the best model

that are visited at least 15% of the time following this method leaves us with a model that includes, $ln(TP)$, $ln(Acres)$, wake restriction, state park classification, $ln(TN)$ and availability of handicap facility. However, aside from the $ln(TP)$, the coefficients on the water quality attributes in most cases have their posterior densities that are more or less evenly divided between the positive and negative values.

We also compare our results to that of Egan *et al.* (2009) that used likelihood ratio values to select the water quality variables that are important in determining site visitation patterns. The differences between our modeling framework and that of Egan *et al.* (2009) has been described in the first paper aside from the variable selection that was introduced in this paper. First, in contrast to Egan *et al.* (2009), we do not make the conclusion that the water variables individually and as a group were consistently significant based on the results in this paper and the first paper. Secondly, Egan *et al.* (2009) concluded that Secchi transparency is clearly the *best* one measure to include, as it is easy to obtain and the most important single measure. However, from our result, not only is the mass of the posterior densities $[P(. > 0|y)]$ hovering around 0.5, it is ranked 9th our of the water quality variables included in our model.

Similar to Egan *et al.* (2009), we find the log of the Total Phosphorus, which determines algae growth to be the variables visited with the highest frequency by the sampler. However, Egan *et al*'s choice of the most important single additional water quality measure is not consistent with our result. Inorganic suspended solids (ISS) and the *log* of Chlorophyll ranks low on our list of water quality variables.

Posterior Calculation and Welfare

Recreational demand models are used primarily to predict how exogenous changes in the attributes of the sites will affect the welfare of the household. These posterior calculations are in particular intuitive and relatively easy to implement in the Bayesian framework. Though the algorithm is similar to that in the first paper, the introduction of model uncertainty in this paper adds another level of complexity to calculating welfare implications of a change in policy scenarios.

The approach we propose is in the spirit of implementing a Bayesian Model Averaging for posterior inference purposes. This approach of averaging over all the variables of the model rather than selecting a subset of the model for welfare analysis takes into consideration the uncertainty related to each of the variables.

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

$$\boldsymbol{\Upsilon_{it}^{s}(\Xi_{-\alpha.}, Q^{s}) = \max_{j}(U_{ijt}^{s}|\Xi_{-\alpha.}, Q^{s}) \quad s = 0, 1} \tag{22}$$

where $\boldsymbol{\alpha}. = (\alpha_1, \ldots, \alpha_J)$ denotes the vector of alternative specific constants. Changes in the site characteristics impact individual consumers by altering the overall appeal of the sites, as reflected in the α_j's. Thus, we no longer have a single set of alternative specific constants, but a set for each scenario (denoted $\boldsymbol{\alpha}_{\cdot}^{s}$). We use the hierarchical structure in equation (6) to simulate the changes to these constants resulting from a change in the site attributes. However, given the structure of the α_0 parameter, we average over the parameter instead of choosing a subset. Thus, the first step of drawing the alternative specific constant will proceed as follows:

Step 1: Draw $\boldsymbol{\alpha}_{(r)}^{s}, \quad s = 0, 1$ using (6).

That is, draw $\boldsymbol{\alpha}_{(r)}^{s}$ from a normal distribution with mean $\boldsymbol{Q^{s}[\alpha_{0(r)} * p(\lambda(r)|Y)]}$ and variance $\sigma_{\alpha(r)}^{2}$. where (r) indexes each iteration of the *posterior* simulator of the stated parameter. What this does is that at each iteration, only variables visited are used to simulate the alternative specific constant which will be used to average the CV estimate. This way the frequency that a variable is included in the model is used to weight the variable and follows the procedure proposed by Chipman, George and McCulloch (2001).

Once we draw the alternative specific constants, the other steps in the algorithm is the same as that of the first paper. The utility levels are drawn using the simulated parameters and used to calculate the simulated based estimate of the compensating variation defined as:

$$\widehat{CV} = \frac{1}{R}\sum_{r=1}^{R}\frac{T}{-\beta}\left[\left(\max_{j} U_{ijt}^{1(r)}\right) - \left(\max_{j} U_{ijt}^{0(r)}\right)\right]. \tag{23}$$

This algorithm is then applied to the Iowa lakes data. The scenario of interest is a case in which the water quality attributes in the 9 zonal lakes are upgraded to the quality of W.

Okoboji lake. This is similar to the first paper and Egan *et al.* (2009). The result is an estimated compensated variation of \$17.47. However, the result should be interpreted with caution. The probability that the density $[P(\widehat{CV} > 0)]$ is greater than zero is approximately 80%. This result is in contrast to Egan *et al.* (2009) that found the CV estimates to range between \$8 and \$40 depending on the model used with their "best" model having a CV estimate of about \$29.

Summary

This study has proposed a method for incorporating model uncertainty in RUM models, with particular attention to welfare measurement. In addition to proposing a method of selecting a "best" model, posterior calculation of welfare implications of changes in policy scenarios is proposed. The resulting model is used to study visitational patterns of Iowans and how site characteristics influence choice of sites. The proposed method is in the spirit of hierarchical model framework employing the stochastic search variable selection (SSVS) method in George and McCulloch (1993). The model can used to identify a preferred model and as part of the process of employing a Bayesian Model Averaging technique. This is particularly useful for welfare analysis and counterfactual calculation. This technique accounts for model uncertainty which is incorporated not only into model selection but also posterior simulation and welfare analysis.

Comparing the results of our approach to Egan *et al.* (2009), our choice of model and subsequent welfare analysis produces a different result. First, we do not find compelling evidence that all water quality attributes are important in the determination of visitation patterns of Iowans. Secondly, welfare estimates accounting for model uncertainty is lower in our model compared to the preferred model in Egan *et al.* (2009).

