2003

Bayesian inference in recreation demand models: on linking disparate data sources

Yimin Liang
Iowa State University

Bayesian inference in recreation demand models
-- On linking disparate data sources

by

Yimin Liang

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, Major Professor
Barry L. Falk
Catherine L. Kling
Jinhua Zhao
Kenneth J. Koehler

Iowa State University

Ames, Iowa

2003

UMI Number: 3085928

Graduate College
Iowa State University

This is to certify that the doctoral dissertation of

Yimin Liang

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

Contents

Acknowledgements

I would like to express my sincere appreciation to Joseph A. Herriges, for his kind guidance, patience, and encouragement. He has given me tons of help throughout all the stages of my dissertation: from topic choosing, model building, to programs coding and checking, to result interpretation and thesis writing. I would like to thank the rest of my committee members: Barry L. Falk, Catherine L. Kling, Jinhua Zhao and Kenneth J. Koehler for their helpful comments and timely assistance. I would like to especially thank my major professor and Catherine L. Kling for providing me with a research assistantship during the course of this study. I would also like to thank Hal S. Stern who has left ISU and used to serve as a member of committee. He taught me Bayesian analysis and gave me a lot of help and suggestions in Bayesian modeling.

I would like to thank my wife Hui Su and my family for supporting and encouraging me during my graduate study. I would also like to thank all my friends for their friendship and encouragement.

Chapter 1
Introduction

In considering changes to the environment, policymakers need information on the value placed in environmental amenities. For example, in recent years, regulatory agencies have taken steps to preserve and restore wetlands. Yet the optimal level of restoration will depend not only on the costs of such restorations, but also on the value placed on these improvements. Similarly, efforts to avoid the loss of a recreation site, such as a lake or wildlife preserve, can more readily be justified in times of tight budgets if it can be shown that the value to society of the site exceeds the cost of its preservation. Unfortunately, information on environmental values is sparse and comes from a variety of disparate data sources, including survey data and behavioral data (such as visitation rates). This makes it all the more important to integrate what information is available, so as to best inform decision makers. Bayesian analysis provides a natural framework in which to integrate different sources of information. The goal of my dissertation is to develop Bayesian models to address these integrating problems in context of two key problems of cost benefit analysis: (1) Benefits transfer and (2) the combining of stated and revealed preference data in valuing the same environmental amenity.

My dissertation will consist of three essays. The first essay develops a Bayesian framework for benefits transfer analysis. In many situations, often due to time, limited research budgets or other resource constraints, it may not be practical for policymakers to collect primary data on which economic value estimates can be based. In such cases, policymakers may turn to benefits transfer techniques, i.e., making inference from existing studies to other unstudied sites, or "transferring" the benefit estimates from study sites to policy sites. Essentially, this is simply a way of combining sources of information from studied sites and the site of interest, which is naturally modeled in a Bayesian framework. Information from outside the site of interest can be viewed as the basis of a prior on the parameters describing preferences towards the site of interest. Available data from the site of interest can then be used to inform that prior and form a posterior distribution. In this first essay, a hierarchical linear model will be developed and used to address the benefits transfer issues. The hierarchical model combines the various information sources in the sense that it establishes a distributional relationship among the parameters of the models from which the economic value estimates are derived. It goes beyond the classical "model transfer" (Loomis [32]) and lets the models communicate through the distributional relationship of their parameters. An application of this model will be given for the study of Iowa wetlands usage. The data used for the study was gathered as part of a large Iowa Wetlands Survey conducted in 1997 and funded by U.S. Environmental Protection Agency. Surveys were designed to elicit travel cost information, contingent valuation and behavior information, and socioeconomic information from Iowa residents concerning their use of Iowa wetlands.

The second essay details how to combine stated preference (SP) and revealed preference (RP) in a Bayesian framework. The RP model is based on the individuals' actual decisions or choices under the present market and environmental conditions, while SP model is based on survey responses to hypothetical changes in either the availability of an environmental amenity or its characteristics. Historically, analysts have viewed RP and SP as competing sources of data. For market valuation studies, revealed preferences are preferred source of value information since prices and quantities are readily measured and based on the actual behavior of agents in the market (Cameron *et al.* [15]). This viewpoint is adopted by many researchers in the context of nonmarket valuation studies. The welfare estimates from RP models are often viewed as a "benchmark" against the estimates from SP models are judged. However, as many researches have pointed out, data limitations often make environmental valuations based solely on revealed preferences difficult. More recently, a number of researchers have suggested instead that SP and RP data sources should be viewed as complementary sources of information revealing different aspects of the same underlying preferences (Cameron [14], Adamowicz *et al.* [1] etc). Thus, combining two methods should provide better estimates of environmental valuations. Again, Bayesian analysis provides a natural framework within which to integrate different sources of information. In this essay, I will describe procedures to link together these two types of information sources in a Bayesian framework and to incorporate prior beliefs regarding the compatibility of RP and SP data. An application to the valuation of Iowa wetlands is provided.

In the third essay I will develop a mixed logit model to address the issues associated with combining revealed and stated preference data. As noted above, most efforts to date at combining stated and revealed data sources have focused on testing the consistency in the underlying preferences. Indeed, most studies rely upon a simple test of consistency, (i.e., H_0: consistent and H_A: inconsistent), with little attention paid to the form or sources of inconsistency. Recently, Azevedo, Herriges and Kling [6] have attempted to isolate the sources of the inconsistency (e.g., respondents ignoring their budget constraint or measurement errors in RP prices). In this essay, I take the analysis further by trying to explain both the sources of the discrepancy between stated and revealed preferences and to capture heterogeneity in the degree of consistency across individuals. Specifically, I propose to model both SP and RP data from the wetland data used in essay 2 using a mixed logit of demand for wetland visits. In this framework, the discrepancies between SP and RP responses can be modeled as having a distribution in the population; a discrepancy whose means and variances can depend on observed attributes of the survey respondents. Understanding the sources of these discrepancies can help to better design SP surveys.

The first essay and the third essay will have appendix material following each essay. The figures and tables for each essay will appear at the end of the essay.

Chapter 2
Benefit Transfer

2.1 INTRODUCTION

Faced with limited resources for the protection and enhancement of the environment, policymakers must weight the benefits and costs associated potential programs. Yet at the same time, budget or other resource constraints often prevent policymakers from gathering primary data upon which economic value estimates can be based. Instead, regulators may be forced to rely on benefits transfer techniques, i.e., making inference from existing studies to other unstudied sites, or "transferring" the benefit estimates from study sites to policy sites. These "transfers" often take the form of simply constructing a value per unit day from studied sites and employing this same value for all unstudied areas. More recently, there have been a number of attempts to develop formal benefits transfer procedures, transferring either values themselves or the functions from where they are derived (e.g. Boyle and Bergstrom [12], McConnell [36] and Smith [45]). Yet the problem of combining sources of information from a studied sites and the site of interest is naturally modeled in a Bayesian framework. Information from outside the site of interest can be viewed as the basis of a prior on the parameters describing preferences towards the site of interest. Available data from the site of interest can then be used to inform that prior and form a posterior distribution.

Few studies have employed the Bayesian framework in studying the transfer of environmental valuations. Parsons and Kealy [40] used a Bayesian prior updating approach in their study, treating the parameter estimates obtained in one study region as prior information and then updating that prior using sample information from the region of interest. Atkinson, Crocker, and Shogren [4] suggested an empirical Bayes method to exploit the research efficiencies in the study[1]. The purpose of this essay is to further extend this line of research by developing a hierarchical linear model in which the issue of benefit transfer can be addressed and to apply the methods to an analysis of wetland usage in Iowa.

The remainder of this chapter is outline as follows. In section 2, I provide a summary of prior efforts to address the problem of benefits transfer, both those using classical methods and the few Bayesian efforts. Section 3 outlines a Bayesian approach to benefits transfer. I start by reviewing the standard linear model and then consider extensions to a hierarchical model that can be used for benefits transfer and to cases with censored variables. The specifics of the Iowa wetlands data base is then described in section 4. Section 5 provides the models to be estimated, the steps used to assessing the effectiveness of the benefits transfer exercise, and the estimation results. The conclusions and future directions are outlined in section 6.

2.2 LITERATURE REVIEW

As noted above, there have been a number of efforts in the past to transfer environmental valuations, most of them relying on classical statistical methods. In this section, I briefly review the prior literature dividing it between classical and Bayesian approaches.

---

[1] Bayesian methods have been employed by a number of analysts in the electicity demand literature. Caves, Herriges, Train and Windle [16] proposed using Bayesian framework to combine engineering data (treated as prior information source) and direct metered end-use load data. Aigner and Leamer [3] also used an "empirical Bayes" method to study the transferability between utilities, which will be reviewed in detail in next section.

2.2.1 Classical Approaches

While benefits transfer has been practiced by policymakers for years, a special section of Water Resources Research, 28(3), 651-722, 1992 [42] provides a useful summary of the early work. In that section, several researchers discussed the conceptual frameworks for value transfer. Boyle and Bergstrom [12] proposed a systematic, conceptual foundation for conducting benefit transfer studies, and suggested a research agenda to identify conditions under which benefit transfer was valid. Brookshire and Neill [13] addressed the ongoing development of the procedures for benefit transfers through a case study approach. McConnell [36] argued that, "while standard hypothesis testing plays a role in model estimation, it may be less important than the research's judgment about how the model ought to work." Smith [45] illustrated the need for guidelines for deciding when benefit transfer method can be used to value changes in environmental resources.

Benefit transfer methods can be divided into three types: fixed value transfer, expert judgment, and benefit function or model transfer (Brookshire and Neill [13]; Bergstrom and Civita [8]; and Rosenberger and Loomis [42]). Rosenberger and Loomis [42] classified these methods as traditional methods to distinguish them from meta analysis methods.

2.2.1.1 *Traditional Methods*

With the fixed value transfer methods, aggregate or average values per unit day derived from existing study site data are used to estimate the total benefits at a proposed policy site, while with the expert judgment methods the values per unit day derived from an expert judgment or opinion process are used (Bergstrom and Civita [8]). The most common approach to benefit transfer, referred to the "unit day value" method, falls into the second category. It was widely used by the US. Forest Service in the 1970s and 1980s (Garrod and Willis [20]) and is still used by some agencies today. The problem with this approach is that it ignores variability in the quality or applicability of existing data sources or relies on ad hoc selection criteria.

Desvousges *et al.* [19] propose a more systematic approach to value transfer in the context of benefits transfer for water quality improvements. Their study is carried out in two steps: the selection of the study sites and the transfer procedure. They proposed criteria for selecting studies to be used for benefits transfer and established the market areas for the policy sites of interest. They then transferred the study results using information on the relationship between distance and benefit estimates, census data for representative households at the policy sites, and policy site water quality changes.

In a break from the value transfer method, Loomis [32] suggested benefits transfer be based on the transfer of the value function (i.e., the travel cost demand equation and the contingent valuation function) from the existing study site to the new site. This would allow the analyst to correct for known difference between study sites and the site being transferred to. Loomis tested the validity of transferability of benefit estimates based on the travel cost method (TCM) by comparing site-specific benefit estimates with those derived from transferring TCM equations. A zonal TCM demand model was set up for that purpose. Loomis' results led to rejection of the equality of demand coefficients for ocean sport salmon fishing in Oregon versus Washington and for freshwater steelhead fishing in Oregon versus Idaho. The benefit transfer by average benefits per trip were also conducted and compared. He concluded that transfer of either demand equations or average benefits per trip were likely to be in error, but the latter would be less accurate. In addition, he also used the demand equation obtained from $n-1$ Oregon rivers to predict benefits at the nth river, which resulted benefit transfers to rivers within the state being accurate to within $5-15\%$.

Parsons and Kealy [40] analyzed a random utility model of lake recreation in the state of Wisconsin. The transfers were for measuring water quality improvement and performed between two non-overlapping samples divided from the original data, Milwaukee residents and non-Milwaukee residents (state). A variety of transfer methods were examined: simple transfers, model transfers and Bayesian updated transfers. Three assumptions about the level of information on the policy site were considered: no information, limited (no behavioral) information, and some behavioral information. In the first case, the mean of per choice occasion benefits of water quality improvement per person was estimated and used as the estimate for the policy site. The model transfer method was used in the second case, where the differences in choice opportunities (lakes) and the differences in incomes were accounted for the transfers. In the third and final case, a small fraction of sample were randomly drawn from the policy sample and used to estimate the benefits in three ways: by itself, by pooling with study site sample, and by Bayesian updating. They found that the state model, transferred to the Milwaukee sample, estimated the benefits of water quality improvements with considerable accuracy. Moreover, updating the state model with behavioral information from small Milwaukee samples could be used to improve the performance of benefits transfer, but only modestly so.

Kirchhoff *et al.* [30] evaluated the performance of direct benefit transfer and benefit function transfer through contingent valuation method (CVM) applications to two pairs of similar non-market amenities. Their tests led to rejection of convergent validity in nearly all cases for the New Mexico policy sites and study sites and for most cases at the Arizona sites.

Their analysis also led to rejection of the transfer of a simple mean site benefit estimate, for it led to large percentage errors. Benefit function transfer performed better in nearly all cases, which is consistent with the observation of Loomis [32].

There have subsequently been a number of benefits transfer studies based on model transfer. For example, Kask and Shogren [28] developed a protocol for benefit transfer of long-term health risk reduction and presented a case study for surface water contamination. Loomis *et al.* [33] examined the interchangeability of two specifications of travel cost demand models (nonlinear least squares and Heckman sample selection models) for recreation at U.S. Army Corps of Engineer reservoirs in Arkansas, California, and Tennessee/Kentucky. Much of the literature has moved, however, to a meta analysis approach, attempting to capture fundamental characteristics of value functions over a large number of studies.

2.2.1.2 *Meta Analysis Methods*

Meta analysis has a long history in the fields of psychology, education, and the health sciences. Most of its applications involve controlled experiments and focus on consistently aggregating the results from different controlled experiments to avoid the subjective nature of most research reviews. Meta analysis is concerned with understanding the influence of methodological and study-specific factors on research outcomes and providing statistical summaries and syntheses of past research (Rosenberger and Loomis [42]). The argument is that it is possible to use meta analysis to improve judgment for benefit transfer. Rosenberger and Loomis [42] summarized several conceptual advantages of meta analysis over the traditional methods, such as being able to utilize information from a greater number of studies, control the methodological differences when calculating values, and account for the differences of independent variables between original study and the transfer site. Of course, the limitations of using the meta analysis for benefit transfer are also recognized. First, there may not be an adequate number of studies for certain recreation activities. Second, a meta analysis can only be as good as the quality of past research efforts. Third, for meta analysis to be successful, the studies included should be similar enough in content and context (Rosenberger and Loomis [42]). Another potential limitation is that meta analyses frequently employ numerous qualitative or dummy variables capturing differences in functional form, analysis method, etc. Choosing levels for these variables in forming transfer values becomes difficult. For example, the dependent variables used in Rosenberger and Loomis [42] are all qualitative. Garrod and Willis [20] also point out that meta analysis may be subject to publication bias, which reduces its credibility for informing policy. Since significant studies are more likely to be published, meta analysis based on such studies is not likely to provide a total perspective on a particular issue. Thus caution must also be exercised in applying results from meta analysis.

Smith and Kaoru [46] examined 77 travel cost recreation demand models prepared from 1970 to 1986. Using the estimates of consumer surplus per unit of use from previous studies, the authors tried to characterize these values by using variables describing the site characteristics, the activities undertaken at each site, the behavioral assumption and specification decisions made in individual studies. The explanatory variables included in the study can be classified according to: the assumptions inherent in the behavioral model underlying the travel cost framework, the specifications of the estimated demand function, and the econometric procedure used in estimation. The meta analysis was carried out using ordinary least squares (OLS) with the Newey-West version of the White consistent covariance estimator. Their findings showed a systematic relationship between the estimates and the features of the empirical models.[2] Smith and Kaoru concluded that meta analysis can serve as a consistency check to the process used in benefit transfer analysis.

Walsh *et al.* [55] studied 287 estimates of net economic value per day drawn from 120 outdoor recreation demand studies from 1968-1988, containing 156 valuation estimates based on the travel cost model (TCM) and 129 values based on contingent valuation method (CVM). They sought to explain the variations in these estimates via meta analysis. Three models were estimated: one for the complete set of observations, the other two separately analyzing TCM and CVM valuations. The explanatory variables in the studies include independent variables that influence recreation demand (such as cost or price, consumer income, and socioeconomic variables), and methodological variables (such as recreation activity, sample size and coverage, elicitation method (i.e., CVM, TCM or other), and the statistical model). The adjusted R^2 showed that $36-44\%$ of the total variation in the reported values was explained by the included variables, i.e., site, location, and methodological variables. The authors suggested that adjustments be considered in the benefit transfer process, which include adjustments for travel time, sample truncation, use of individual observation approach, substitution, CVM method, site quality, recreation activity and so on.

More recently, Rosenberger and Loomis [42] have further refined and tested meta analysis as a benefit transfer tool. Their meta analysis database consists of 682 estimates from 131 separate studies spanning 1967-1998. The basic analysis tools they used are the same as those used in Smith and Kaoru [46]. Besides carrying out a fully specified model for

---

[2] Again, the problem still remains in deciding what settings to use for some of these variables when constructing "transfer" values.

traditional investigations of factor effects on study outcomes, they obtained the meta analysis benefit transfer models by optimizing the basic meta regression models: retaining only those variables that were significant at an 80% level of confidence or better based on t-statistic. There were five models optimized for benefit transfer applications: a national model and four regional models. They also used the fitted benefit transfer models to forecast use values sensitive to methodology, recreation activity, and site characteristics. The in-sample convergent validity tests were performed based on adapting the meta models to regional and activity-specific conditions. They claimed that "all the models fit the data reasonably well" and the national model minimized percent difference across all activities by region. They suggested, "...the national meta analysis regression model would be a starting point for agencies estimating recreation benefits arising from broad environmental policies that effect recreation use over a large number of sites in a region" ([42], p.1106).

2.2.2 Bayesian Approaches

There are few applications of Bayesian analysis in benefit transfer in the recreation demand literature. Atkinson *et al.* [4] is perhaps the first. In this paper, the authors specified an economic model for the policymaker's extrapolation problem and derived a series of conditions for achieving efficient benefit transfer. They found that random coefficient approaches are fully consistent with those conditions. Assuming exchangeability, they proposed the use of an empirical Bayesian estimator to exploit the research efficiencies.[3] When the empirical Bayes estimation procedure is applied to a specific site, the estimate

[3] The n values v_i $(i = 1, ...n)$, are regarded as exchangeable if the joint probability density $p(v_1, ...v_n)$ is invariant to permutations of the indexes (Gelman *et al.* [22]). Exchangeability is different from independence.

from that site is shrunk towards the mean estimate from the other sites. The degree of this shrinkage depends on the variance of the parameters that defined the benefit structure as well as the sampling variance. They concluded that, for their data set, the research efficiency gains from increased sample size generally dominate the gains from shrinkage, which implied that policymakers would have little incentive to extrapolate the hedonic benefits estimates from previous studies to new sites. Of course, this conclusion is less useful in those cases in which there are resource constraints and obtaining the sample data from the policy site is difficult or impossible.

As noted above, Parsons and Kealy [40] use a Bayesian updating approach as one method to estimate and test the benefit transfer between sites (Milwaukee residents and non-Milwaukee residents). This method was carried out in the case where they assumed they had behavioral information for a fraction of the Milwaukee sample. They assumed alternatively that they had information for 13, 28, and 55 randomly chosen individuals from the Milwaukee sample. These were referred as "small" Milwaukee samples. In their study, three transfer models were considered: (1) a RUM with the small Milwaukee sample only, (2) a RUM with pooled small Milwaukee sample and the state sample, and (3) a Bayesian model. In approach 3, the parameters estimated in the state model (a RUM with non-Milwaukee sample) were treated as prior information and the parameters estimated in Milwaukee model (a RUM with small Milwaukee sample) were treated as sample information. They used the sample information to update the prior information with a conventional Bayesian statistic

$$\hat{\beta} = \left[Var(b_N)^{-1} + Var(b_M)^{-1}\right]^{-1} \left[Var(b_N)^{-1} b_N + Var(b_M)^{-1} b_M\right],$$

where $\hat{\beta}$ is Bayesian estimator, b_N and b_M are parameter estimates from the state model and Milwaukee model respectively, $Var(b_N)$ and $Var(b_M)$ are the corresponding variance-covariance matrices. This updated estimate $\hat{\beta}$ is a weighted average of b_N and b_M, weighted by the inverse of their variance-covariance matrices. As we can see from Table 2.1, the Bayesian estimates provide the best results in those cases where data were available for 13, 55 individuals. Parsons and Kealy suggested that updating the model with behavioral information from policy site could improve the performance of benefits transfer. A limitation of their approach is that their sampling process was not repeated, i.e., only one sample was drawn for each of the "small" Milwaukee samples. Thus their results might not represent the general cases. It is possible that the results would differ if another set of the "small" samples was drawn and used in the analysis.

Some efforts have been undertaken in applying Bayesian methods to investigate transferability in the electricity demand literature. Caves, Herriges, Train and Windle [16] use a Bayesian framework to model appliance specific electricity demand, combining engineering data (treated as prior information source) and direct metered end-use load data. Aigner and Leamer [3] employ an "empirical Bayes" estimation approach fostered by Lindley and Smith [31] to the problem of inferring time-of-use (TOU) pricing response in a subject utility ("policy site") based on information from other TOU pricing experiments. Their goal was to determine, through analyzing the results on experiments done other utilities, whether it was necessary for the subject utility to conduct its own TOU experiment. The intuition behind this is that, if all of the past TOU experiments produced the same set of findings, the subject utility could with confidence predict the effects of TOU pricing on individuals within its service territory and further experimentation would be unnecessary. On the other hand, if the findings vary widely for unexplainable reasons, the subject utility might want to conduct its own experiment.

The model used by Aigner and Leamer [3] was based on a random coefficients linear regression model where the coefficient vector was assumed to come from a multivariate normal distribution. The model structure is

$$y_i = X_i\beta_i + \varepsilon_i \qquad \varepsilon_i \sim N(0, \sigma_i^2 I), \tag{2.1}$$

and

$$\beta_i \sim N\left(\bar{\beta}, \Sigma\right), \tag{2.2}$$

where the subscript i represents the region i $(i = 1, ..., m)$, y_i is the vector of observations and X_i is the corresponding explanatory variable matrix, β_i is the parameter vector, ε_i is the error vector (with variance σ_i^2), $\bar{\beta}$ is the mean vector for the $\beta_i's$, and Σ is the variance-covariance matrix for the $\beta_i's$. This structure makes the issue of transferability transparent; i.e., if Σ is a zero matrix, then the normal distribution (2.2) degenerates to $prob\left(\beta_i = \bar{\beta}\right) = 1$, so that $prob(\beta_i = \beta_j) = 1\ \forall i, j$ and the regression equations in (2.1) are perfectly transferable from region to region. On the other hand, as Σ tends to a infinite matrix, $prob\left(|\beta_i - \beta_j| < \varepsilon\right) \to 0$ for all finite positive ε and there is no transferability between the regions. In other cases, the transferability will be somewhere between these two extremes and the degree of transferability will be determined by the value of Σ.

In the case where $\Sigma = 0$, with $\beta_i = \bar{\beta}\ \ \forall i$,the regression coefficient vector is estimated by pooling data. The posterior mean of β_i is

$$\hat{\beta}_i = \hat{\bar{\beta}} = \left(\sum_i \sigma_i^{-2} X_i' X_i\right)^{-1} \left(\sum_i \sigma_i^{-2} X_i' y_i\right), \tag{2.3}$$

and corresponding variance-covariance matrix is

$$V_{\min} = \left(\sum_i \sigma_i^{-2} X_i' X_i \right)^{-1}, \tag{2.4}$$

which is the smallest covariance matrix. This estimate and variance would apply to all experiments. Aigner and Leamer [3] refer to this situation as "perfect transferability", since all data are pooled as if there were no differences among experiments.

At the other extreme (i.e., as $\Sigma \longrightarrow \infty$) the coefficient vectors across the regressions are totally unrelated. Each coefficient vector has to be estimated from its own regression data source. Without an experiment or study for a policy site, the estimation of a particular coefficient vector is impossible. In this case, the posterior means and variances are

$$\hat{\beta}_i = (X_i' X_i)^{-1} X_i' y_i \,, \tag{2.5}$$

and

$$V_{i\,\max} = (\sigma_i^{-2} X_i' X_i)^{-1}. \tag{2.6}$$

In intermediate cases, the estimates can be expressed as a function of Σ :

$$\hat{\beta}_i = (\sigma_i^{-2} X_i' X_i + \Sigma^{-1})^{-1} (\sigma_i^{-2} X_i' y_i + \Sigma^{-1} \bar{\beta}), \tag{2.7}$$

$$\begin{aligned} \widehat{\bar{\beta}} \;=\; & \left[\sum_i (\sigma_i^{-2} X_i' X_i + \Sigma^{-1})^{-1} \sigma_i^{-2} X_i' X_i \right]^{-1} \cdot \\ & \left[\sum_i (\sigma_i^{-2} X_i' X_i + \Sigma^{-1})^{-1} \sigma_i^{-2} X_i' y_i \right], \end{aligned} \tag{2.8}$$

and the posterior variances are

$$Var(\beta_i | \bar{\beta}) = (\sigma_i^{-2} X_i' X_i + \Sigma^{-1})^{-1}, \tag{2.9}$$

$$Var(\bar{\beta}) = \left[\sum_i X_i'(X_i \Sigma X_i' + \sigma_i^2 I)^{-1} X_i \right]^{-1}, \tag{2.10}$$

$$Var(\beta_i) = Var(\beta_i|\bar{\beta}) + Var(\beta_i|\bar{\beta})\Sigma^{-1}Var(\bar{\beta})\Sigma^{-1}Var(\beta_i|\bar{\beta}). \tag{2.11}$$

Aigner and Leamer [3] use an index function to indicate the degree of overall data transferability:

$$\omega_i = \frac{|Var(\beta_i)^{-1}| - |V_{\max}^{-1}|}{|V_{\min}^{-1}| - |V_{\max}^{-1}|}. \tag{2.12}$$

When this measure is applied to an individual parameter, it becomes

$$\omega_{ik} = \frac{\sigma_{\Sigma}^{-2} - \sigma_{\max}^{-2}}{\sigma_{\min}^{-2} - \sigma_{\max}^{-2}}, \tag{2.13}$$

where σ^2 refers to the variance of the coefficient in question. In their application, the prior for $\bar{\beta}$ was non-informative, and the values of σ_i^2 and Σ used the posterior modal estimates suggested by Lindley and Smith [31]. Though the results of their empirical study for the data transferability were not very good, they established a Bayesian framework for benefit transfer studies.

The idea of the transferability index is very intuitive. However, just like the concept of R-square, while the two extreme cases (perfect transferability and non transferability) are defined clearly, the intermediate cases are not. This index provides a relative measure of transferability, and absolute criteria for judging when transferability is adequate have not been established. Since their sample size was small (only 3 utilities were considered in the model), a full Bayesian analysis for the hierarchical linear model was very difficult to carry out. If a full Bayesian analysis is possible, one can investigate the posterior distributions of the parameters and determine the transferability between experiments.

2.3 BAYESIAN FRAMEWORK

As noted in the previous section, most of the efforts aimed at addressing benefits transfer issues in environmental recreation demand literature rely on classical statistical methods. In this section, after reviewing basic Bayesian inference methods for Linear Regression Models, I develop a Bayesian framework for benefits transfer in a recreational demand setting.

2.3.1 Review of Basic Bayesian Inference Methods for Linear Regression Models

Linear regression is a basic statistical tool widely used in econometrics. Bayesian models for linear regression have been well developed for decades. In this subsection, I will review the basic results for the normal regression models in the Bayesian framework. Throughout, I use notation such as $N(\mu, \sigma^2)$ for random variables and $N(\theta|\mu, \sigma^2)$ for density functions.

2.3.1.1 *Noninformative Prior*

Using vector notation, the simplest linear regression model is

$$y|\beta, \sigma^2 \sim N(X\beta, \sigma^2 I), \tag{2.14}$$

where the data are assumed to be i.i.d. random variables from a normal distribution.

A conventional noninformative or diffuse prior distribution for (β, σ^2) takes the form

$$p(\beta, \sigma^2) \propto \sigma^{-2}. \tag{2.15}$$

The joint posterior distribution is then

$$\begin{aligned} p(\beta,\sigma^2|y) &\propto p(\beta,\sigma^2)p(y|\beta,\sigma^2) \\ &\propto \sigma^{-2}N(y|X\beta,\sigma^2 I), \end{aligned} \tag{2.16}$$

Since

$$p(\beta,\sigma^2|y) = p(\beta|\sigma^2,y)p\left(\sigma^2|y\right), \tag{2.17}$$

we can determine first the posterior distribution for β, conditional on σ^2, and then the marginal posterior distribution of σ^2. In this way, it is easy to simulate draws from the posterior distribution since both densities $p(\beta|\sigma^2,y)$ and $p\left(\sigma^2|y\right)$ are from standard distributions:

$$\beta|\sigma^2,y \sim N(\hat{\beta},\hat{\Sigma}), \tag{2.18}$$

where

$$\hat{\beta} = (X'X)^{-1}X'y, \tag{2.19}$$

$$\hat{\Sigma} = \sigma^2(X'X)^{-1}; \tag{2.20}$$

and

$$\sigma^2|y \sim \text{Inv-}\chi^2(n-K,s^2), \tag{2.21}$$

where

$$s^2 = \frac{1}{n-K}(y-X\hat{\beta})'(y-X\hat{\beta}). \tag{2.22}$$

Here n is the number of observations and K is the number of elements in vector β. Also, the marginal posterior distribution of β can be recognized as a multivariate student t with $n-K$ degrees of freedom.

2.3.1.2 *Including Prior Information*

If we have prior information about β and σ^2, it is not difficult to incorporate it into the analysis. For instance, a conjugate prior distribution makes the analytical computation easier,

$$\begin{aligned} p(\beta, \sigma^2) &= p(\sigma^2)p(\beta|\sigma^2) \\ &= \text{Inv-}\chi^2(\sigma^2|n_0, \sigma_0^2)N(\beta|\beta_0, \sigma^2\Sigma_0). \end{aligned} \tag{2.23}$$

The joint posterior distribution of (β, σ^2) will be

$$p(\beta, \sigma^2|y) \propto \text{Inv-}\chi^2(\sigma^2|n_0, \sigma_0^2)N(\beta|\beta_0, \sigma^2\Sigma_0)N(y|X\beta, \sigma^2 I). \tag{2.24}$$

The posterior density can also be factored the same way as in equations (2.17) - (2.22), but now equations (2.19) and (2.20) are replaced with

$$\hat{\beta} = (X'X + \Sigma_0^{-1})^{-1}((X'X)\tilde{\beta} + \Sigma_0^{-1}\beta_0), \tag{2.25}$$

$$\hat{\Sigma} = \sigma^2(X'X + \Sigma_0^{-1})^{-1}; \tag{2.26}$$

where $\tilde{\beta}$ is the OLS estimate for β, i.e.,

$$\tilde{\beta} = (X'X)^{-1}X'y \tag{2.27}$$

and the posterior distribution of σ^2 is also Inv-χ^2 with degree of freedom $(n_0 + n)$, scale parameter

$$\phi = \frac{(n-K)s^2 + n_0\sigma_0^2 + (\tilde{\beta} - \beta_0)'\Sigma_0^{-1}(X'X + \Sigma_0^{-1})^{-1}X'X(\tilde{\beta} - \beta_0)}{n + n_0}, \tag{2.28}$$

where

$$s^2 = (y - X\tilde{\beta})'(y - X\tilde{\beta})/(n - K). \tag{2.29}$$

2.3.1.3 *The Regression Models with a Known Covariance Matrix*

A more general case for the normal linear regression is when the data covariance matrix is not necessarily proportional to the identity matrix:

$$y|\beta \sim N(X\beta, \Omega). \tag{2.30}$$

But here we only consider the case where the covariance matrix Ω is known. In this case, the posterior distribution of β is still normal. If a noninformative uniform prior distribution for β is used, the posterior distribution of β is given by (2.18) with σ^2 fixed at 1 and (2.19) and (2.20) replaced by

$$\hat{\beta} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y, \tag{2.31}$$

$$\hat{\Sigma} = (X'\Omega^{-1}X)^{-1}. \tag{2.32}$$

If a normal prior

$$\beta \sim N(\beta_0, \Sigma_0)$$

is used instead, then (2.19) and (2.20) are replaced with

$$\hat{\beta} = (X'\Omega^{-1}X + \Sigma_0^{-1})^{-1}(X'\Omega^{-1}y + \Sigma_0^{-1}\beta_0), \tag{2.33}$$

$$\hat{\Sigma} = (X'\Omega^{-1}X + \Sigma_0^{-1})^{-1}. \tag{2.34}$$

2.3.2 Hierarchical Linear Model for Benefit Transfer

From the work of Aigner and Leamer [3], we have seen that the Bayesian framework provides a natural structure for combining disparate regressions and facilitates benefit transfer studies. The hierarchical linear model, introduced by Lindley and Smith [31], has been studied by many statisticians, including, e.g., Gelfand *et al.* [21], Wakefield *et al.* [54], and

Gelman *et al.* [22]. In Aigner and Leamer [3], the variance parameters were treated as known in the stage of Bayesian analysis and finally replaced with the modal estimates. A more formal hierarchical linear model for the benefit transfer can be expressed as follows.

Let $y_j = (y_{1j}, ..., y_{n_j j})$, where y_{ij} is the ith observation for the jth experiment and n_j is the total observations of the jth experiment. Let X_j be the explanatory variable matrix for the jth experiment. Then the model for the jth experiment is

$$y_j|\beta_j, \sigma_j^2 \sim N(X_j\beta_j, \sigma_j^2 I), \quad \text{for } j = 1, ..., J \tag{2.35}$$

where β_j is the coefficient vector, σ_j^2 is the variance. This model is a little different from the population model in Gelfand *et al.* [21] and Wakefield *et al.* [54], where σ_j^2's are assumed to be equal across experiments. We assume that β_j's are from a normal population with mean $\bar{\beta}$ and covariance matrix Σ_β :

$$\beta_j|\bar{\beta}, \Sigma_\beta \sim N(\bar{\beta}, \Sigma_\beta). \tag{2.36}$$

For completeness of the model, we need prior distributions for σ_j^2's and the hyper-parameters $\bar{\beta}$ and Σ_β. The conjugate priors are often used for computational convenience, for example, Wakefield *et al.* [54] use

$$\sigma_j^2 \sim \text{Inv-}\chi^2(n_{j0}, \sigma_{j0}^2), \forall j \tag{2.37}$$

$$\bar{\beta} \sim N(\bar{\beta}_0, \Sigma_{\bar{\beta}}), \tag{2.38}$$

$$\Sigma_\beta^{-1} \sim W((\rho R)^{-1}, \rho) \tag{2.39}$$

as the prior for σ_j^2, $\bar{\beta}$ and Σ_β^{-1}, where n_{j0}, σ_{j0}^2, $\bar{\beta}_0$, $\Sigma_{\bar{\beta}}$, ρ and R are known, and W denotes a Wishart distribution. One can also set the prior of σ_j^2's as drawn from an Inv-χ^2 distribution:

$$\sigma_j^2 \sim \text{Inv-}\chi^2(n_0, \sigma_0^2) \ \forall j, \tag{2.40}$$

where n_0, σ_0^2 are known.