In summary, this paper proposes a framework for incorporating model uncertainty in recreation demand. In contrast to Egan *et al.* (2009) that proposes a likelihood ratio dominance method with unknown sample properties, we propose a method that selects a model(s) with maximum posterior probabilities conditional on the data. In general, the results from this paper indicate that incorporating model uncertainty is important in understanding visitation patterns of Iowans.

The findings of this study suggests that there is uncertainty surrounding the inclusion of site attributes in a model of recreational demand. Ignoring this uncertainty may lead to wrong prioritization of clean-up activities especially for water quality attributes. The ability to average over all possible models makes this method appealing for identifying lakes to target for improvements and in an efficient manner with probability based ranking of the site characteristics. This will be particularly useful for policy makers and stakeholders concerned with water quality.

Tables and Figures

Table 1: Posterior Results for Generated Data Experiment

Parameter		Model with SSVS					Ignoring Model Uncertainty	
		No correlation			High correlation			
	True	Mean	$P(\cdot > 0 \mid y)$	$\sqrt{\text{Inefficiency Factor}}$	Mean	$P(\cdot > 0 \mid y)$	Mean	$P(\cdot > 0 \mid y)$
$\alpha_0(1)$	-1.91	-1.80	0.00	2.05	-1.58	0.00	-1.60	0.00
$\alpha_0(2)$	-0.53	-0.57	0.10	2.26	-0.15	0.40	-0.90	0.02
α_0(added regressor)	0	-0.07	0.22	1.26	-0.45	0.13	-0.16	0.10
β	-3.71	-3.72	0.00	20.40	-3.73	0.00	-3.73	0.00
γ_{01}	0.40	0.40	1.00	7.93	0.46	1.00	0.44	1.00
γ_{02}	0.50	0.56	1.00	6.52	0.56	1.00	0.57	1.00
σ^2_φ	0.50	0.49	1.00	20.81	0.50	1.00	0.48	1.00
σ^2_α	0.1	0.19	1.00	1.19	0.20	1.00	0.19	1.00
Alternative specific constants								
α_1	-2.49	-2.48	0.00	15.32	-2.48	0.00	-2.46	0.00
α_2	-2.42	-2.40	0.00	12.97	-2.41	0.00	-2.39	0.00
α_3	-2.22	-2.20	0.00	13.63	-2.21	0.00	-2.19	0.00
α_4	-1.94	-1.91	0.00	13.50	-1.91	0.00	-1.89	0.00
α_5	-1.99	-1.99	0.00	15.26	-1.98	0.00	-1.97	0.00
α_6	-1.60	-1.59	0.00	11.49	-1.59	0.00	-1.57	0.00
α_7	-2.26	-2.24	0.00	14.33	-2.25	0.00	-2.23	0.00
α_8	-2.76	-2.73	0.00	17.94	-2.74	0.00	-2.72	0.00
α_9	-2.54	-2.50	0.00	14.81	-2.51	0.00	-2.49	0.00
α_{10}	-1.31	-1.29	0.00	11.53	-1.29	0.00	-1.27	0.00

Table 2: Posterior Means of Travel Cost,

Table 2: Demographic Variables and Variance Parameters

Parameter	Mean	$P(\cdot > 0\|y)$
Travel cost	-0.0134	0.0000
Demographic Variables		
Age	0.1083	1.0000
Gender	0.1342	0.9985
Education	-0.1633	0.0000
Adults	-0.0956	0.0000
Child	-0.0045	0.0381
Variance parameters		
σ_{φ}^2	1.91	1.0000
σ_{α}^2	0.08	1.0000

Table 3: Posterior Means of hierarchical Parameters (Site Characteristics)

Site Characteristics	Posterior Mean	$P(\cdot > 0\|y)$	Proportion $[P(\lambda\|Y)]$
α_0	-4.2366	0	1
ln(Total Phosphorus)	-0.2584	0.0317	0.5878
ln(Acres)	0.1771	1	0.3312
Wake	0.1558	0.9977	0.3019
State Park	0.1116	0.98	0.1974
ln(Total Nitrogen)	-0.0079	0.4734	0.1925
Handicap	0.0944	0.9667	0.1625
ln(N03)	-0.0687	0.125	0.1433
Total Nitrogen	-0.0375	0.2956	0.1281
Ramp	0.0416	0.7326	0.1273
ln(Cyanobacteria)	-0.0705	0.0358	0.1252
N03	0.0342	0.7053	0.1212
ln(Chlorophyll)	0.0173	0.5981	0.1184
Water Quality Interaction[a]	0.0559	0.9408	0.1135
Secchi	0.0022	0.5249	0.1005
Total Phosphorus	0.0009	0.7166	0.0932
Chlorophyll	0.0006	0.641	0.0922
Cyanobacteria	7.50E-06	0.5749	0.0911
ISS	-0.0017	0.2287	0.091
VSS	-0.0089	0.0718	0.0904
Acres	9.00E-07	0.5207	0.0901

[a]Water Quality interaction was introduced as a way of capturing a single water quality indicator. This is a crude way to capture this and is defined as the log of the product of all the water quality attributes