After the model is set up, a Bayesian analysis can be performed using a variety of methods. Due to the complexity of the hierarchical linear models, the integrals required for a fully Bayesian analysis do not have closed form expressions, with numerical or analytical approximations often used instead. In Aigner and Leamer [3], an empirical Bayes estimation procedure is used. In our application to the Iowa wetlands data, the Gibbs sampling algorithm will be employed to simulate the posterior distribution of the parameters.

2.4 DATA SOURCE: IOWA WETLANDS SURVEY DATA SET

This data set was gathered as part of a large Iowa Wetlands Survey conducted in 1997 and funded by U.S. Environmental Protection Agency. The surveys were carried out in two stages: pre-test and final survey mailing. The pre-test survey was used to test the survey on a sample of Iowa residents and to seek improvement on the survey design. The final survey was mailed in February 1998. Of the 6000 surveys sent, 594 were undeliverable, 3139 were returned with a 58% response rate among deliverable surveys. Among those returned surveys, some surveys were incomplete and some were unusable due to the unrealistic answers.[4] These surveys are of a very small amount and not used in our study.

Surveys were designed to elicit travel cost information, contingent valuation and behavior information, and socioeconomic information from Iowa residents concerning their use of Iowa wetlands. The state was divided into 15 zones based roughly on the Iowa crop reporting districts (see Figure 2.1). A copy of the survey is provided in Appendix.

---

4 For example, some survey respondents indicated that their total number of trips taken to the wetlands exceeded 100 per year.

Respondents were asked to indicate the number of trips they had taken to each zone over the past year and how their behavior would have changed if the travel cost increased. The actual number of trips taken by individuals in 1996 provides the basis for the recreation demand model used in my analysis. Travel costs, from each individual's home to the wetland zones, are based on travel distance (priced at 0.21 cents per mile) plus the cost of travel time (priced at one-third the individual's wage rate). The socioeconomic information such as gender, age, and income was also collected and will be used in my analysis. Table 2.2 provides some summary statistics for the wetlands survey data.

2.5 APPLICATION TO THE IOWA WETLANDS DATA

One objective of the Iowa Wetlands study was to develop and test Bayesian procedures for benefits transfer (Herriges and Kling [24]). The design of the study, from sample point selection to survey questionnaire design, gave some consideration to this objective. The sample can be divided into 15 sub-samples according to the Iowa crop reporting districts (see Figure 2.1). These zoned sub-samples can be further grouped into similar types. For examples, zones 4, 5 and 8 consist of the Prairie Pothole region of Iowa. One would expect trip demands for this region to be similar across zones 4, 5, and 8. Likewise, one might expect similar behavior along the border regions (zones 1, 2, 3, 13, 14, and 15) as these areas are dominated by riverine wetlands.

2.5.1 Conceptual Issues

There are three primary difficulties in the application of the hierarchical model (2.35)-(2.39). First, even though we have 15 "experiments", it is still a small sample for use in estimating the unknown hyper-parameters $\bar{\beta}$ and Σ_β for the population distribution of β_j (2.36). Second, the number of parameters involved in the model is large. We need to calculate the posterior distribution for the zone specific beta coefficients and variance parameters as well as the hyper-parameters. Third, our trip data are censored (non-negative), resulting in a Tobit style regression model and further complicating Bayesian analysis.

2.5.1.1 *Identification*

To overcome the identification problem, we must put some restrictions on the model structure. One reasonable restriction is imposing a simpler prior structure (i.e., a diagonal matrix) on to the covariance matrix Σ_β :

$$\Sigma_\beta \equiv diag(\tau_1^2, ..., \tau_K^2), \tag{2.41}$$

where K is the number of elements in β. The priors are diffuse:

$$p(\sigma_j^2) \propto \sigma_j^{-2}, \text{ for } j = 1, ..., J;$$

$$p(\bar{\beta}, \Sigma_\beta) \propto \prod_{k=1}^{K} \tau_i^{-1},$$

equivalently, uniform priors are assigned to $\bar{\beta}$, τ_i and $log(\sigma_j)$ $(\forall i, j)$ to ensure the posterior distribution to be proper. If uniform priors were assigned to $log(\tau_i)$ $\forall i$, the posterior distribution would be improper.

2.5.1.2 *Gibbs Sampler*

Complicated models such as hierarchical models are most conveniently summarized by draws from the posterior distribution of the model parameters. In our application, the Gibbs sampler will be used to obtain draws from the joint posterior distribution for the parameters of interest.

The Gibbs sampler, also called alternating conditional sampling, is a particular Markov chain algorithm often used in multidimensional problems (Gelman *et al.* [22]). It breaks a high-dimensional problem into lower-dimensional pieces using conditional distributions. More specifically, the Gibbs Sampler can be defined in terms of subvectors of a target vector. Suppose we want to draw a sample from the distribution of a parameter vector θ, which can be divided into k subvectors, i.e. $\theta = (\theta_{1,} ..., \theta_k)'$. Then Gibbs sampler processes as follows: starting by drawing an initial θ from an approximation to $p(\theta)$, one repeats the following steps using the most recently drawn values for variables in the conditioning set:

draw θ_1 from $p(\theta_1|\theta_2, ..., \theta_k)$;

draw θ_2 from $p(\theta_2|\theta_1, \theta_3 ..., \theta_k)$;

$\vdots$

draw θ_k from $p(\theta_k|\theta_1, ..., \theta_{k-1})$.

The sequence drawn from this procedure will converge to the distribution of the target distribution $p(\theta)$ as the number of iterations gets large.

To see how the Gibbs sampling procedure applies to problem at hand, first consider the case where data are not censored. The Gibbs sampling algorithm processes as follows. Defining $y = (y_1', ..., y_J')'$, $\beta = (\beta_1', ..., \beta_J')'$, $\sigma^2 = (\sigma_1^2, ..., \sigma_J^2)$, $\tau^2 = (\tau_1^2, ..., \tau_K^2)$, the likelihood function has the form

$$p(y|\beta, \sigma^2) = \prod_{j=1}^{J} N(y_j|X_j\beta_j, \sigma_j^2 I); \tag{2.42}$$

the joint posterior density function is

$$p(\beta, \sigma^2, \bar{\beta}, \Sigma_\beta | y) \propto p(\bar{\beta}, \Sigma_\beta) p(\sigma^2) p(\beta | \bar{\beta}, \Sigma_\beta) p(y | \beta, \sigma^2)$$

$$\propto \prod_{k=1}^{K} \tau_k^{-1} \prod_{j=1}^{J} \{\sigma_j^{-2} N(\beta_j | \bar{\beta}, \Sigma_\beta) N(y_j | X_j \beta_j, \sigma_j^2 I)\}. \quad (2.43)$$

Let

$$D_j \equiv (\sigma_j^{-2} X_j' X_j + \Sigma_\beta^{-1})^{-1}, \quad (2.44)$$

$$\beta_{-j} \equiv (\beta_1', ..., \beta_{j-1}', \beta_{j+1}', ..., \beta_J')', \quad (2.45)$$

$$\sigma_{-j}^2 \equiv (\sigma_1^2, ..., \sigma_{j-1}^2, \sigma_{j+1}^2, ..., \sigma_J^2), \quad (2.46)$$

$$\tau_{-k}^2 \equiv (\tau_1^2, ..., \tau_{k-1}^2, \tau_{k+1}^2, ..., \tau_K^2). \quad (2.47)$$

The full conditional distributions for Gibbs sampler are defined by:

$$\beta_j | y, \beta_{-j}, \sigma^2, \bar{\beta}, \tau^2 \sim N\left(D_j(\sigma_j^{-2} X_j' y_j + \Sigma_\beta^{-1} \bar{\beta}), D_j\right), \quad j = 1, ..., J; \quad (2.48)$$

$$\sigma_j^2 | y, \beta, \sigma_{-j}^2, \bar{\beta}, \tau^2 \sim \text{Inv-}\chi^2 \left(n_j, \frac{1}{n_j}(y_j - X_j\beta_j)'(y_j - X_j\beta_j) \right); \quad (2.49)$$

$$\bar{\beta} | y, \beta, \sigma^2, \tau^2 \sim N\left(J^{-1} \sum_j \beta_j, J^{-1} \Sigma_\beta \right); \quad (2.50)$$

$$\tau_k^2 | y, \beta, \sigma^2, \bar{\beta}, \tau_{-k}^2 \sim \text{Inv-}\chi^2 \left(J - 1, \frac{1}{J-1} \sum_j (\beta_{jk} - \bar{\beta}_k)^2 \right), \quad (2.51)$$

where β_{jk} is the kth element of β_j, and $\bar{\beta}_k$ is the kth element of $\bar{\beta}$. The Gibbs sampling procedure will iteratively generate draws according to conditional distributions (2.48)-(2.51) which are of standard form and easy to draw samples from. As the number of iterations becomes large, the sequence will converge to the target distribution (2.43).

2.5.1.3 *Censoring*

To circumvent the third difficulty, I use the data augmentation and Gibbs sampler algorithm proposed by Chib [17]. The basic insight of Chib's data augmentation procedure is that conditional on the underlying parameters, the unobserved (censored) latent variables can be readily simulated as draws from a truncated normal distribution. In turn, conditional on these simulated versions of the latent variables, the model reduces to the standard linear regression model and standard results apply. A Gibbs sampling procedure can then be used to iterate back and forth between drawing simulated latent variables (i.e., data augmentation) and posterior values for the structural parameters of the model, thus obtaining draws from the joint posterior distribution for the parameters of interest.

Essentially, we introduce latent variables z_{ij} to get around the censoring problem. Define a set of new observations:

$$y_{ij}^* = \begin{cases} y_{ij} & \text{if } \; y_{ij} > 0 \\ z_{ij} & \text{if } \; y_{ij} = 0 \end{cases}, \quad i = (1, ..., n_j),\ j = (1, ...J) \tag{2.52}$$

where z_{ij} are drawn from the truncated normal distribution with mean $(x_{ij})'\beta_j$, and variance σ_j^2, and support $(-\infty, 0]$. Thus

$$y_j^* | \beta_j, \sigma_j^2 \sim N(X_j \beta_j, \sigma_j^2 I), \quad \text{for } j = 1, ..., J, \tag{2.53}$$

where $y_j^* = (y_{1j}^*, ..., y_{n_j j}^*)$. This enable us to use $y^* \equiv (\, y_1^{*\prime}, ..., y_J^{*\prime})'$ in place with y in the above algorithm. To complete the analysis, we need to add one more set of conditional distributions for z_{ij}'s:

$$z_{ij}|y,\beta,\sigma^2,\alpha,\tau^2 \sim N\left((x_{ij})'\beta_j,\sigma_j^2\right) \text{ truncated at the right by 0 if } y_{ij}=0. \tag{2.54}$$

At this point, all the conditional distributions are from known distributions, so it is straightforward to draw the samples.

2.5.2 Empirical Analysis

The primary goal of benefit transfer is to take advantage of past studies and to obtain as precise as possible estimates for the parameters or their functions of interest. In this context, the transferability index function may not be very useful. In this study, I investigate the trade-off between collection of on-site data and the resulting information gain by comparing posterior distributions of the parameters of interest across four stages. The first stage consists of analyzing zone specific data using both classical and Bayesian methods. In stages two, three, and four, I then utilize the hierarchical linear model developed in previous section to investigate the transferability issues with different information structures, i.e., with no on-site data (stage 2), partial on-site data (stage 3), and with all data available (stage 4).

2.5.2.1 *Stage 1: Zone Specific Estimates*

Using data from each zone separately, a standard linear regression model was estimated using the classical Tobit procedure given the specification

$$y_{ij}=\begin{cases} q_{ij} & \text{if } q_{ij}>0 \\ 0 & \text{if } q_{ij}\le 0 \end{cases}, \tag{2.55}$$

and

$$q_{ij}=\beta_{0j}+\beta_{1j}Price_{ij}+\beta_{2j}Income_{ij}+\beta_{3j}Age_{ij}+\varepsilon_{ij}, \tag{2.56}$$

where

$Price_{ij}$ denotes the travel cost for individual i in zone j;

$Income_{ij}$ denotes the income of individual i in zone j;

Age_{ij} denotes the age of individual i in zone j.

The same specification then was estimated using the Bayesian methods outlined in section 2.5.1. These later estimates will be treated as a benchmark to compare with in subsequent evaluations of benefits transfer procedures below.

The zone specific estimation results are summarized in Table 2.3 and Table 2.4. Table 2.3 presents the maximum likelihood estimates for the parameters in zone specific Tobit model. The standard errors for the parameter estimates are shown in the parenthesis below the estimates. As we can see, most estimates for constants and age coefficients are significant at a 5% level, with the number of trips decreasing with age. While both the estimates for income and cost coefficients generally have the expected signs (positive and negative respectively), most are not statistically significant. Table 2.4 presents the posterior means and posterior standard deviations for the parameters in zone specific Tobit model. At a glance, these two tables look very similar to each other. But as a matter of fact, they are computed from different models with different fundamental concepts. Even with the same numerical values, they have different meanings. Table 2.3 provides estimates and standard errors for those parameters that are assumed to be fixed in the Tobit model. Table 2.4 provides the posterior means and posterior standard deviations for those parameters that are assumed to be random and have some prior distributions. Since diffuse priors are used in the Bayesian analysis, it is not surprising that the posterior means and standard deviations of the model parameters turn out to be similar to their counterparts in the classical framework. In both of these two tables, the means and standard deviations for the column statistics are presented in the bottom two lines. These statistics show that the variation of the parameters across zones is substantial and there may be not much similarity among the parameters across zones.

2.5.2.2 *Stage 2: Transfer with No On-site Data*

The second stage of the analysis mimics the situation in which demand data (or model) are available from sites other than the policy site, but no data are available at the site of interest itself. Essentially, each of the 15 sites was treated in turn as the policy site, excluding it from use in construction of the hierarchical model described in section 2.5.1. For example, let zone 10 be the policy site. We use the data from the other 14 zones to run a hierarchical model with 14 experiments, i.e., $J = 14$ in equation (2.35). Using data augmentation and the Gibbs sampling simulation method, we then construct a numerical approximation to the posterior distributions of the hyper-parameters. Since we do not have any information from zone 10, we have to rely on the model assumption (2.35) – (2.36) to predict the distribution of the model coefficient vector for zone (β_{10}). Specifically, a set of samples will be drawn from the posterior distribution of the hyper-parameters and from each of these samples; a new β_{10} will be randomly drawn according to the distribution (2.36). The later drawn samples will be random selections from the conditional distribution of β_{10} given y. Table 2.5 shows the posterior means for the hyper-parameters resulting from this analysis. Essentially, these values are the same as the means of the predicted distributions. Comparing these statistics with the zone-specified results in Table 2.4, we can see big differences between the corresponding statistics for most zones. Nevertheless, these predicted or "transferred" means still lie in the corresponding posterior distributions that resulted from stage 1 within two standard deviations from the posterior means. An interesting fact is that these transferred statistics are very close to those estimates from Tobit regressions pooling all the off-site data for estimation (Table 2.6).

2.5.2.3 *Stage 3: Partial On-site Data*

While policymakers may not have sufficient resources to use to conduct a full scale analysis for the policy site, a small sample of data points may be feasible. In this third stage, I consider augmenting the off-site data with on-site observations. Specifically, I consider the impact of including a small fraction of the actual data from the policy site.[5]

The bootstrap method was used here as a simulation tool. The simulation proceeded as follows:

1) 500 sets of partial data samples were randomly drawn from the data set of the assumed policy site;

2) For each set of these partial data, a hierarchical model analysis was carried out using this set of data together with the data from other 14 sites, and a numerical approximation for the posterior distributions of the parameters of the policy site were generated;

3) Based on these 500 sample distributions, a kernel density estimation procedure was used to summarize the overall impact of adding the partial data into the analysis.

In our analysis, 25 and 50 sample points were considered. As an experiment, we considered zone 10 and zone 11 as policy sites respectively. Zone 10 has a sample size with

---

5 Another possibility that could also be investigated, but was not pursued in this analysis, is to incorporate prior aggregate data for the policy site (which may be more readily available to policymakers) into the posterior distribution for the site. For example, the total number of trips taken to the site may be known.

289 observations while zone 11 is the zone with the largest sample size with 770 observations. The basic results from zone specific Tobit analysis (Table 2.3) show that the estimated coefficients for these two zones have the expected signs. And these estimates provide us a contrast for the cost coefficients, our main concern. The estimates of cost coefficients for zone 10 and 11 are -0.037 and -0.155 respectively, lying on either side of the overall average of the zone specific cost coefficients, -0.139. Table 2.7 presents the summary statistics resulted from our study. The kernel density plots for the parameters are also shown in Figures 2.2 – 2.9.

2.5.2.4 *Stage 4: Full Hierarchical Model*

The full model hierarchical Bayesian analysis was carried out at stage 4. The posterior means and posterior standard deviations of the zone-specified parameters are shown in Table 2.8, and the summary statistics for the hyper-parameters in Table 2.9. The posterior density plots for the hyper-parameters are also shown in Figure 2.10. As we can see from these density plots, the $\bar{\beta}$'s are roughly normally distributed and the τ^2's roughly χ^2 distributed. Note that there is some non-negative support for the density of the posterior distribution of cost coefficient, the parameter of most concern. Although we do not anticipate the number of trips to wetland to increase when the cost of the trips becomes larger, we do observe insignificant cost coefficients in the Tobit regression results (Table 2.3). Thus numerically it is not abnormal to see $\bar{\beta}_{\text{cost}}$ be non-negative in its posterior distribution. A confusing result from this stage is that the posterior means of zone specific parameters are close to each other even if their counterparts differ a lot in the classical Tobit regression estimation. This may not be a problem if we look at the numbers of observations across zones, which vary a lot and range from 46 to 770. The Bayesian weighting is a complicated procedure and it is not so easy to interpret the results for an unbalanced data. However, this is not the case if we have a balanced data. The appendix to this chapter (2.A) shows the simulation results (Tobit estimation results and Full hierarchical Bayesian analysis results) of the same analysis with 400 observations bootstrapped from the data of each zone. By comparing, we can see now the zonal characteristics of parameters are observed.

Having finished analysis for all stages, one can see how the results compare across stages. The first stage provides us a basic view of zone specific estimations. One might treat the estimates from this stage as true estimates for the zone specific models as Parsons and Kealy [40] did in their study. However in a Bayesian context, it might not be a good idea to do so: first, the zone specific data are not complete sample; second, there is information from other similar zones which will be ignored if the zone specific data are used alone. Here I use the estimates from the first stage as a basis of comparison. The second stage goes to the other extreme: transfer the estimates from other zones without any on-site information. In the third and fourth stages, on-site information is gradually added into the off-site data for estimation, with the on-site sample size changing from 25 to 50 to the full sample. The results from these two stages are essentially the compromise of the on-site data and off-site data. The comparisons of the summary statistics for the results from these stages are shown in Table 2.10 (Zone 10) and Table 2.11 (Zone 11). These comparisons are also illustrated graphically in Figures 2.2 – 2.9, where the kernel densities of the posterior distributions for the coefficients are plotted. Generally speaking, with more on-site information adding into the model two things happen. First, the mode of the posterior distribution gets closer to the mode based on the on-site data alone. Second, the variance of the posterior distribution shrinks. It is somewhat disappointing that the results do not turn out exactly in this way. If one looks at the three extreme cases ("on-site only", "full model", and "off-site only"), one can observe a clear moving pattern for the posterior mode: "off-site only" $\rightarrow$ "full model" $\rightarrow$ "on-site only", which is in the expected direction and well demonstrated in the graphs (Figures 2.2 – 2.9). However, the results from the partial transfer (with either 25 or 50 sample points) turn out to be too noisy to have a clear moving pattern for the posterior modes. This may be due to the fact that the sample sizes of the off-site data (2438 for the zone 10 case and 1961 for zone 11 case) are much bigger than that of partial on-site data (25 or 50). The partial data are dominated by the off-site data and do not provide much additional information about the specific site. Another reason is that these partial transfer results are the average across 500 samples of 25 or 50. There may exist large variation in selected samples. Nevertheless, we do observe that the variance of the posterior distribution shrinks when more information is added into the model. This can be seen numerically through the comparison tables for all the coefficients for both zones (10 and 11) with a little deviation in the results of the constants.

2.5.3 Welfare Analysis and Transferability Index

One of the primary reasons for conducting this type of study is to be able to measure the welfare gain or consumer surplus associated with various sites. Here, the measure of consumer surplus is a target of our Bayesian analysis. As a result of Gibbs sampling, the posterior distribution of the price coefficient β_{1j} will be generated at each stage. Using the following form:

$$CS_{ij} = \frac{q_{ij}^2}{-2\beta_{1j}}, \tag{2.57}$$

we can calculate the consumer surplus (*CS*) for each simulated value of the price coefficient and get an approximated posterior distribution of *CS*. Since q_{ij} is different for each individual, the resulting CS will also be different. In order to study this, I normalized the number of trips to unity; i.e., I only considered the CS for an individual who took only one trip to the wetland during the year of survey. Another problem is that CS function involves a reciprocal relationship with the cost coefficient and we would not expect a non-negative value for the cost coefficient. But in our analysis, we did find some support in the non-negative region for the posterior distributions of the cost coefficients (Table 2.10 and 2.11, Figure 2.4 and 2.7). However, in reality it is irrational for an individual to visit wetlands more often as the costs of such visits increase and the other conditions do not change. So it is reasonable to believe that the cost coefficients are always negative. This is an aspect of our prior beliefs ignored to this point in the analysis. However, the posterior reweighting technique (Geweke [23]) provides a convenient way to incorporate this information. That is, regarding the posterior data distribution of the cost coefficients from previous analysis as an importance sampling density, and reweighting this density with a weighting function

$$w(\beta) = \frac{p(\beta|M_{new})}{p(\beta|M_{old})}$$

where $p(\beta|M_{old})$ stands for the prior density used in previous analysis and $p(\beta|M_{new})$ for the later defined prior density. To implement our belief on cost coefficients $(prob(\beta_{1j} < 0) = 1)$, we can simply use the following weighting scheme.[6]

$$w(\beta) = \begin{cases} 1 & \text{if } \beta_{1j} < 0\,\forall j; \\ 0 & \text{otherwise.} \end{cases}$$

Table 2.12 provides the summary statistics of the resulting posterior distributions for this standardized CS over stages, for zone 10 and 11. Figure 2.11 and 2.12 provide a graphical view of these distributions. From these tables and figures, we can easily see that the posterior

[6] In practice, the condition $\beta_{1j} < -0.001$ was used instead of $\beta_{1j} < 0$.

distributions of the CS resulting from the partial data transfer as well as the full model are the compromises of those resulting from the on-site data model and the off-site data model. Apparently, the results are very different for zone 10 and zone 11 if one looks at the figures. In the case of zone 10 (Figure 2.11), we do not see much movement going on for the posterior distributions of CS. As to zone 11 (Figure 2.12), one may think that the densities are abnormal at the first thought. But as a matter of fact, these plots are as reasonable as those of zone 10 if one realizes that the CS has an inverse relationship with the cost coefficient. Due to this inverse relationship, the closer to zero the cost coefficient lies, the bigger change reflected in CS with the same amount of change in the cost coefficient. For example, a change from -0.1 to 0 for the cost coefficient results in a change of 5 to infinity in CS; while with the same amount of change in cost coefficient, but from -1 to -0.9 results in a much smaller change in CS (from 0.5 to 0.5555). Thus one may not observe shrink in variance or increase in precision for the estimation of CS as more information is added into the model.

Although the Aigner-Leamer transferability index (2.12) is not of our main concern in this study, it was also calculated. Table 2.13 shows the results. The indices for the overall coefficient vector are very small and vary across zones with values ranging from 0.0008 to 0.0526. The indices for individual coefficients are much larger compared with the overall indices. These value disagreements between overall indices and individual parameter indices may be due to the covariance relationship between individual coefficients. In terms of interpretation of these indices, it is not clear to what degree the transferability is for values between 0 and 1, though we know there is no transferability when the index equals to 0 and perfect transferability when the index equals to 1. Another problem is that this index function in fact only cares about the relative variances or precision between parameter estimates rather than the degree of transferability of parameters across experiments. Table 2.13 shows that zone 11 has the highest overall index value (0.0526) while zone 2 the lowest one (0.0008). But if one looks at the numbers of observations for these two zones (Table 2.2), one sees that zone 2 has only 46 observations while zone 11 has 770 and it is not surprising that we have the above results.

2.6 DISCUSSION AND CONCLUSION

This essay extended the Aigner and Leamer's benefit transfer framework from by conducting a full Bayesian rather than an "Empirical Bayes" analysis. The posterior distributions of the parameters of interest were numerically characterized through the Gibbs sampling algorithm and a data augmentation method. In our application to the Iowa wetland survey data, we investigated the situations where we have only on-site data, off-site data, partial on-site data and off-site data, and both on-site and off-site data, respectively. Apparently we do gain tightening knowledge about the coefficients when more information is added into the Bayesian analysis. However, due to the inverse relationship of consumer surplus and cost coefficient, the consumer surplus does not necessarily become less variable with the additional information. We also observe that the modes of the posterior distributions for the model parameters generally move closer to the modes based on the on-site data alone. The movement of the posterior mode depends on how much information is added into the model. When a large amount of on-site data are added (full model) the movement is substantial. When small portion of the on-site data (25 or 50) is added into the model, the movement is vague. This may be due to the following reasons:

(1) There exist interactions among the individual parameters within regressions;

(2) Outliners exist, which could have more influence in the stage of partial on-site data transfer where only a few observations were drawn.

(3) The variation in selecting a single sample may be large.

In this setting, a small sample of on-site data does not help, substantially to inform the model of recreation demand in the region.

Table 2.1. Comparison Results for Transfer models with Behavioral information Transferred Benefit Estimates (1978 Dollars) (Parsons and Kealy [43])

	Per Choice Occasion Benefits		Deviation from the true
	Mean	95% C.I	Model Estimate,%
True Milwaukee Model	0.67	0.54-0.80	
Transfer models			
Data available for 13 individuals			
Milwaukee only	0.78	0.51-1.05	16
Pooled transfer	0.65	0.56-0.71	-3
Bayesian transfer	0.66	0.58-0.73	-1
Data available for 28 individuals			
Milwaukee only	1.17	0.74-1.59	75
Pooled transfer	0.66	0.60-0.72	-1
Bayesian transfer	0.65	0.58-0.72	-3
Data available for 55 individuals			
Milwaukee only	0.23	0.08-0.38	-66
Pooled transfer	0.57	0.51-0.63	-15
Bayesian transfer	0.61	0.54-0.68	-10

The deviation from the true model (Milwaukee sample model) estimates was calculated as $[cv^c/cv^t-1]$, where cv^c is the per choice occasion benefit from the current model and cv^t from true model.

Table 2.2 Mean Statistics for Zone Specific Data
(Standard Deviation is shown in parenthesis)

Zone	#Obs	#Trips	Cost ($)	Time (hours)	Income (k$)	Wage ($/hr)	License[1]	Age (years)	Gender[2]
1	101	6.37 (10.81)	14.82 (3.66)	0.88 (0.22)	46.78 (35.30)	22.15 (28.98)	0.74 (0.44)	50.00 (16.61)	0.83 (0.38)
2	46	6.57 (12.33)	16.32 (3.16)	0.82 (0.18)	36.85 (21.75)	17.15 (12.11)	0.70 (0.47)	52.38 (16.06)	0.59 (0.50)
3	74	5.25 (8.60)	11.64 (7.53)	0.64 (0.41)	46.08 (30.65)	19.50 (20.12)	0.69 (0.47)	48.17 (15.13)	0.70 (0.46)
4	49	9.31 (12.53)	17.24 (17.54)	1.01 (0.94)	39.13 (25.03)	16.95 (12.09)	0.84 (0.37)	50.26 (14.12)	0.71 (0.46)
5	166	7.50 (10.30)	17.25 (5.58)	1.06 (0.29)	42.95 (26.13)	18.23 (11.77)	0.81 (0.39)	47.05 (15.44)	0.80 (0.40)
6	85	4.07 (6.49)	20.47 (10.13)	1.13 (0.51)	35.41 (19.45)	13.19 (6.50)	0.76 (0.43)	47.89 (15.52)	0.75 (0.43)
7	97	7.69 (11.25)	26.04 (6.76)	1.49 (0.39)	39.48 (29.22)	16.97 (15.38)	0.82 (0.38)	45.26 (14.86)	0.76 (0.43)
8	66	6.78 (10.62)	13.52 (3.05)	0.73 (0.17)	37.73 (22.61)	15.79 (8.31)	0.61 (0.49)	49.85 (16.19)	0.67 (0.48)
9	416	6.07 (9.30)	17.64 (7.15)	0.93 (0.36)	51.73 (33.78)	22.75 (18.90)	0.65 (0.48)	46.36 (15.42)	0.73 (0.44)
10	293	7.28 (10.05)	20.42 (8.99)	1.18 (0.46)	40.20 (27.42)	18.89 (17.52)	0.81 (0.39)	47.18 (15.89)	0.79 (0.41)
11	770	6.22 (9.91)	19.16 (4.61)	1.07 (0.23)	42.73 (27.82)	17.67 (11.92)	0.66 (0.47)	49.94 (16.11)	0.76 (0.42)
12	200	8.01 (10.91)	20.71 (6.46)	1.21 (0.36)	37.22 (24.46)	15.55 (9.79)	0.80 (0.40)	47.35 (15.48)	0.76 (0.43)
13	120	6.03 (9.03)	21.25 (10.22)	1.28 (0.58)	40.75 (28.66)	17.40 (17.36)	0.69 (0.46)	49.68 (14.80)	0.78 (0.42)
14	127	6.99 (10.02)	12.51 (3.35)	0.65 (0.17)	47.85 (30.07)	20.25 (13.21)	0.62 (0.49)	46.84 (16.08)	0.75 (0.44)
15	121	8.26 (10.89)	15.66 (5.16)	0.91 (0.28)	48.35 (33.76)	22.39 (16.23)	0.75 (0.43)	48.69 (15.30)	0.79 (0.41)
Total	2731	6.68 (10.05)	18.39 (7.36)	1.04 (0.40)	43.45 (29.15)	18.75 (15.47)	0.71 (0.45)	48.29 (15.75)	0.76 (0.43)

[1] License=1 if individual owns a hunting or fishing license, =0 otherwise;
[2] Gender=1 if respondent is male, =0 if female.

Table 2.3. Tobit Estimation for Zone Specific Data
(Standard Error is shown in parenthesis)

Zone	Constant	Income	Cost	Age	sigma	#Obs
1	27.34*	-0.008	-0.008	-0.530*	14.64*	101
	(6.34)	(0.060)	(0.124)	(0.111)	(1.45)	
2	39.59*	-0.017	-0.693	-0.411*	17.38*	46
	(14.88)	(0.203)	(0.472)	(0.197)	(2.62)	
3	13.83*	0.005	0.105	-0.305*	11.27*	74
	(5.72)	(0.049)	(0.088)	(0.099)	(1.30)	
4	8.64	0.220*	-0.065	-0.192	15.45*	49
	(9.60)	(0.094)	(0.106)	(0.172)	(1.97)	
5	24.16*	0.032	-0.140	-0.370*	13.15*	166
	(5.00)	(0.049)	(0.117)	(0.077)	(0.94)	
6	14.87*	0.059	-0.430*	-0.068	8.54*	85
	(5.00)	(0.055)	(0.118)	(0.067)	(0.89)	
7	19.63*	0.100	-0.181	-0.261*	14.53*	97
	(6.79)	(0.064)	(0.118)	(0.109)	(1.34)	
8	29.69	0.012	-0.560	-0.318*	14.25*	66
	(11.04)	(0.108)	(0.412)	(0.128)	(1.67)	
9	6.75*	0.035	0.013	-0.139*	12.83*	416
	(2.70)	(0.022)	(0.046)	(0.045)	(0.59)	
10	10.36*	0.040	-0.037	-0.153*	14.03*	293
	(3.63)	(0.035)	(0.059)	(0.057)	(0.78)	
11	16.60	0.078*	-0.155*	-0.268*	13.92*	770
	(2.57)	(0.024)	(0.058)	(0.036)	(0.49)	
12	15.29*	0.056	0.019	-0.279*	13.87*	200
	(4.40)	(0.047)	(0.091)	(0.071)	(0.88)	
13	14.35*	0.115*	-0.141*	-0.240*	12.25*	120
	(4.89)	(0.050)	(0.064)	(0.085)	(1.09)	
14	0.52	0.009	0.248	-0.050	13.43*	127
	(5.61)	(0.049)	(0.217)	(0.080)	(1.10)	
15	24.58*	0.020	-0.063	-0.380*	13.17*	121
	(5.11)	(0.046)	(0.104)	(0.088)	(1.07)	
mean	17.75	0.050	-0.139	-0.264	13.51	
std	10.00	0.060	0.249	0.131	1.97	

* Denotes significance at $\alpha = 5\%$ level.