Table 4: Posterior Means of Alternative Specific Constants

Lake	Mean	$P(\cdot > 0\|y)$	Lake	Mean	$P(\cdot > 0\|y)$	Lake	Mean	$P(\cdot > 0\|y)$
Arbor	-3.9275	0.0000	Hooper	-4.3465	0.0000	North Twin	-3.3558	0.0000
Arrowhead	-3.5753	0.0000	Indian	-3.7789	0.0000	Oldham	-4.3417	0.0000
Arrowhead	-4.1592	0.0000	Ingham	-3.4519	0.0000	Otter Creek	-4.0339	0.0000
Ave. of the Saints	-4.2301	0.0000	Kent Park	-3.6818	0.0000	Ottumwa Lagoon (proper)	-3.4404	0.0000
Badger Creek	-3.7581	0.0000	Lacey-Keosauqua	-3.4008	0.0000	Pierce Creek	-3.9949	0.0000
Badger	-3.3999	0.0000	Ahquabi	-3.5265	0.0000	Pleasant Creek	-3.4225	0.0000
Beaver	-4.1718	0.0000	Anita	-3.4787	0.0000	Pollmiller	-3.7134	0.0000
Beed's	-3.4307	0.0000	Cornelia	-3.5217	0.0000	Prairie Rose	-3.5208	0.0000
Big Creek	-3.0041	0.0000	Darling	-3.4689	0.0000	Rathbun	-2.8085	0.0000
Spirit Lake	-2.6911	0.0000	Geode	-3.3150	0.0000	Red Haw	-3.6192	0.0000
Black Hawk	-3.2365	0.0000	Hendricks	-3.7819	0.0000	Red Rock	-2.8233	0.0000
Blue	-3.4784	0.0000	Icaria	-3.2498	0.0000	Robert's Creek	-3.9423	0.0000
Bob White	-4.0734	0.0000	Iowa	-3.8190	0.0000	Rock Creek	-3.6220	0.0000
Brigg's Woods	-3.6687	0.0000	Keomah	-3.7978	0.0000	Rogers	-4.1931	0.0000
Brown's	-3.3885	0.0000	Manawa	-2.9209	0.0000	Saylorville	-2.8201	0.0000
Brushy Creek	-3.2466	0.0000	Macbride	-3.1417	0.0000	Silver	-3.3566	0.0000
Carter	-3.6157	0.0000	Miami	-3.7688	0.0000	Silver	-4.1963	0.0000
Casey	-3.9098	0.0000	Minnewashata	-3.4602	0.0000	Silver	-4.0143	0.0000
Center	-3.6519	0.0000	Lake of The Hills	-3.6529	0.0000	Silver	-3.5440	0.0000
Central	-3.9834	0.0000	Three Fires	-3.3757	0.0000	Slip Bluff	-4.3711	0.0000
Clear	-2.6311	0.0000	Orient	-4.1194	0.0000	South Prairie	-4.1744	0.0000
Cold Springs	-3.7565	0.0000	Pahoja	-3.4359	0.0000	Spring	-3.8895	0.0000
Coralville	-2.9538	0.0000	Smith	-3.7267	0.0000	Springbrook	-3.5933	0.0000
Crawford Creek	-3.9497	0.0000	Sugema	-3.3857	0.0000	Storm Lake	-2.8789	0.0000
Crystal	-3.4714	0.0000	Wapello	-3.3570	0.0000	Swan	-3.3453	0.0000
Dale Maffit	-3.9819	0.0000	Little River	-3.6989	0.0000	Thayer	-4.2660	0.0000
DeSoto Bend	-3.2837	0.0000	Little Sioux Park	-3.6266	0.0000	Three Mile	-3.3365	0.0000
Diamond	-3.8340	0.0000	Little Spirit	-3.1163	0.0000	Trumbull	-3.7913	0.0000
Dog Creek	-3.7958	0.0000	Little Wall	-3.8735	0.0000	Tuttle	-3.8808	0.0000
Don Williams	-3.4618	0.0000	Littlefield	-3.8006	0.0000	Twelve Mile	-3.4214	0.0000
East Osceola	-3.7125	0.0000	Lost Island	-3.2040	0.0000	Union Grove	-3.9034	0.0000
East Okoboji	-2.5189	0.0000	Lower Gar	-3.3525	0.0000	Upper Gar	-3.4705	0.0000
Easter	-3.6519	0.0000	Lower Pine	-3.6610	0.0000	Upper Pine	-3.5401	0.0000
Eldred Sherwood	-4.0616	0.0000	Manteno Pond	-4.2881	0.0000	Viking	-3.3972	0.0000
Five Island	-3.4527	0.0000	Mariposa	-4.0683	0.0000	Volga	-3.4062	0.0000
Fogle	-4.0403	0.0000	Meadow	-4.4744	0.0000	West Okoboji	-2.3348	0.0000
George Wyth	-3.3615	0.0000	Meyers	-4.0864	0.0000	West Osceola	-3.6274	0.0000
Green Belt	-4.4293	0.0000	Mill Creek	-3.7472	0.0000	White Oak	-4.5455	0.0000
Green Castle	-4.2100	0.0000	Mitchell Impoundment	-4.4165	0.0000	Williamson Pond	-4.6817	0.0000
Green Valley	-3.5536	0.0000	Moorehead	-3.8495	0.0000	Willow	-3.9528	0.0000
Greenfield Lake	-4.114	0.0000	Mormon Trail	-4.0983	0.0000	Wilson	-4.1120	0.0000
Hannen	-3.9010	0.0000	Nelson Park	-4.1805	0.0000	Windmill	-3.9269	0.0000
Hawthorn	-3.7806	0.0000	Nine Eagles	-3.7967	0.0000	Yellow Smoke	-3.5167	0.0000
Hickory Grove	-3.8078	0.0000						

Figure 1 Posterior Distribution for the true regressor ($\alpha_0(2)$) (No correlation Experiment)

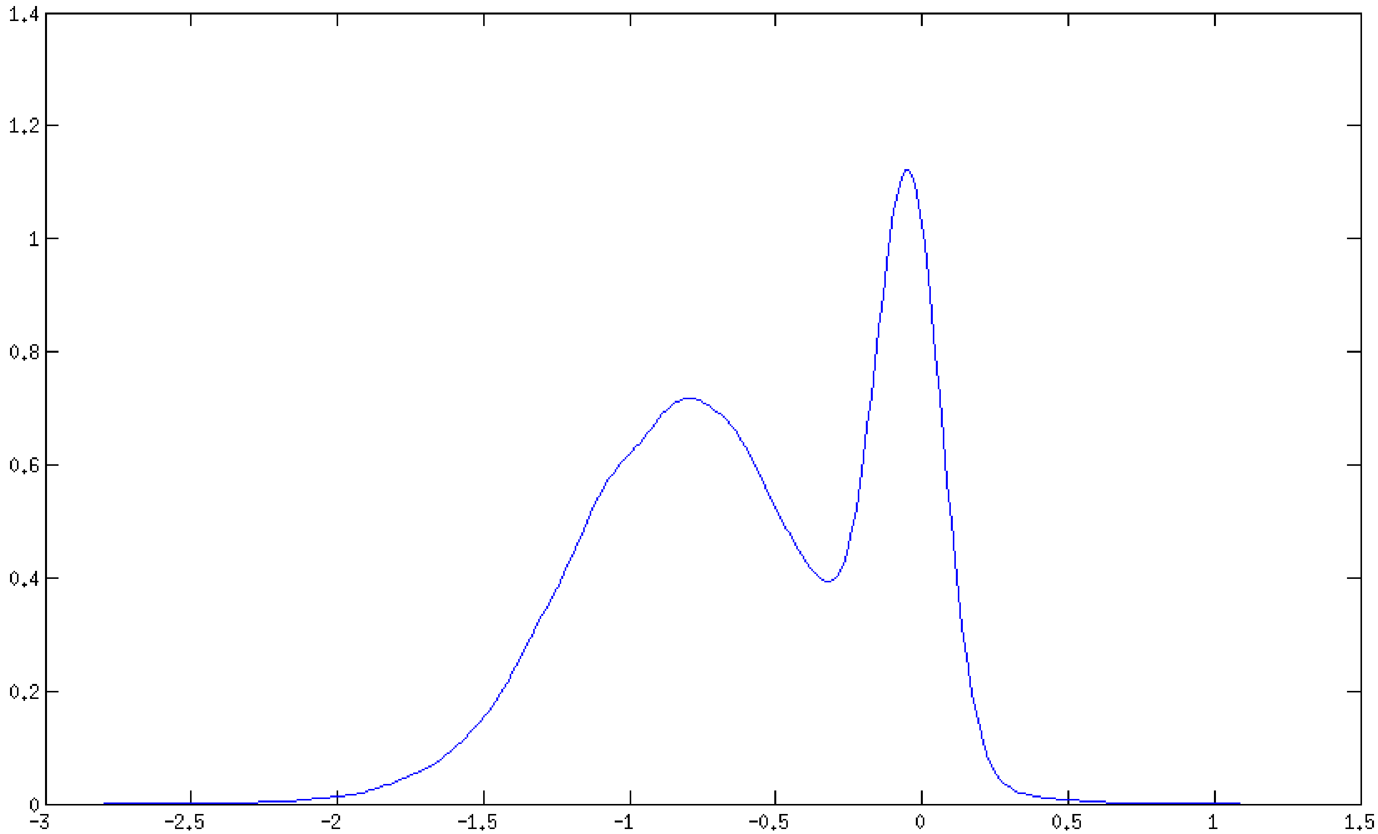


Figure 2 Posterior Distribution for the added regressor ($\alpha_0(addedregressor)$) (No correlation Experiment)