Table 2.4. Bayesian Analysis for Zone Specific Data -- Posterior Mean and Posterior Standard Deviation (in parenthesis)

Zone	Constant	Income	Cost	Age	sigma
1	27.54	-0.005	-0.012	-0.538	14.90
	(6.53)	(0.063)	(0.127)	(0.114)	(1.49)
2	41.03	-0.017	-0.719	-0.436	18.03
	(15.71)	(0.213)	(0.490)	(0.213)	(2.83)
3	14.11	0.002	0.104	-0.311	11.49
	(5.88)	(0.051)	(0.090)	(0.103)	(1.40)
4	8.96	0.225	-0.076	-0.199	15.82
	(10.04)	(0.098)	(0.115)	(0.178)	(2.05)
5	24.47	0.033	-0.145	-0.377	13.24
	(5.12)	(0.050)	(0.120)	(0.079)	(0.94)
6	15.17	0.060	-0.446	-0.069	8.77
	(5.16)	(0.056)	(0.122)	(0.068)	(0.94)
7	20.10	0.103	-0.192	-0.266	14.74
	(6.97)	(0.066)	(0.123)	(0.112)	(1.39)
8	30.29	0.007	-0.570	-0.327	14.63
	(11.58)	(0.116)	(0.426)	(0.134)	(1.74)
9	6.74	0.035	0.014	-0.139	12.86
	(2.66)	(0.022)	(0.046)	(0.045)	(0.61)
10	10.32	0.041	-0.037	-0.153	14.08
	(3.64)	(0.034)	(0.061)	(0.057)	(0.78)
11	16.58	0.077	-0.155	-0.268	13.93
	(2.62)	(0.024)	(0.059)	(0.037)	(0.49)
12	15.31	0.056	0.020	-0.281	13.93
	(4.48)	(0.048)	(0.091)	(0.071)	(0.89)
13	14.45	0.115	-0.144	-0.241	12.33
	(4.94)	(0.050)	(0.065)	(0.086)	(1.13)
14	0.42	0.008	0.250	-0.050	13.59
	(5.68)	(0.050)	(0.222)	(0.081)	(1.16)
15	24.52	0.021	-0.064	-0.380	13.29
	(5.26)	(0.046)	(0.105)	(0.090)	(1.01)
mean	18.00	0.051	-0.144	-0.269	13.71
std	10.29	0.062	0.256	0.134	2.05

Table 2.5. Benefit Transfer with No On-site Data - Posterior Means for the Hyper-parameters

Zone	$\bar{\beta}_{\text{constant}}$	τ_{constant}	$\bar{\beta}_{\text{income}}$	τ_{income}	$\bar{\beta}_{\text{cost}}$	τ_{cost}	$\bar{\beta}_{\text{age}}$	τ_{age}
1	13.68	0.998	0.048	0.015	-0.069	0.038	-0.222	0.017
2	13.84	0.936	0.046	0.015	-0.063	0.031	-0.228	0.015
3	14.08	0.895	0.052	0.014	-0.078	0.034	-0.228	0.016
4	13.80	0.957	0.044	0.013	-0.063	0.036	-0.231	0.018
5	12.90	0.979	0.047	0.016	-0.062	0.038	-0.214	0.018
6	14.17	0.958	0.043	0.011	-0.045	0.020	-0.243	0.016
7	13.74	1.054	0.046	0.018	-0.064	0.043	-0.230	0.022
8	13.85	0.893	0.047	0.014	-0.061	0.037	-0.233	0.020
9	15.07	0.920	0.050	0.014	-0.081	0.037	-0.246	0.019
10	14.64	1.201	0.048	0.016	-0.073	0.040	-0.245	0.023
11	13.51	1.083	0.039	0.015	-0.053	0.034	-0.221	0.021
12	13.95	1.045	0.045	0.015	-0.071	0.037	-0.231	0.021
13	13.52	1.041	0.043	0.015	-0.057	0.039	-0.226	0.020
14	14.79	1.108	0.048	0.014	-0.073	0.040	-0.245	0.020
15	13.43	0.859	0.046	0.015	-0.061	0.031	-0.226	0.017

Table 2.6. Tobit Estimation with No On-site Data (Standard Error is shown in parenthesis)

Zone	Constant	Income	Cost	Age	Sigma	#Obs.
1	13.54 (1.20)	0.044 (0.011)	-0.054 (0.021)	-0.228 (0.019)	13.74 (0.25)	2630
2	13.83 (1.18)	0.044 (0.011)	-0.050 (0.021)	-0.237 (0.019)	13.71 (0.25)	2685
3	14.20 (1.20)	0.046 (0.011)	-0.065 (0.022)	-0.236 (0.019)	13.83 (0.26)	2657
4	14.10 (1.18)	0.040 (0.011)	-0.050 (0.021)	-0.241 (0.019)	13.73 (0.25)	2682
5	13.40 (1.21)	0.044 (0.011)	-0.050 (0.021)	-0.230 (0.019)	13.82 (0.26)	2565
6	14.21 (1.20)	0.040 (0.011)	-0.045 (0.021)	-0.245 (0.019)	13.89 (0.26)	2646
7	13.92 (1.20)	0.042 (0.011)	-0.053 (0.022)	-0.238 (0.019)	13.76 (0.26)	2634
8	13.90 (1.19)	0.043 (0.011)	-0.052 (0.021)	-0.238 (0.019)	13.78 (0.25)	2665
9	15.48 (1.30)	0.047 (0.012)	-0.071 (0.024)	-0.258 (0.021)	13.92 (0.28)	2315
10	14.49 (1.24)	0.043 (0.011)	-0.057 (0.023)	-0.249 (0.020)	13.75 (0.26)	2438
11	13.44 (1.34)	0.035 (0.012)	-0.036 (0.022)	-0.225 (0.022)	13.72 (0.29)	1961
12	13.84 (1.22)	0.044 (0.011)	-0.061 (0.022)	-0.235 (0.020)	13.77 (0.26)	2531
13	13.87 (1.22)	0.039 (0.011)	-0.045 (0.022)	-0.238 (0.019)	13.85 (0.26)	2611
14	14.63 (1.21)	0.043 (0.011)	-0.057 (0.021)	-0.249 (0.019)	13.79 (0.26)	2604
15	13.50 (1.21)	0.042 (0.011)	-0.049 (0.021)	-0.232 (0.019)	13.80 (0.26)	2610
All Data	14.02 (1.18)	0.043 (0.011)	-0.053 (0.021)	-0.239 (0.019)	13.79 (0.06)	2731

Table 2.7. Summary Statistics for Partial On-site Data Transfer

(a) Zone 10 with 25 On-site Sample Points

Parameter	mean	std	Min	2.50%	median	97.50%	max
Constant	14.469	1.917	4.704	11.194	14.378	18.574	27.264
Income	0.051	0.025	-0.126	0.001	0.050	0.104	0.273
Cost	-0.072	0.050	-0.412	-0.166	-0.073	0.033	0.296
Age	-0.234	0.033	-0.419	-0.294	-0.232	-0.163	0.075

(b) Zone 10 with 50 On-site Sample Points

Parameter	mean	std	Min	2.50%	median	97.50%	max
Constant	14.441	1.755	5.611	11.261	14.341	18.277	24.805
Income	0.049	0.022	-0.121	0.007	0.048	0.098	0.214
Cost	-0.061	0.052	-0.398	-0.169	-0.060	0.045	0.243
Age	-0.240	0.029	-0.387	-0.290	-0.243	-0.176	-0.012

(c) Zone 11 with 25 On-site Sample Points

Parameter	mean	std	Min	2.50%	median	97.50%	max
Constant	12.916	2.039	1.724	9.238	12.940	17.045	30.777
Income	0.043	0.023	-0.134	-0.003	0.044	0.088	0.211
Cost	-0.056	0.050	-0.543	-0.165	-0.057	0.042	0.343
Age	-0.220	0.035	-0.518	-0.296	-0.218	-0.156	-0.033

(d) Zone 11 with 50 On-site Sample Points

Parameter	mean	std	Min	2.50%	median	97.50%	Max
Constant	12.876	1.681	4.211	9.731	12.756	16.411	22.478
Income	0.042	0.020	-0.081	0.001	0.043	0.082	0.170
Cost	-0.051	0.041	-0.382	-0.143	-0.048	0.017	0.151
Age	-0.220	0.029	-0.479	-0.284	-0.217	-0.168	-0.048

Table 2.8. Full Hierarchical Model Results
Posterior Means and Standard Deviations for Zone Specified Parameters

Zone	Constant	Income	Cost	Age	Sigma
1	13.89 (1.53)	0.043 (0.018)	-0.071 (0.039)	-0.248 (0.029)	14.80 (1.45)
2	13.99 (1.61)	0.043 (0.021)	-0.080 (0.048)	-0.239 (0.027)	18.83 (2.80)
3	13.61 (1.50)	0.040 (0.019)	-0.054 (0.041)	-0.246 (0.027)	11.85 (1.39)
4	14.53 (1.59)	0.054 (0.023)	-0.056 (0.041)	-0.229 (0.028)	17.11 (2.25)
5	14.59 (1.51)	0.047 (0.017)	-0.061 (0.037)	-0.238 (0.024)	13.45 (0.97)
6	13.89 (1.46)	0.042 (0.019)	-0.109 (0.052)	-0.227 (0.027)	9.72 (1.05)
7	14.46 (1.54)	0.050 (0.019)	-0.063 (0.034)	-0.231 (0.026)	15.00 (1.39)
8	14.18 (1.52)	0.044 (0.019)	-0.073 (0.046)	-0.235 (0.026)	14.74 (1.75)
9	13.32 (1.46)	0.041 (0.014)	-0.059 (0.031)	-0.233 (0.023)	13.06 (0.60)
10	14.18 (1.39)	0.047 (0.016)	-0.060 (0.030)	-0.227 (0.024)	14.34 (0.80)
11	14.09 (1.30)	0.051 (0.015)	-0.081 (0.033)	-0.241 (0.021)	13.93 (0.49)
12	14.62 (1.48)	0.052 (0.019)	-0.040 (0.039)	-0.232 (0.024)	14.13 (0.91)
13	14.23 (1.43)	0.052 (0.019)	-0.074 (0.033)	-0.235 (0.024)	12.57 (1.13)
14	13.85 (1.46)	0.044 (0.017)	-0.065 (0.044)	-0.229 (0.026)	14.22 (1.17)
15	14.79 (1.60)	0.049 (0.017)	-0.055 (0.039)	-0.233 (0.025)	13.75 (1.14)

Table 2.9. Full Hierarchical Model Results
Summary Statistics for Hyper-parameters

Parameter	mean	std	Min	2.50%	median	97.50%	Max
$\bar{\beta}_{\text{constant}}$	14.1600	1.1525	10.6785	11.9980	14.1434	16.5850	18.6576
$\bar{\beta}_{\text{income}}$	0.0467	0.0122	-0.0061	0.0201	0.0474	0.0705	0.1101
$\bar{\beta}_{\text{cost}}$	-0.0670	0.0261	-0.1836	-0.1194	-0.0669	-0.0184	0.0280
$\bar{\beta}_{\text{Age}}$	-0.2349	0.0187	-0.3129	-0.2758	-0.2348	-0.1996	-0.1651
τ^2_{constant}	1.6214	2.6919	0.0000	0.0015	0.8241	8.1980	49.9053
τ^2_{income}	0.0003	0.0006	0.0000	0.0000	0.0001	0.0018	0.0092
τ^2_{cost}	0.0017	0.0022	0.0000	0.0000	0.0010	0.0076	0.0300
τ^2_{age}	0.0005	0.0010	0.0000	0.0000	0.0002	0.0031	0.0157
τ_{constant}	1.0261	0.7541	0.0029	0.0385	0.9078	2.8632	7.0644
τ_{income}	0.0136	0.0113	0.0003	0.0011	0.0106	0.0425	0.0957
τ_{cost}	0.0343	0.0224	0.0002	0.0020	0.0311	0.0873	0.1732
τ_{age}	0.0169	0.0150	0.0001	0.0005	0.0134	0.0560	0.1253

Table 2.10. Comparison of Summary Statistics for Zone 10

(a) Constant Coefficient Estimates

Stage	Mean	std	Min	2.50%	median	97.50%	max
On-site Only	10.306	3.649	-2.228	3.181	10.262	17.615	24.305
Off-site Only	14.402	1.901	-3.251	10.670	14.409	18.237	25.950
25 On-site	14.469	1.917	4.704	11.194	14.378	18.574	27.264
50 On-site	14.441	1.755	5.611	11.261	14.341	18.277	24.805
Full Model	14.176	1.385	8.227	11.319	14.181	16.871	19.280

(b) Income Coefficient Estimates

Stage	mean	std	Min	2.50%	median	97.50%	max
On-site Only	0.040	0.035	-0.096	-0.028	0.040	0.111	0.172
Off-site Only	0.048	0.026	-0.099	-0.008	0.049	0.098	0.234
25 On-site	0.051	0.025	-0.126	0.001	0.050	0.104	0.273
50 On-site	0.049	0.022	-0.121	0.007	0.048	0.098	0.214
Full Model	0.047	0.016	-0.035	0.013	0.047	0.079	0.130

(c) Cost Coefficient Estimates

Stage	mean	std	Min	2.50%	median	97.50%	max
On-site Only	-0.036	0.059	-0.238	-0.153	-0.036	0.080	0.161
Off-site Only	-0.072	0.059	-0.460	-0.198	-0.072	0.050	0.359
25 On-site	-0.072	0.050	-0.412	-0.166	-0.073	0.033	0.296
50 On-site	-0.061	0.052	-0.398	-0.169	-0.060	0.045	0.243
Full Model	-0.060	0.030	-0.172	-0.120	-0.060	-0.002	0.079

(d) Age Coefficient Estimates

Stage	mean	std	Min	2.50%	median	97.50%	max
On-site Only	-0.153	0.057	-0.349	-0.269	-0.152	-0.042	0.059
Off-site Only	-0.238	0.034	-0.496	-0.306	-0.238	-0.167	-0.008
25 On-site	-0.234	0.033	-0.419	-0.294	-0.232	-0.163	0.075
50 On-site	-0.240	0.029	-0.387	-0.290	-0.243	-0.176	-0.012
Full Model	-0.227	0.024	-0.324	-0.273	-0.230	-0.174	-0.106

Table 2.11. Comparison of Summary Statistics for Zone 11

(a) Constant Coefficient Estimates

Stage	mean	std	Min	2.50%	median	97.50%	max
On-site Only	16.575	2.622	6.784	11.445	16.567	21.742	27.280
Off-site Only	13.134	2.005	0.457	9.539	12.862	17.189	30.994
25 On-site	12.916	2.039	1.724	9.238	12.940	17.045	30.777
50 On-site	12.876	1.681	4.211	9.731	12.756	16.411	22.478
Full Model	14.094	1.297	8.927	11.607	14.078	16.697	19.814

(b) Income Coefficient Estimates

Stage	mean	std	Min	2.50%	median	97.50%	max
On-site Only	0.077	0.024	-0.007	0.029	0.077	0.126	0.172
Off-site Only	0.042	0.023	-0.188	-0.004	0.041	0.091	0.276
25 On-site	0.043	0.023	-0.134	-0.003	0.044	0.088	0.211
50 On-site	0.042	0.020	-0.081	0.001	0.043	0.082	0.170
Full Model	0.051	0.015	-0.004	0.022	0.050	0.084	0.115

(c) Cost Coefficient Estimates

Stage	mean	std	Min	2.50%	median	97.50%	max
On-site Only	-0.155	0.059	-0.359	-0.271	-0.155	-0.039	0.072
Off-site Only	-0.050	0.053	-0.584	-0.167	-0.047	0.053	0.407
25 On-site	-0.056	0.050	-0.543	-0.165	-0.057	0.042	0.343
50 On-site	-0.051	0.041	-0.382	-0.143	-0.048	0.017	0.151
Full Model	-0.081	0.033	-0.234	-0.151	-0.079	-0.023	0.041

(d) Age Coefficient Estimates

Stage	mean	std	Min	2.50%	median	97.50%	max
On-site Only	-0.268	0.037	-0.409	-0.341	-0.267	-0.196	-0.137
Off-site Only	-0.218	0.036	-0.542	-0.287	-0.218	-0.150	0.109
25 On-site	-0.220	0.035	-0.518	-0.296	-0.218	-0.156	-0.033
50 On-site	-0.220	0.029	-0.479	-0.284	-0.217	-0.168	-0.048
Full Model	-0.241	0.021	-0.331	-0.286	-0.239	-0.203	-0.173

Table 2.12. Comparison of Summary Statistics for Consumer Surplus Estimates

(a) Zone 10

Stage	mean	std	Min	2.50%	median	97.50%	max
On-site Only	11.383	9.116	2.098	3.089	8.076	38.730	49.991
Off-site Only	8.314	6.436	1.087	2.463	6.437	28.908	49.607
25 On-site	8.332	6.258	1.213	2.970	6.529	27.645	49.993
50 On-site	9.874	7.538	1.255	2.875	7.441	33.535	49.974
Full Model	10.343	6.991	2.910	4.146	8.139	31.877	49.622

(a) Zone 11

Stage	mean	std	Min	2.50%	median	97.50%	max
On-site Only	3.997	3.034	1.392	1.846	3.220	11.143	49.944
Off-site Only	11.404	8.196	0.856	2.877	9.050	36.049	49.743
25 On-site	10.624	7.916	0.921	2.924	7.706	34.822	49.917
50 On-site	11.765	8.134	1.308	3.387	9.005	36.196	49.985
Full Model	7.500	4.481	2.139	3.316	6.310	19.200	49.094

Table 2.13. Aigner-Leamer Transferability Index

Zone	Overall	Individual coefficients			
		Intercept	Income	Cost	Age
1	0.0024	0.5654	0.3248	0.2527	0.4004
2	0.0008	0.5153	0.2387	0.1804	0.4735
3	0.0021	0.5815	0.2741	0.1980	0.4538
4	0.0009	0.5249	0.1917	0.2109	0.4592
5	0.0068	0.5692	0.3587	0.2721	0.5634
6	0.0022	0.6123	0.2758	0.1252	0.4354
7	0.0034	0.5559	0.2779	0.3262	0.4925
8	0.0017	0.5825	0.2887	0.1965	0.5215
9	0.0318	0.5526	0.4189	0.2919	0.6267
10	0.0155	0.6696	0.3642	0.3771	0.5582
11	0.0526	0.7504	0.3862	0.3009	0.7613
12	0.0050	0.5913	0.2733	0.2363	0.5840
13	0.0069	0.6421	0.2828	0.3147	0.5819
14	0.0033	0.6143	0.3437	0.2058	0.4976
15	0.0040	0.4986	0.3543	0.2428	0.5356