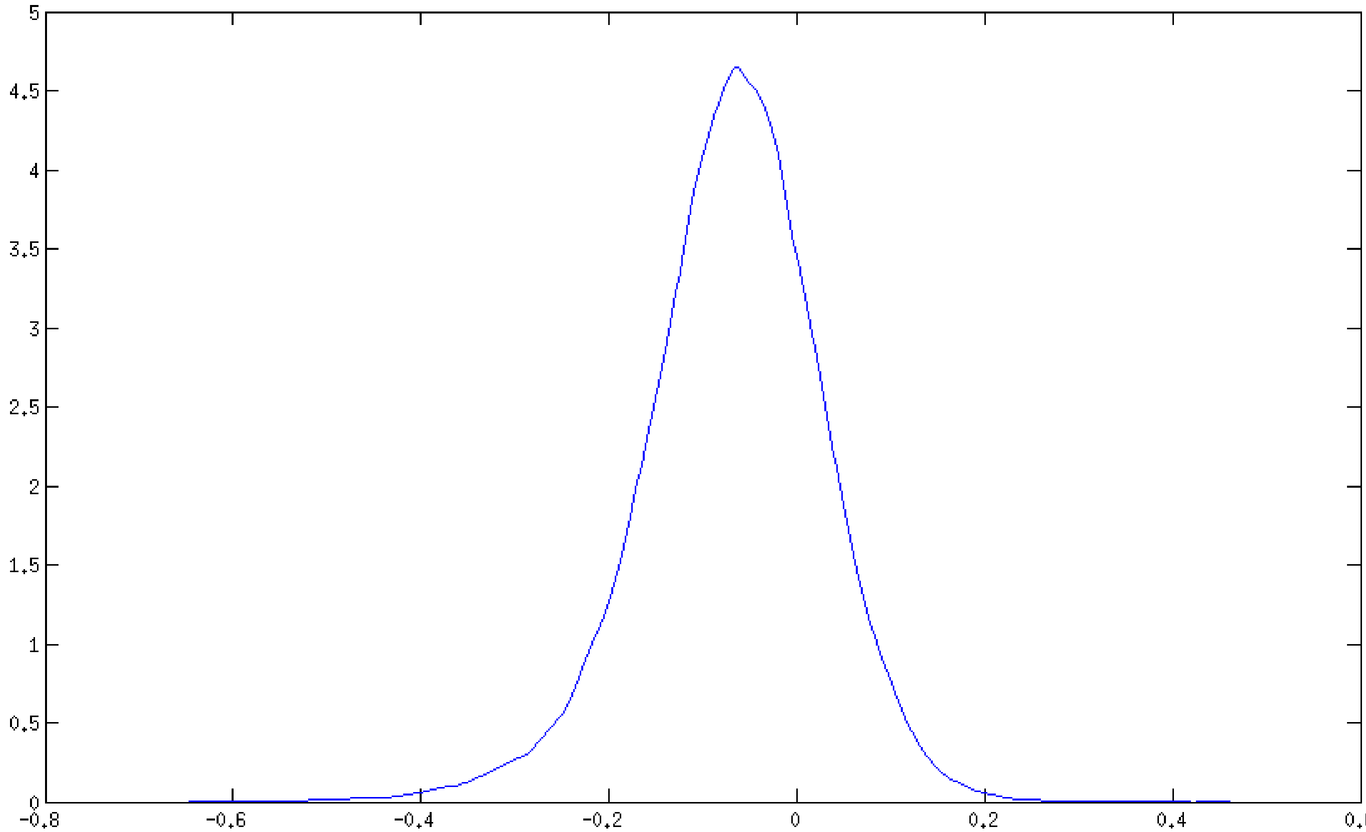


Figure 3 Posterior Distribution for the true regressor ($\alpha_0(2)$) (Correlation Experiment)

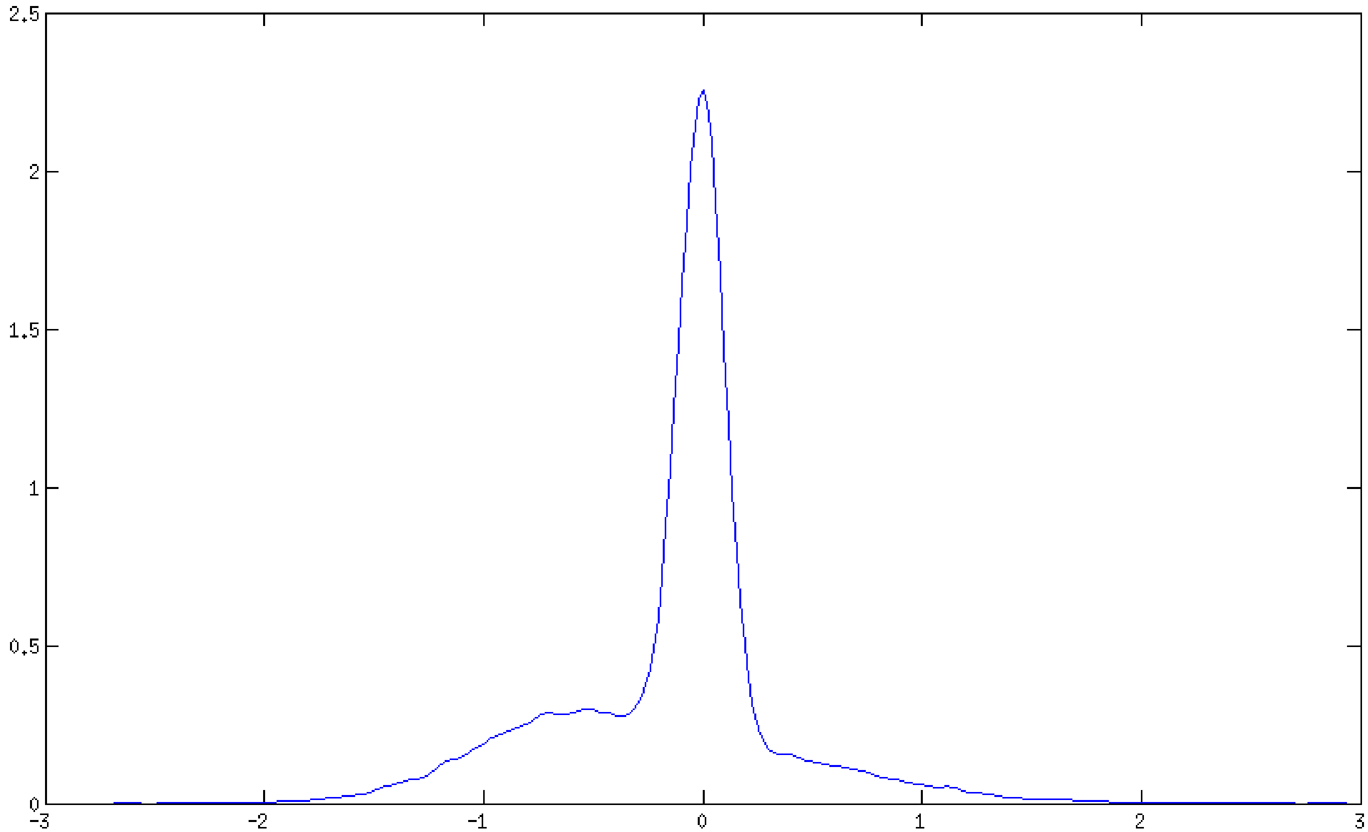


Figure 4 Posterior Distribution for the added regressor ($\alpha_0(addedregressor)$) (Correlation Experiment)

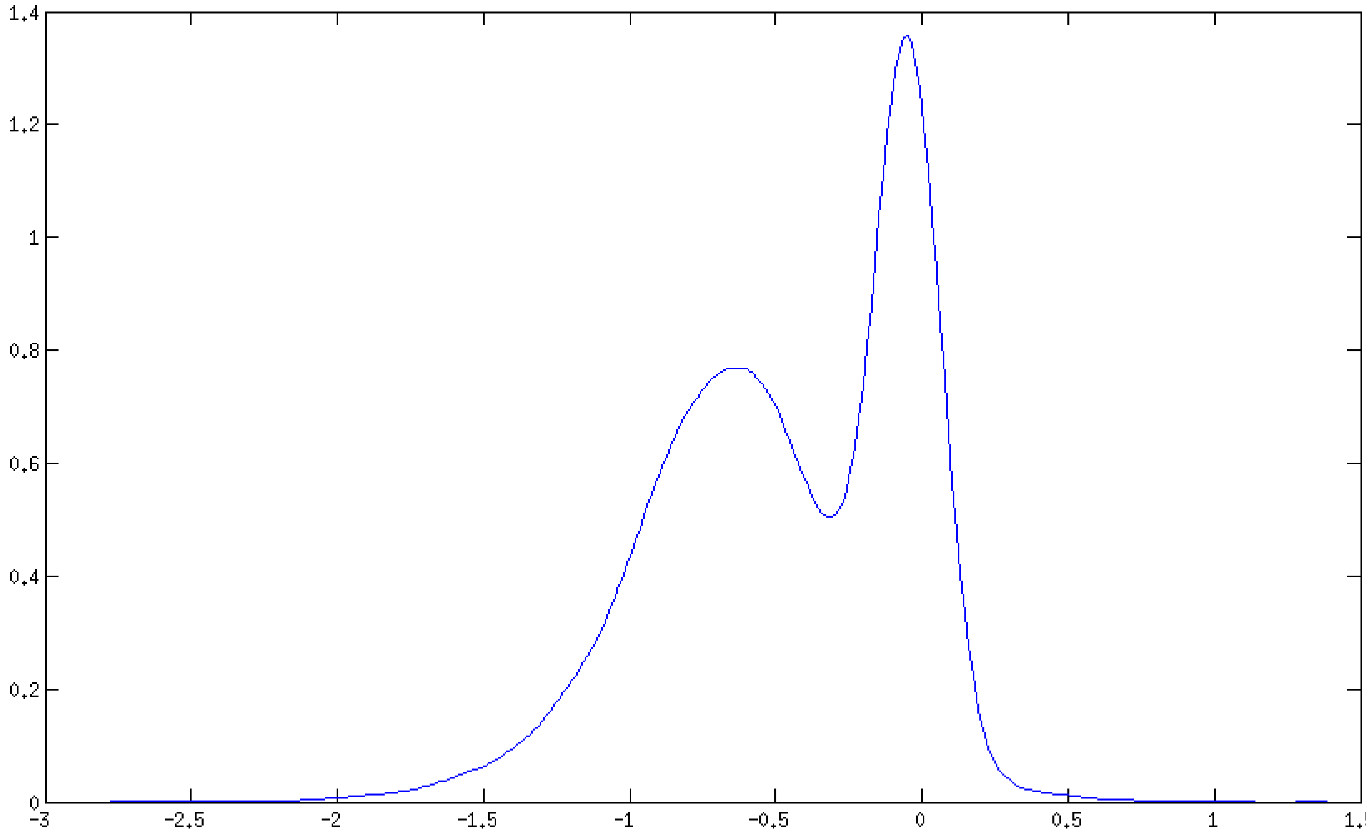
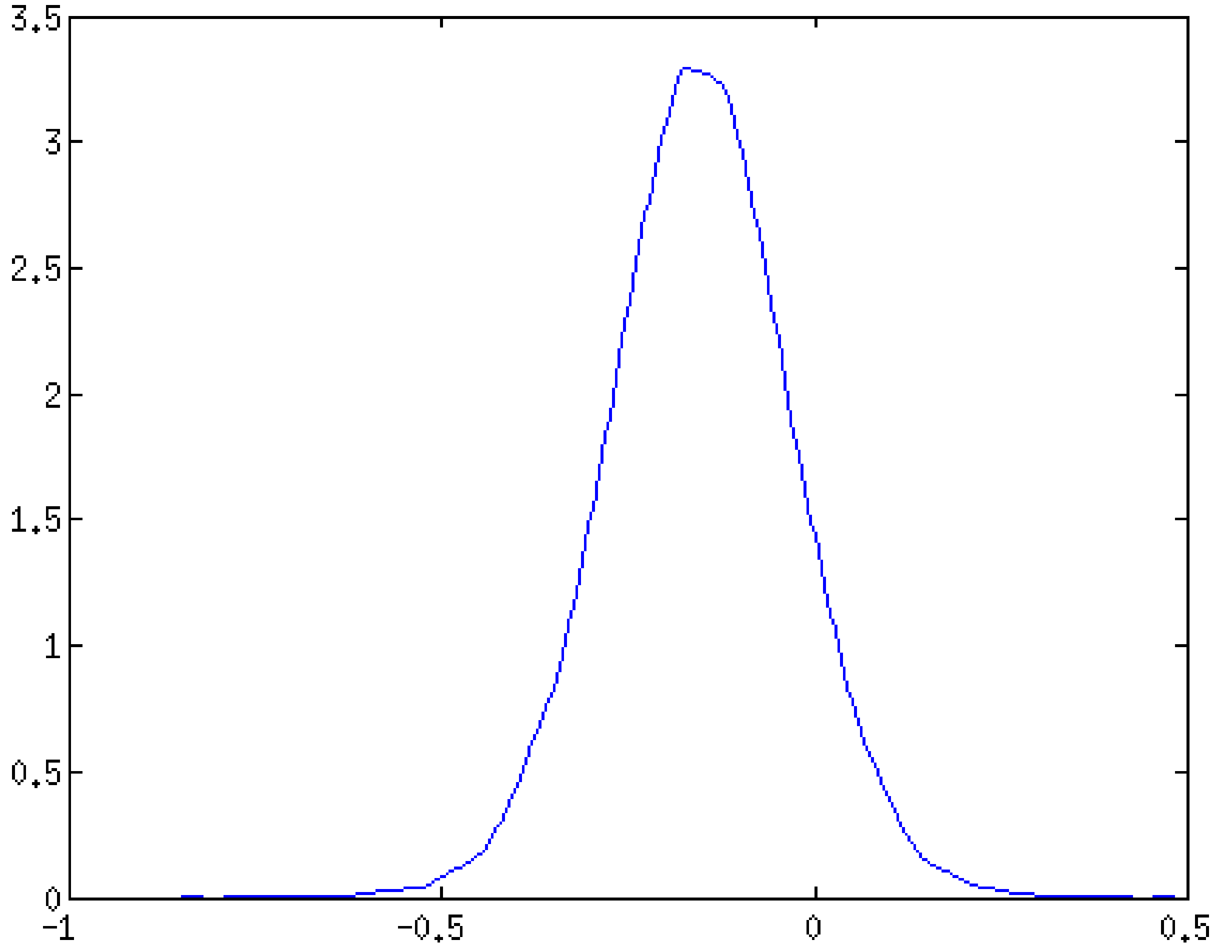