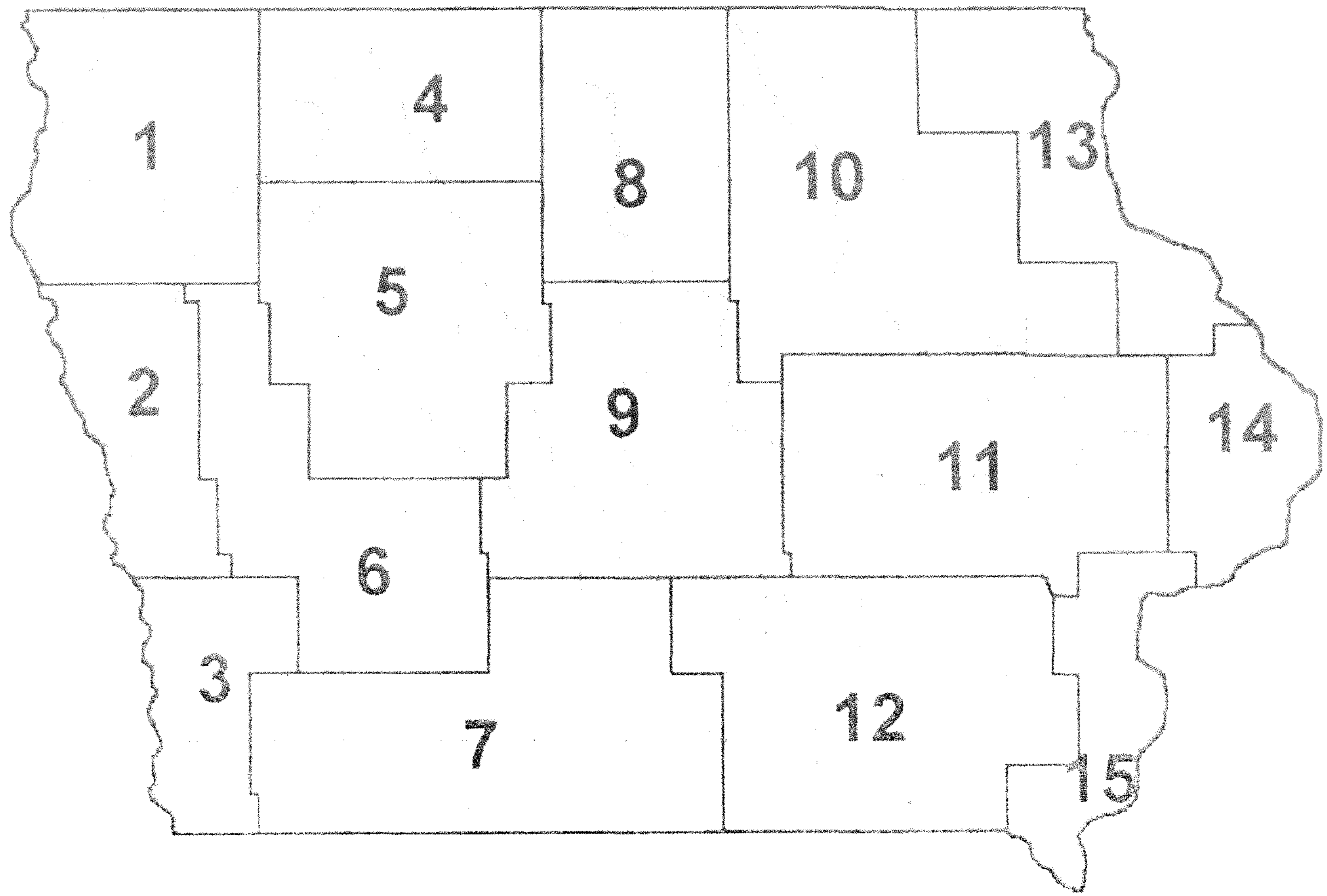


Figure 2.1. Iowa Wetlands Zone Map

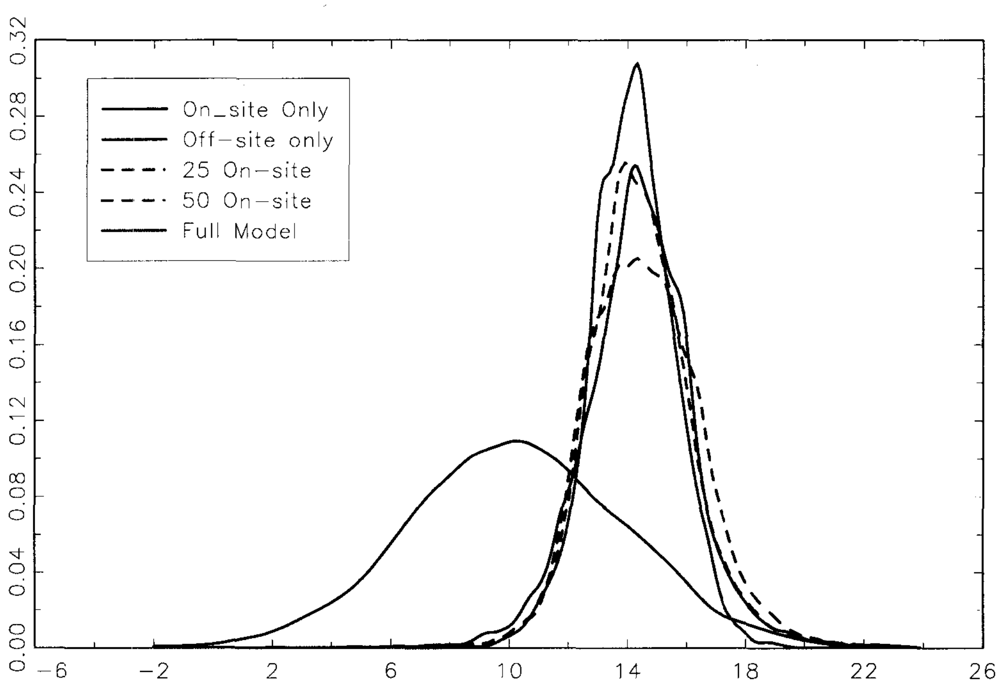


Figure 2.2. Kernel Density Plots for Constant – Zone 10

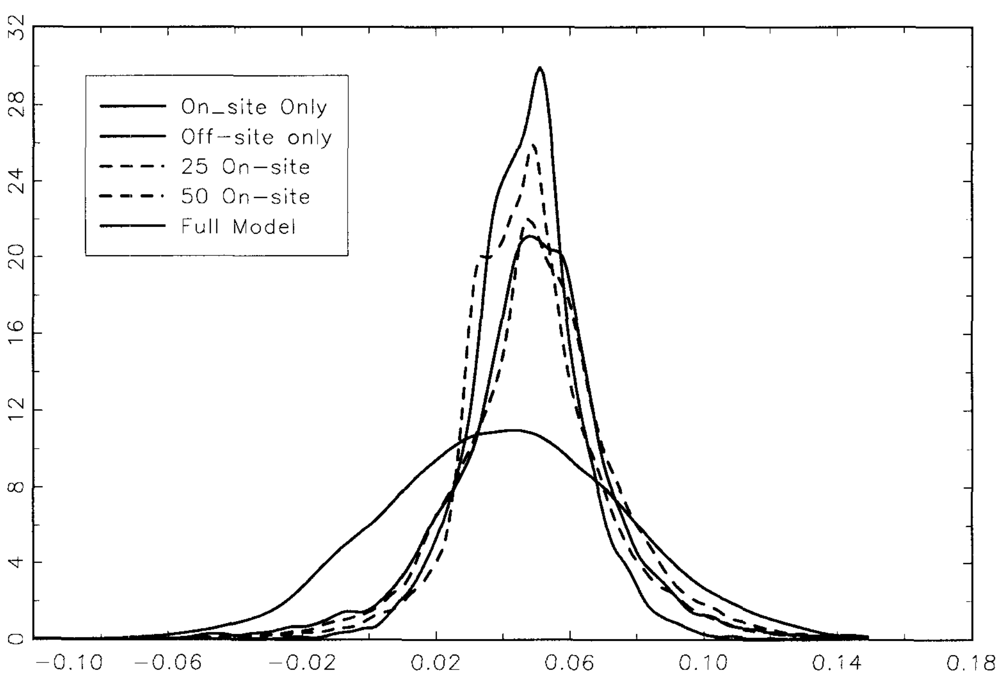


Figure 2.3. Kernel Density Plots for Income Coefficient – Zone 10

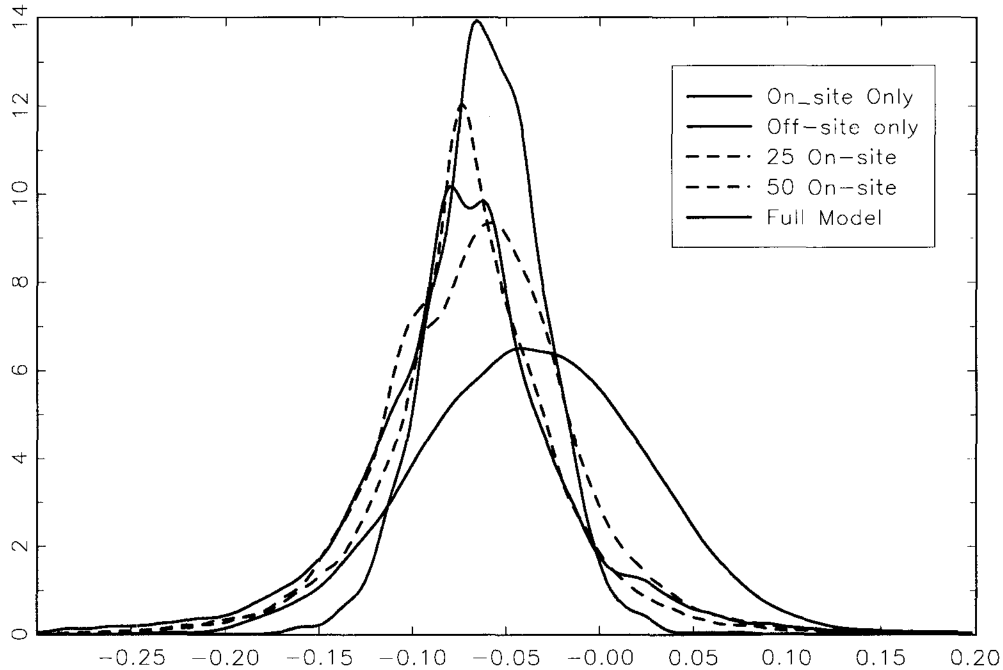


Figure 2.4. Kernel Density Plots for Cost Coefficient – Zone 10

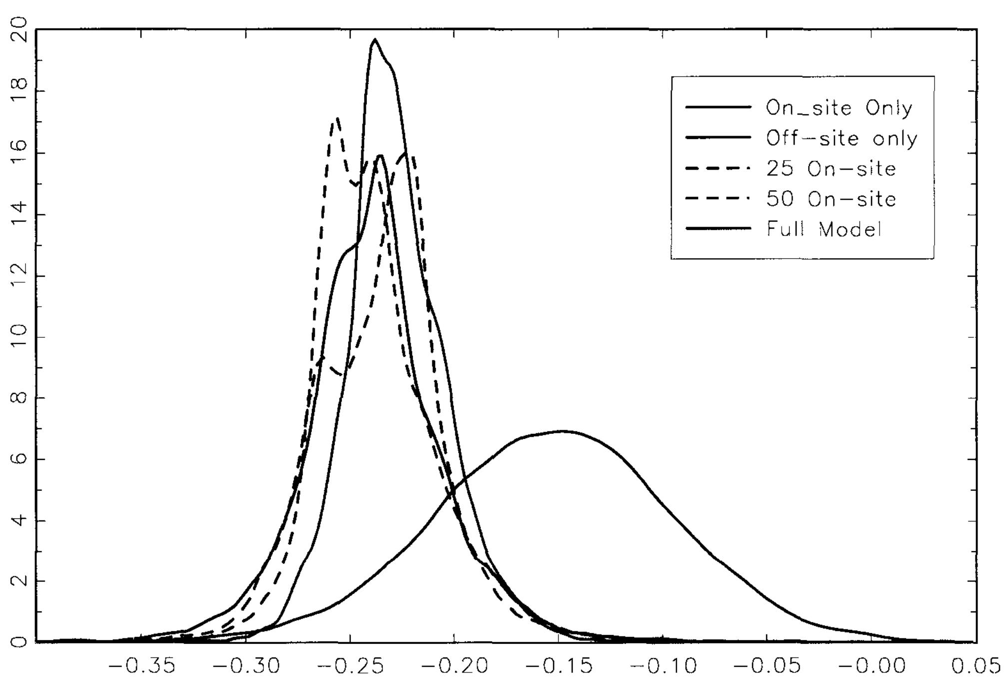


Figure 2.5. Kernel Density Plots for Age Coefficient – Zone 10

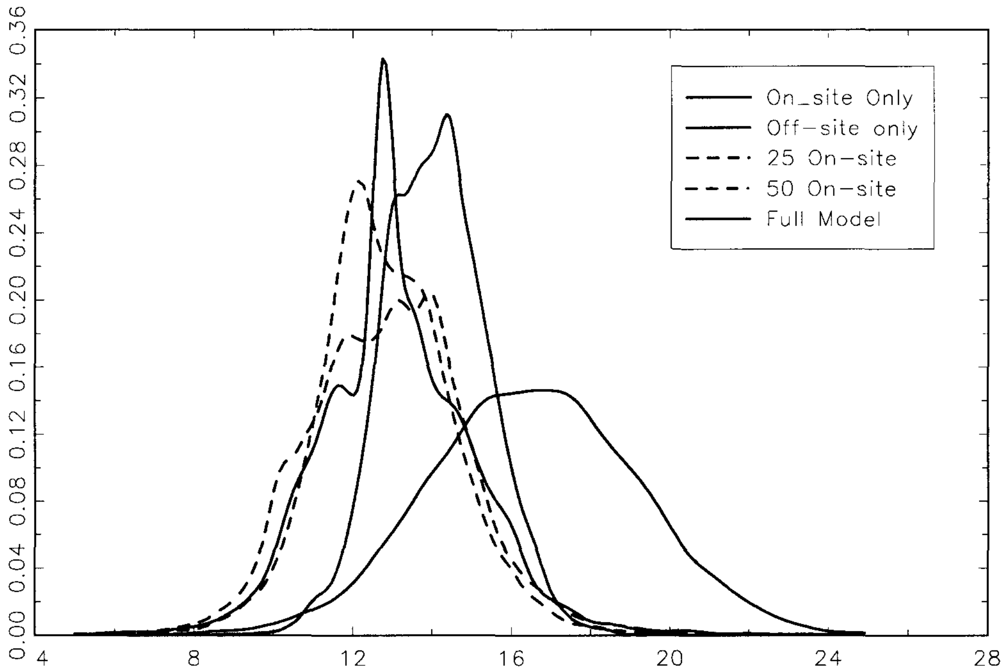


Figure 2.6. Kernel Density Plots for Constant – Zone 11

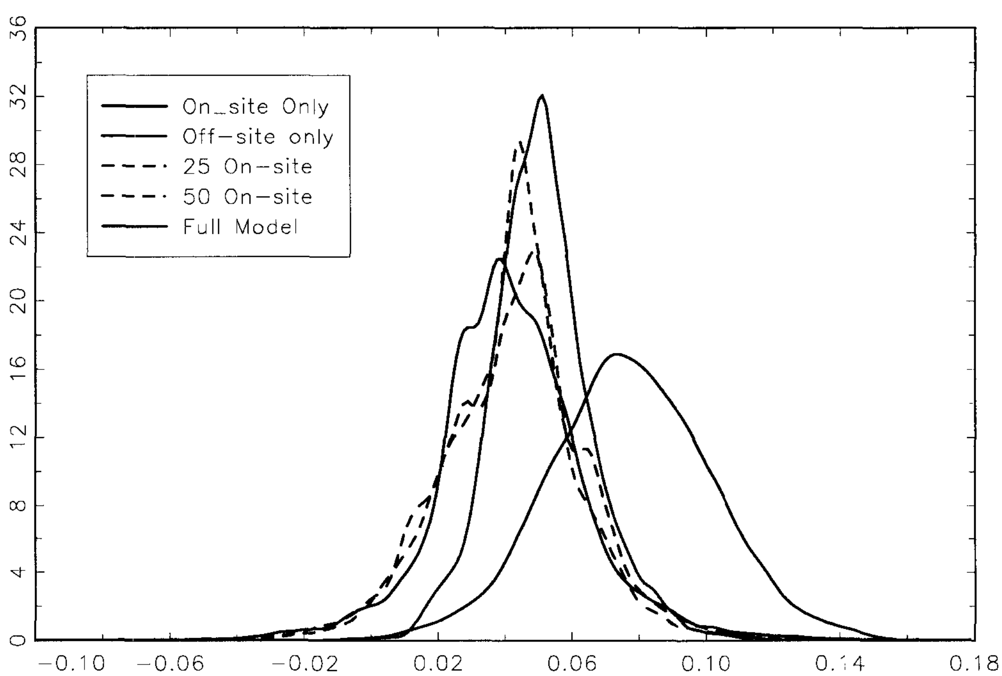


Figure 2.7. Kernel Density Plots for Income Coefficient – Zone 11

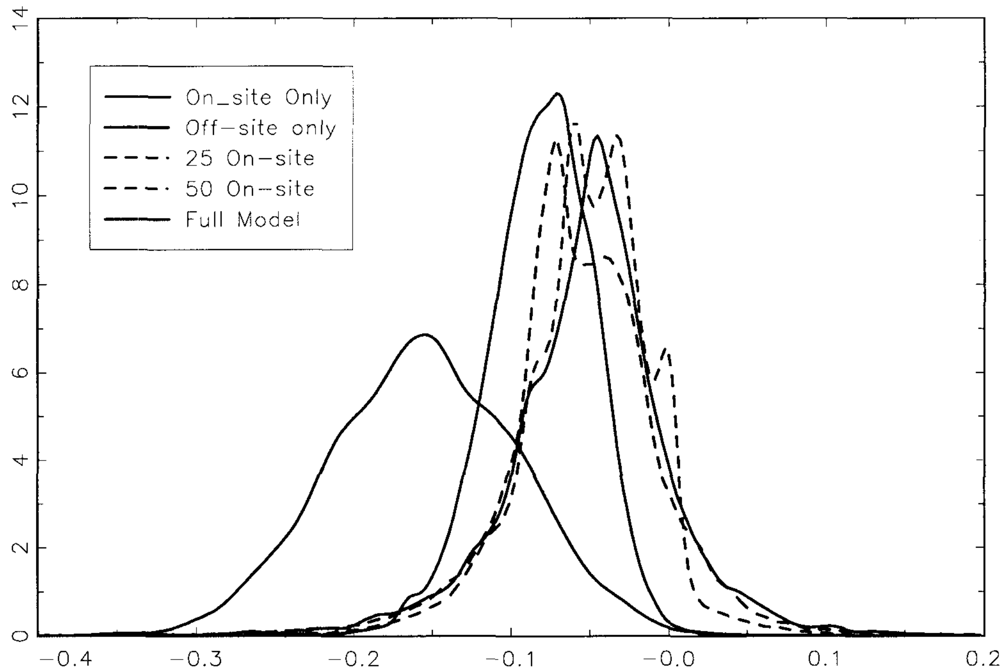


Figure 2.8. Kernel Density Plots for Cost Coefficient – Zone 11

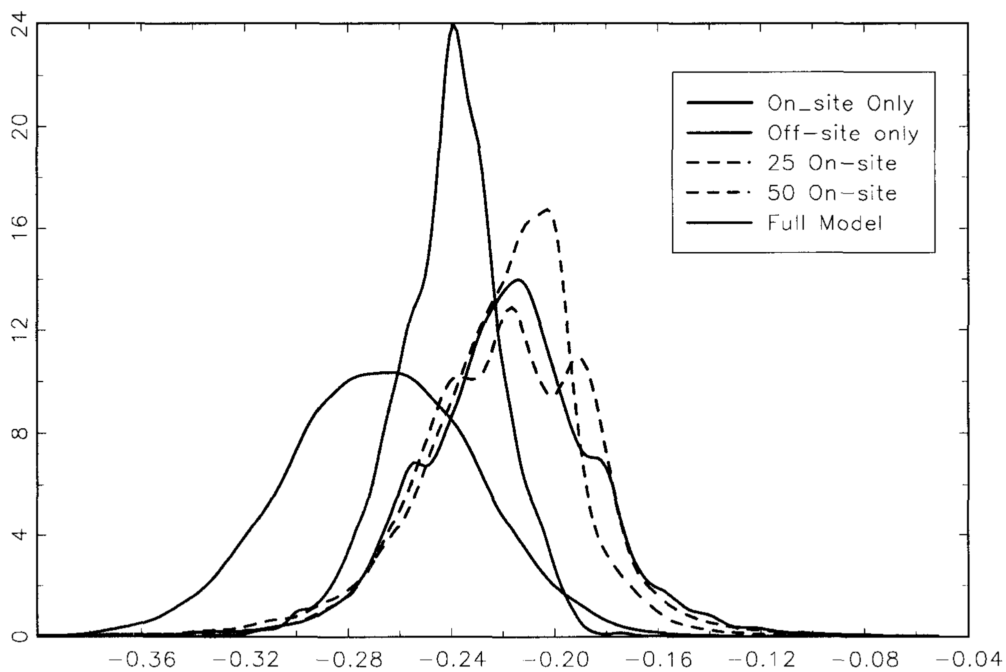


Figure 2.9. Kernel Density Plots for Age Coefficient – Zone 11

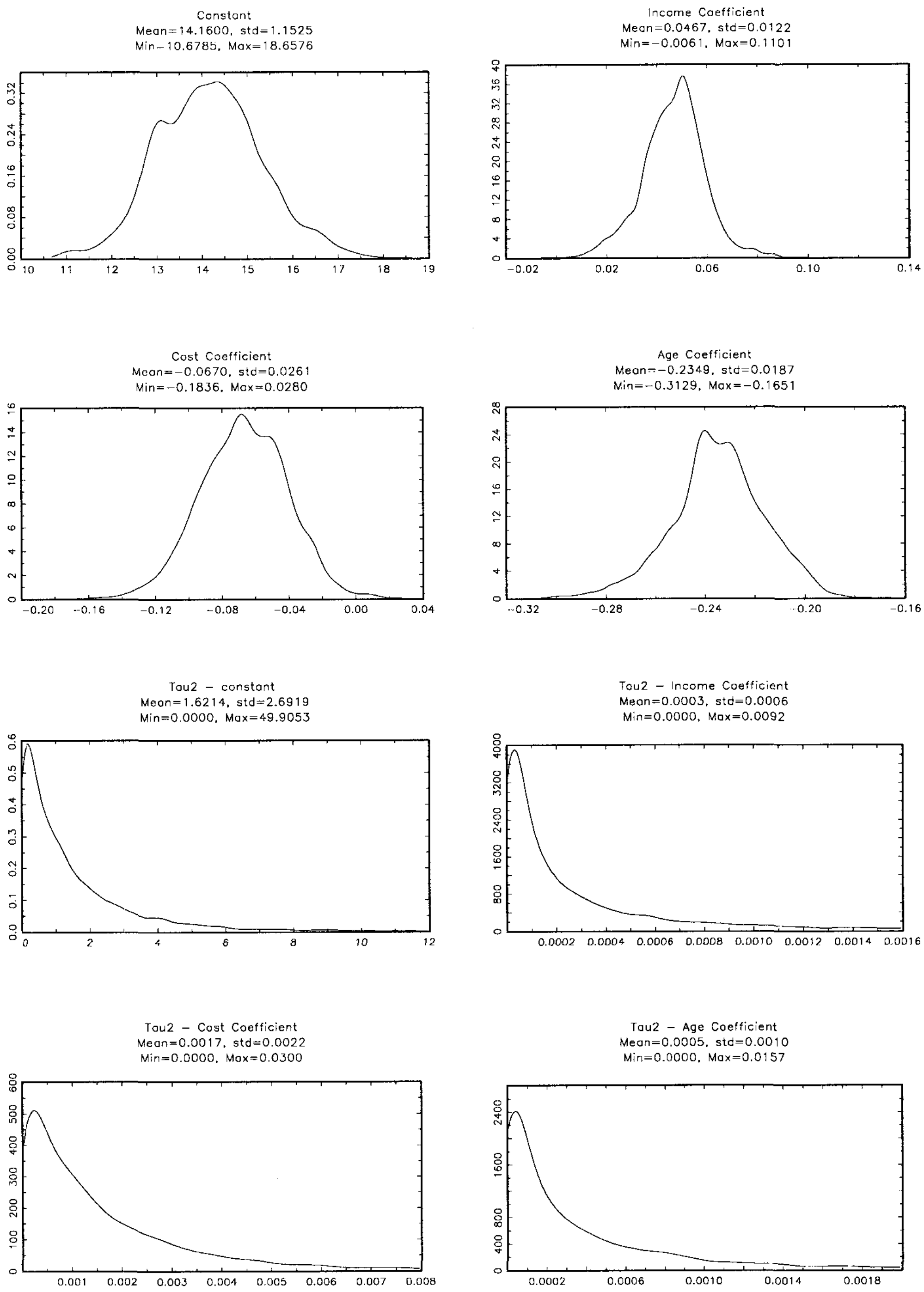


Figure 2.10. Kernel Density Plots for Hyper-parameters

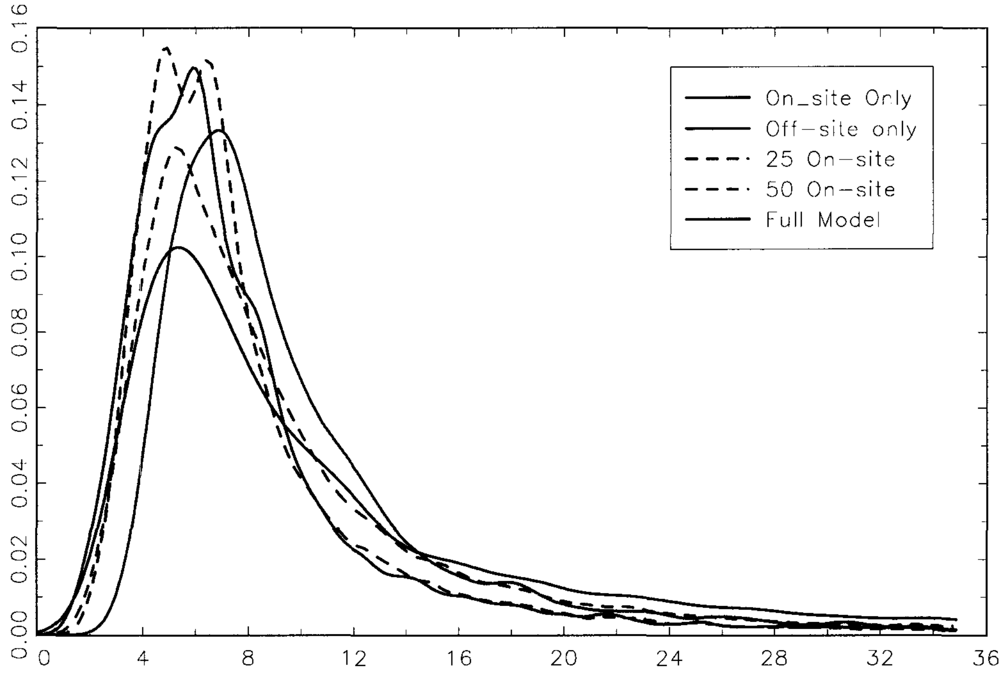


Figure 2.11. Kernel Density Plots for Consumer Surplus – Zone 10

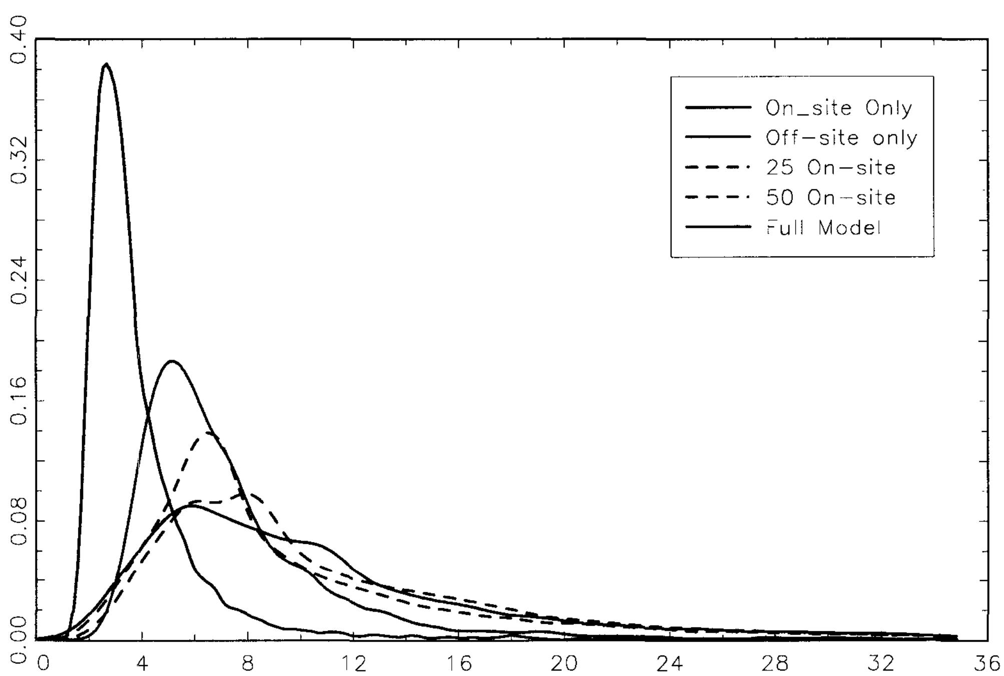


Figure 2.12. Kernel Density Plots for Consumer Surplus – Zone 11

Appendix 2.A RESULTS FROM BOOTSTRAPPING STUDY

(a) Tobit Estimation with Bootstrap Data (Standard Error is shown in parenthesis)

Zone	Constant	Income	Cost	Age	Sigma
1	23.23	-0.007	0.027	-0.486	16.40
	(3.40)	(0.033)	(0.070)	(0.061)	(0.83)
2	41.87	-0.065	-0.629	-0.451	17.63
	(5.16)	(0.076)	(0.172)	(0.069)	(0.90)
3	13.17	0.020	0.118	-0.311	9.96
	(2.32)	(0.018)	(0.040)	(0.042)	(0.50)
4	9.34	0.262	-0.045	-0.254	15.53
	(3.79)	(0.048)	(0.036)	(0.062)	(0.71)
5	24.40	-0.016	-0.072	-0.400	14.71
	(3.81)	(0.033)	(0.090)	(0.057)	(0.69)
6	13.72	0.114	-0.475	-0.062	8.68
	(2.42)	(0.025)	(0.059)	(0.031)	(0.43)
7	16.40	0.085	-0.182	-0.192	12.13
	(2.74)	(0.027)	(0.048)	(0.046)	(0.55)
8	30.08	-0.028	-0.447	-0.346	14.76
	(4.60)	(0.045)	(0.173)	(0.056)	(0.69)
9	9.83	0.071	-0.099	-0.173	11.69
	(2.71)	(0.023)	(0.065)	(0.044)	(0.56)
10	13.13	0.044	-0.070	-0.180	13.17
	(2.87)	(0.027)	(0.052)	(0.044)	(0.61)
11	18.48	0.043	-0.137	-0.278	14.31
	(3.63)	(0.033)	(0.075)	(0.051)	(0.70)
12	10.05	0.070	-0.021	-0.138	13.70
	(2.88)	(0.036)	(0.059)	(0.046)	(0.61)
13	18.16	0.089	-0.096	-0.324	13.81
	(3.09)	(0.029)	(0.033)	(0.057)	(0.68)
14	1.17	0.036	0.041	0.014	14.02
	(3.22)	(0.029)	(0.128)	(0.046)	(0.64)
15	30.40	-0.029	-0.076	-0.449	13.26
	(2.92)	(0.027)	(0.060)	(0.050)	(0.60)
Mean	18.23	0.046	-0.144	-0.269	13.58
Std	10.33	0.079	0.210	0.147	2.31

(b) Full Hierarchical Model Results with Bootstrap Data
Posterior Means and Standard Deviations*

Zone	Constant	Income	Cost	Age	Sigma
1	21.05 (3.05)	0.011 (0.030)	-0.002 (0.064)	-0.442 (0.054)	16.47 (0.83)
2	30.69 (4.29)	-0.056 (0.056)	-0.409 (0.130)	-0.355 (0.060)	17.72 (0.92)
3	13.51 (2.18)	0.022 (0.018)	0.102 (0.040)	-0.315 (0.040)	10.09 (0.51)
4	13.17 (3.24)	0.183 (0.047)	-0.049 (0.035)	-0.271 (0.052)	15.63 (0.70)
5	22.41 (3.25)	0.000 (0.030)	-0.080 (0.079)	-0.369 (0.051)	14.85 (0.72)
6	13.89 (2.34)	0.100 (0.024)	-0.441 (0.057)	-0.075 (0.031)	8.76 (0.43)
7	16.75 (2.55)	0.078 (0.025)	-0.175 (0.046)	-0.201 (0.043)	12.26 (0.56)
8	24.26 (3.57)	-0.020 (0.037)	-0.284 (0.126)	-0.304 (0.049)	14.85 (0.69)
9	11.15 (2.52)	0.068 (0.022)	-0.107 (0.060)	-0.194 (0.041)	11.84 (0.56)
10	14.22 (2.63)	0.045 (0.025)	-0.082 (0.048)	-0.196 (0.041)	13.32 (0.62)
11	18.14 (3.16)	0.045 (0.030)	-0.133 (0.066)	-0.276 (0.047)	14.47 (0.72)
12	11.81 (2.65)	0.066 (0.032)	-0.036 (0.055)	-0.163 (0.044)	13.87 (0.63)
13	17.90 (2.78)	0.083 (0.027)	-0.093 (0.031)	-0.317 (0.052)	13.95 (0.68)
14	5.44 (3.00)	0.037 (0.026)	-0.053 (0.105)	-0.035 (0.046)	14.21 (0.66)
15	27.67 (2.76)	-0.012 (0.025)	-0.085 (0.057)	-0.406 (0.047)	13.36 (0.60)
$\bar{\beta}$	17.47	0.043	-0.129	-0.262	
$\bar{\tau}$	8.06	0.071	0.178	0.138	

*The last two rows provide the posterior means for hyper-parameters.