Figure 5 Posterior Distribution for the added regressor ($\alpha_0(addedregressor)$) (Ignoring Model Uncertainty Experiment)



General Conclusions

General Discussion

The three papers in this dissertation contributes to the recreation demand literature by correcting for model misspecification. The first paper focuses on consistently estimating the travel cost parameter by isolating it from the effect of unobserved site characteristics, the second paper allows for more flexibility in the estimation of the marginal utility of income parameter and the third paper accounts for uncertainty in model specification of the observed site characteristics.

The first paper corrects for omitted variable bias misspecification by extending the RUM model to allow for unobserved site characteristics. While analysts studying recreation demand can boast of wide variation in the price data, this variation is frequently offset by a paucity of information characterizing the attributes of the sites. The introduction of alternative specific constants with a distribution that depend upon the observed site attributes enhances the estimation of the parameters of the observed site characteristics and simultaneously capture unobserved site characteristics.

The second paper accounts for nonlinear income effect by estimating a function that has a different effect on the level of utility depending on the income category the individual belongs to. The method makes use of the Taylor series function to approximate the nonlinear function. For the applied researcher, this approach of capturing income effect offers a number of advantages. First of all, it is not as computationally intensive as a full semi-parametric model. Also, the discrete nature of the income data in most cases reduces the number of parameters to be estimated. This approach also partly abstracts away the problem of what time frame to use in computing income in the semi-parametric model (though implicitly it affects the validity of the Taylor series approximation).

The third paper on incorporating model uncertainty in recreation demand is a follow up to the first paper. Researchers are forced to use the few observed attributes available to them, however, economic theory does not dictate which of these variables are to be included or in what form. While all the site attributes can be important determinant of demand, there is huge uncertainty as to the link through which these variables can impact demand. The model proposed draws from the literature on Stochastic Search Variable Selection (SSVS) of George and McCulloch (1993) and Bayesian Model Averaging.

Welfare scenarios were calculated in all the three papers exploring how the closure of West Okoboji lake will affect households in Iowa. In papers one and three, we also explored the possibility of improving the nine zonal lakes in Iowa to the water quality level of West Okoboji which is considered to be one of the cleanest lakes in the state. The results indicate that the closure of West Okoboji lake will result in a loss of welfare but no convincing evidence that improving the zonal lakes will result in welfare gain.

BIBLIOGRAPHY

Albert, James H., Siddhartha Chib (1993). Bayesian Analysis of Binary and Polychotomous Response Data, *Journal of American Statistical Association*, 88, 669-679.

Azevedo, C.D., K.J. Egan, J.A. Herriges, and C.L. Kling (2003). *Iowa Lakes Valuation Project: Summary and Findings from Year One. Final Report to the Iowa Department of Natural Resources*, August.

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

Berry, S.,J. Levinsohn and A. Pakes (1995). Automobile Prices in market Equilibrium, *Econometrica*, 63 (4), 841-890.

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 agreeement CR-811043-01-0, Washington, D.C.

Buckland, S. T., Burnham, K. P., Augustin, N. H. (1997) Model Selection: An Integral Part of Inference. *Biometrics* 53: 603-618

Casella, George and Edward I. George (1992). Explaining the Gibbs Sampler, *The American Statistician*, 46 (3) (August), 167 - 174.

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.

Chib, S and B.P. Carlin (1999). On MCMC sampling in hierarchical longitudinal models, *Statistics and Computing*, (9), 17-26.

Casella, George and Edward I. George (1992). Explaining the Gibbs Sampler, *The American Statistician*, 46 (3) (August), 167 - 174.

Chipman, H., George, E., and McCulloch, R. E. (2001), The Practical Implementation of Bayesian Model Selection, Hayward, CA: IMS, pp. 67116.

Clyde, M. (2000). Model Uncertainty and Health Effect Studies for Particulate Matter, *Environmetrics* 11, 745-763.

Clyde, M., Guttorp, P., Sullivan, E. (2000). *Effects of Ambient Fine and Coarse Particles on Mortality in Phoenix*, Arizona ISDS Discussion Paper 00-05

Dagsvik, J.K., and Karlstrom, A. (2005), Compensating Variation and Hicksian Choice Probabilities in Random Utility Models that are Nonlinear in Income., *Review of Economic Studies*, 72, 57-76.

Draper, D. (1995), Assessment and Propagation of Model Uncertainty (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 57, 45-97.

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.

Fernandez, C., Ley, E., Steel, M. F. (2002) Bayesian Modelling of Catch in a North-West Atlantic Fishery, *Applied Statistics* 51: 257-280

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.

George, E.I., and McCulloch, R.E. (1993), Variable Selection via Gibbs Sampling, *Journal of the American Statistical Society*, 88, 881-889.

George, E.I., and McCulloch, R.E. (1997), Approaches for Bayesian Variable Selection, *Statistica Sinica*, 7, 339373.

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.

Geweke, J. (1996). Bayesian inference for linear models subject to linear inequality constraints, in: *Modeling and Prediction: Honoring Seymour Geisser* (eds. W.O. Johnson, J.C. Lee, and A. Zellner), Springer-Verlag, New York, 248-263.

Hanemann, W.M (1981). *Water Quality and the Demand for Recreation,* Department of Agricultural & Resource Economics, UCB. CUDARE Working Paper 164, March.

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.

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.

Herriges, J.A., and Kling, C.L. (1999), Nonlinear Income Effects In Random Utility Models., *The Review of Economics and Statistics*, 81(1), 62-72.

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.

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.

Koop, Gary (2003). *Bayesian Econometrics*, Inc NetLibrary, NetLibrary, Inc, Published by J. Wiley.

Koop, Gary and Poirier, Dale J., 2004. Bayesian variants of some classical semiparametric regression techniques, *Journal of Econometrics*, Elsevier, vol. 123(2), pages 259-282, December.

Koop, G., D.J. Poirier and J.L. Tobias (2007). *Bayesian Econometric Methods*, Cambridge, Cambridge University Press.

Koop, G and L. Tole, (2004) Measuring the health effects of air pollution: to what extent can we really say that people are dying from bad air? *Journal of Environmental Economics and Management* 47, pp. 30–54.

Lancaster, Tony (2004). *An Introduction to Modern Bayesian Econometrics*, Published by Blackwell Publishing Ltd, pp 294-295.

Layton, D. F and S. T Lee, (2006) Embracing Model Uncertainty: Strategies for response Pooling and Model Averaging, *Environmental and Resource* Economics 34, pp. 51-85.

Leamer, E. E. (1983), Let's Take the Con out of Econometrics. *The American Economic Review* 73(1), 31-43

Leon, R., Leon, C. J. (2003) Single or Double Bounded Contingent Valuation? A Bayesian Test, *Scottish Journal of Political Economy* 50: 174-188

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.

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.

McFadden, D. (1981). Econometric models of probabilistic choice. In Manski, C. and McFadden, D., editors, *Structural Analysis of Discrete Data with Econometric Applications*, pages 198-272. MIT Press, Cambridge, MA

McFadden, Daniel. (1995), *Computing Willingness-to-Pay in Random Utility Models.*, Mimeograph, Department of Economics, University of California, Berkley.

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.

Morey, E.R, V.R. Sharma and Karlstrom, A. (2003), A Simple Method of Incorporating Income Effects Into Logit and Nested-Logit Models: Theory and Application., *American Journal of Agricultural Economics*, 85(1), 248-253.

Murdock, Jennifer (2006). Handling unobserved site characterisitics in random utility models of recreation demand, *Journal of Environmental Economics and Management*, 51, 1-25.

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.

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 Statisitcs*, 82 (1) (February), 83-92.

Phaneuf, D., and K. Smith. 2005. "Recreation Demand Models." in: *Handbook of Environmental Economics.* (ed. K. Maeler and J. Vincent) Amsterdam: Elsevier, pp. 671–751.

Poirier, Dale J. (1995), *Intermediate Statistics and Econometrics, A Comparative Approach*, The MIT Press

Raftery, Adrian E. (1995). Bayesian Model Selection in Social Research [with discussion]. Pp. 111-195 in: *Sociological Methodology*, edited by Peter V. Marsden. Cambridge, MA: Blackwell.

Regal, R. and Hook, E. B. (1991). The effects of model selection on confidence intervals for the size of a closed population. *Statistics in Medicine* 10 717-721.

Rossi, P.E., G. M. Allenby, and R. McCulloch (2005). *Bayesian Statistics and Marketing*, John Wiley & Sons Ltd.

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.

Train, K. (2009). *Discrete Choice Methods with Simulation*, Cambridge University Press, 2nd edition, 2009.

Yang, S., Y. Chen, and G. Allenby (2003). Bayesian Analysis of Simultaneous Supply and Demand, *Quantitative Marketing and Economics*, (1), 251-275.

Yu, J. K. and M. M. Fogel. (1978). The Development of a Combined Water Quality Index. *Water Resources Bulletin* 14:1239-1250.