Chapter 3
Combining Revealed and Stated Preference

3.1 INTRODUCTION

In order to predict consumers' market behavior response to changing environmental conditions or to estimate the social consequences of these changes for non-market goods, surveys are designed to elicit revealed preference (RP) or stated preference (SP) towards environmental goods and/or services. Since RP data is based on the individuals' actual decisions or choices under the present market or non-market conditions, while SP data is based on the responses to hypothetical scenarios, models based on these two types of data often result in different estimates of consumer preferences.

Historically, analysts have viewed the RP and SP approaches as two competing methodologies. For market valuation studies, revealed preferences are typically considered the preferred source of value information [15]. This viewpoint is "naturally" adopted in nonmarket valuation studies by many researchers; and the welfare estimates from RP models are often viewed as a benchmark against which the estimates from SP models are "checked". More recently, however, an alternative view has emerged; i.e., that both RP and SP methods tend to reveal imperfect insights into the same underlying preferences, each having their own strength and weaknesses. By combining the two methods, a more complete picture of preferences will emerge.

Bayesian analysis provides a natural tool for integrating different sources of information, and thus represents a promising framework for combining RP and SP data. If treating the RP and SP data as independent data sources, one might view the estimates from SP data as providing the basis for a prior distribution on preferences and then use the RP data to update this prior. However, this would treat the two sources of information as independent, which might be too strong a restriction. A alternative way to combine RP and SP data is to model them jointly, allowing for correlation between these two data sources. The purpose of this essay is to develop a Bayesian framework to link together these two types of information sources. By introducing a set of discrepancy parameters, one can easily incorporate into the analysis prior beliefs about the disagreement between RP and SP data as well as carry out hypothesis tests in regards to that disagreement within a Bayesian context.

The remainder of this chapter has five sections. Section 3.2 provides a literature review, while section 3.3 describes the data used in the analysis. Details of the model specification and the methodologies used in the analysis are provided in section 3.4. Results and conclusions are provided in the last two sections of the chapter.

3.2 LITERATURE REVIEW

As noted above, little attention was historically paid to integrating revealed and stated preference information. Instead, RP and SP were viewed as two competing methodologies in the environmental recreation literature. Analysts focused on contrasting value estimates derived from these two methods. For example, Bishop and Heberlin (1979) [10], Bishop *et al.* (1983) [11], and Sellar *et al.* (1985) [44] compared the results from contingent valuation method (CVM) with those from travel cost model (TCM). In their studies, external validity tests were carried out and the TCM results were used to validate the results form the CVM method.

The move to combining RP and SP data sources is more recent. The argument for combining (as noted by Adamowicz, Louviere, and Williams [1]), is based on the fact that each of two approaches have advantages and drawbacks. "Direct [SP] methods are commonly criticized because of the hypothetical nature of the questions and the fact that actual behavior is not observed (Cummings *et al.* [18], Mitchell and Carson [37]). However, direct methods currently provide the only viable alternative for measuring non-use values and they are commonly used to elicit values in case in which the environmental quality change involves a large number of attribute changes. Indirect methods avoid the criticism of being based on hypothetical behavior, but the models of behavior developed constitute a maintained hypothesis about the structure of preferences which may or may not be testable. Indirect [RP] methods may also suffer on the grounds that the new situation (after environmental quality change) may be outside the current set of experiments (or outside the range used to estimate the model. Finally, indirect method methods may suffer from colinearity among attributes." ([1], p. 272) Here, the direct and indirect methods refer to RP and SP methods respectively. Since these two methods both reflect the process of choosing recreation sites based on attributes, they complement each other to some degree and by combining we may obtain a better overall perspective on consumer preferences. Since there is little research into combining RP and SP data based on Bayesian methodologies, I will focus on reviewing the classical approaches in this section. This section is divided into two parts. In the first part, I review some of the prominent studies aimed at combining RP and SP data. The second part considers those combining efforts that draw on discrete choice models which allow for differing scale effects for the RP and SP data.

3.2.1 Combining CVM Model and TCM Model

Cameron was one of the first to argue for the combining of RP and SP data. In her paper [14] she develops a prototypical empirical model to combine the *RP* data and *SP* data. The underlying logic is that both the *RP* data and *SP* data come from the consumers' optimization of same underlying utility function, though under different conditions (existing conditions for *RP*, hypothetical conditions for *SP*).

Cameron uses survey data collected on recreational fishing. In the study, a referendum CVM question was used that can be interpreted as asking whether the respondent would quit fishing if an annual access fee ("tax") T is charged. To model consumers' response, Cameron starts with a direct utility function

$$U(z,q) = U(Y - Mq, q), \tag{3.1}$$

where Y denotes the respondent's income, q is the number of trips per year to the recreation site, M is the typical travel costs, z is the composite of all other goods and services. Given this structure of preferences, an individual responding to the CVM question would indicate a willingness to continue fishing in the face of an access fee of T if

$$\Delta U(Y,M,T,q) = \max_{q} U(Y - Mq - T, q) - U(Y,0) > 0. \tag{3.2}$$

Since $\Delta U(Y,M,T,q)$ in equation (3.2) cannot be actually observed, Cameron uses the assumption that

$$\Delta U_i = f(x_i,\beta) + \varepsilon_i, \tag{3.3}$$

where $\varepsilon_i \sim N(0,\sigma^2)$, $f(x_i,\beta)$ is the systematic portion of the utility difference on right-hand side of equation (3.2), x_i denotes individual characteristics (including travel costs), and β denotes unknown parameters. Hence the log-likelihood function for the SP portion of the data is:

$$\log L^{SP} = \sum_i \{I_i \log[\Phi(f(x_i,\beta)/\sigma)] + (1-I_i)\log[1-\Phi(f(x_i,\beta)/\sigma)]\}, \tag{3.4}$$

where I_i is an indicator variable.

The separate TCM model comes from the maximization problem

$$\max_q U(Y-Mq,q) \quad \text{s.t.} \quad Y = z + Mq. \tag{3.5}$$

The solution of equation (3.5) implies the ordinary demand function

$$q_i = g(x_i,\beta) + \eta_i, \tag{3.6}$$

where

$$\eta_i \sim N(0,\nu^2). \tag{3.7}$$

As noted by Cameron, here the stochastic structure is ad hoc, since the ordinary demand function will not have error term. Instead, the error term is added to capture measurement or recall error on the part of the survey respondent. The log-likelihood function for the RP portion of the data becomes

$$\log L^{RP} = -(n/2)\log(2\pi) - n\log\nu - 1/2\sum_i\{[q_i - g(x_i,\beta)]/\nu\}^2. \tag{3.8}$$

Since both the CVM model and TCM model originate from the direct utility function (3.1) it is possible to estimate the coefficients jointly by imposing the constraint that the corresponding coefficients in two models are the same. If the error terms ε_i and η_i are independent, then the joint log-likelihood function will be just the summation of the two separate ones; i.e.,

$$\log L = \log L^{SP} + \log L^{RP}. \tag{3.9}$$

But the independent error structure may be unrealistic, since the unobservable factors that affect the respondents' actual demand for fishing trips are simultaneously likely their answers to the contingent behavior questions. It is more reasonable to allow a correlation ρ between the error terms ε_i and η_i. Using the fact that the joint density can be represented as the product of marginal and conditional densities, and the properties of bivariate normal distribution, Cameron derives a more general version of the joint log-likelihood function,

$$\log L = \log L^{RP} + \sum_i \{I_i \log[\Phi(A)] + (1 - I_i)\log[1 - \Phi(A)]\}, \tag{3.10}$$

where

$$A = [f(x_i, \beta)/\sigma + \rho(q_i - g(x_i, \beta))/\nu]/(1 - \rho^2)^{1/2}. \tag{3.11}$$

In the empirical example, Cameron employs a quadratic utility function

$$\begin{aligned} U(z, q) &= \beta_1 z + \beta_2 q + \beta_3 z^2/2 + \beta_4 zq + \beta_5 q^2/2 \\ &= \beta_1(Y - Mq) + \beta_2 q + \beta_3(Y - Mq)^2/2 + \beta_4(Y - Mq)q + \beta_5 q^2/2, \end{aligned} \tag{3.12}$$

and derives the separate models and the joint model for this utility function. Finally, maximum likelihood estimation is performed and welfare measurements are calculated. Cameron concludes that travel cost and contingent valuation data can be usefully combined in one joint model of preferences.

Huang *et al.* [26] also study the relationship between willingness to pay (WTP) CVM and travel demand model. The contingent valuation question they are concerned about is the willingness to pay for a quality improvement. Based on the comparative static analysis of variation function, they derive an analytical model to describe the relationship between the CVM and TCM models. Two recreation areas in North Carolina (Albemarle and Pamlico Sounds) are studied. They use a simple linear variation function to describe the WTP:

$$w_i = \alpha + \beta P_i + (\gamma - 1)Y_i + \delta_1 D + \varepsilon_{1i}, \tag{3.13}$$

where w_i is the Hicksian variation for a specified quality change, P_i is the cost of the trip, Y_i is the income, D is a dummy variable to indicate the component good. According to this model, if w_i is greater than the dollar amount requested in exchange for the improvement, then the respondents answer "yes". From equation (3.13) and the analytical results, a recreation demand function is derived as

$$x_i = \beta + \gamma(x_i^* + \delta_2 D + \varepsilon_{2i}), \tag{3.14}$$

where x is the number of trips taken, $x^* + \varepsilon_2$ is the ex ante number of trips stated for a quality improvement, ε_2 is the measurement error. Note that here the error term for the trip demand is defined differently from (3.7). By imposing a bivariate normal distribution $N(0, 0, \sigma_1^2, \sigma_2^2, \rho)$ on the error terms ε_1 and ε_2, a joint log likelihood function analogous to equation (3.11) in Cameron [14] is derived. In their empirical study, they reject the hypothesis that the parameters β and γ are equal across the equations (3.13) and (3.14). They suggest that the joint estimation or RP and SP data should proceed with caution.

Whitehead *et al.* [56] propose using panel recreation demand models to combine revealed and stated behavior data for measuring recreation benefits. The data are from 1995 telephone survey of eastern North Carolina households that proposed a management plan to restore Albemarle and Pamlico Sound resources. The survey has two versions: one focused on Pamlico Sound, while the other focused on both Albemarle and Pamlico Sound. The survey respondents were asked about trips taken or anticipated under three circumstances:

Scenario 1: current recreation participation and intensity with current quality levels ($t = 1$);

Scenario 2: expected recreation participation and intensity with current quality levels ($t = 2$);

Scenario 3: expected recreation participation and intensity with hypothetically improved quality ($t = 3$).

A Poisson regression model is used to analyze these panel data. Let x_{it} denote the number of trips taken by individual i ($i = 1, ..., N$) in a particular trip scenario t ($t = 1, 2, 3$), μ_{it} denote the mean that depends on the explanatory variables, then the model used by Whitehead *et al.* is given as

$$p(X_{it} = x_{it}) = \frac{e^{-\mu_{it}} \mu_{it}^{x_{it}}}{x_{it}!} \qquad x_{it} = 0, 1, 2, ..., \tag{3.15}$$

and

$$\ln \mu_{it} = \alpha_t + \beta_t TCP_{it} + \delta_t TCF_{it} + \phi_t INC_{it} + \varphi_t Paml_{it} + u_i, \tag{3.16}$$

where $t = 1, 2, 3$, μ_i is a random effect for individual i, *TCP* is trip cost to Pamlico sound, *TCF* is trip cost to Cape Fear, *INC* is income, *Paml* equals to 1 for Pamlico version and 0 otherwise. In order to pool the three sets of data together, two sets of dummy variables were also introduced. Their hypothesis testing results indicate that revealed and stated behavior data with current quality do not belong to the same model.

3.2.2 Allowing for Differing Scale Effects

The information provided by SP and RP data, while based on the same underlying preferences, may have differing degrees of variability. Morikawa [39] was one of the first to take this into account when combining RP and SP models. In his Ph.D. thesis, Morikawa proposes a method to combine SP and RP data in the discrete choice model framework. He uses the following structural functions to associate individual's utilities with a single trip to a recreation site for RP and SP data respectively:

$$U_{RP} = \beta' X + \gamma'_{RP} W_{RP} + \varepsilon_{RP}; \tag{3.17}$$

$$U_{SP} = \beta' X + \gamma'_{SP} W_{SP} + \varepsilon_{SP}, \tag{3.18}$$

where X is a vector of observed variables common to both data sets, W_{RP} and W_{SP} are vectors of observed variables specific to one data set or the other, β, γ_{RP} and γ_{SP},are unknown parameters vectors with β assumed to be the same for both RP and SP data. Morikawa assumes that the error terms ε_{RP} and ε_{SP} are independently distributed with zero means and finite variances. The available information then takes the form of the discrete choice variables

$$d_i^{RP} = \begin{cases} 1 & U_{RP} \geqslant 0 \\ 0 & U_{RP} < 0 \end{cases}$$

and

$$d_i^{SP} = \begin{cases} 1 & U_{SP} \geqslant 0 \\ 0 & U_{SP} < 0\,. \end{cases}$$

Each of the decisions can be analyzed separately using standard discrete choice methods (e.g., logit or probit). In this case, the parameters of each model can only be identified up to a scale factor. Morikawa notes, however, that joint estimation would allow for the identification of the relative variances of ε_{RP} and ε_{SP}. Specifically, after normalizing one of the variances to unity, δ can be identified, where δ is defined as

$$\delta = \sqrt{var(\varepsilon_{RP})/var(\varepsilon_{SP})}. \tag{3.19}$$

There are a variety of reasons one might expect δ to differ from one. For example, stated preference questions often ask individuals about situations on environmental conditions that are new to them. This may cause them to be less certain about their preference or responses or to rely on factors unobserved by the analyst. All of these would tend to increase the variability of preferences in the sample and, hence, the size of $var(\varepsilon_{SP})$. This would cause $\delta \in (0, 1]$. In any case, Morikawa suggests two methods to carry out the joint estimation procedure. One method is traditional maximum likelihood estimation, which maximizes the summation of the log-likelihood functions for those two models. Another method is sequential estimation.

A number of authors have since allowed for differing scale efforts in RP/SP studies. Adamowicz *et al.* (1994) [1] combine RP and SP in an analysis of recreation demand, where SP data is based on hypothetical choice sets and the RP data from survey actual choice sets. The authors specify a RUM model in which the systematic component V_{in} is assumed to be a linear function of the attribute variables,

$$V_{in} = \beta_1 + \beta_2 x_{in2} + ... + \beta_k x_{ink} + \alpha(Y - P_i), \tag{3.20}$$

where x_{ink} are attributes of the site, Y is income and P_i is the travel cost. The error terms are assumed to be i.i.d. type I extreme value for each of the RP and SP data sets, resulting in a MNL specification for each.

After examining the separate models, Adamowicz *et al.* [1] find a significant correlation between the predicted proportions of visits to each site under these two models. Viewing that as evidence that the underlying revealed and stated preferences are similar, they suggest that the difference of the two models may lie in differing scale factors. This leads them to specify a joint model with different scales:

$$\text{Revealed} \quad : \qquad p(i) = \exp^{\mu_r V_i} / \sum_{j \in C_n} \exp^{\mu_r V_j}, \tag{3.21}$$

$$\text{Stated} \quad : \qquad p(i) = \exp^{\mu_s V_i} / \sum_{j \in C_n} \exp^{\mu_s V_j}, \tag{3.22}$$

Again, a single multinomial logit model, it is impossible to identify the scale factor and one can only estimate the parameters up to a scale factor. When jointly estimated, however, the ratio of the scales, μ_r/μ_s, is identified. The authors use a grid search method suggested by Swait and Louviere [47] to calculate for maximum likelihood estimates. In their likelihood ratio test, the null hypothesis of equal parameters is not rejected after incorporating the relative scale effect. However, the null hypothesis of equal scale factor is clearly rejected. Adamowicz *et al.* (1997) [2] subsequently use the same method as in [1] to jointly examine the revealed and stated preferences in the context of recreation site choice for moose hunting in Alberta, Canada.

A number of authors, including Cameron *et al.* [15], have employed similar methods and found consistency between RP and SP methods once different scale factors were permitted. However, Azevedo, Herriges, and Kling [6] and Azevedo [5] find inconsistency remains even when differing scale factors are used. They assume that individual i's (Marshallian) demand for a single recreation good takes the following form:

$$q_i = f(p_i, y_i; \beta) + \varepsilon_i, \tag{3.23}$$

where q_i denotes the quantity consumed by individual i, p_i denotes the associated price, y_i is the individual i's income, and β is a vector of unknown. The error term ε_i is assumed to be drawn from a normal distribution and used to capture heterogeneity across individuals. In addition, q_i is censored from the left at zero (only non-negative value is observed). For RP and SP data, the demand functions can be written as

$$q_i^R = f^R(p_i^R, y_i^R; \beta^R) + \varepsilon_i^R;$$
$$q_i^S = f^S(p_i^S, y_i^S; \beta^S) + \varepsilon_i^S,$$

with superscript R and S denoting RP and SP models respectively. As in Cameron [14], with $(\varepsilon_i^R, \varepsilon_i^S)$ assumed to be from the bivariate normal distribution $N(0, 0, \sigma_R^2, \sigma_S^2, \rho)$, RP and SP models can be combined and the likelihood function can be derived accordingly. In their application to the Iowa Wetlands Survey Study, simple linear forms are specified for these demand functions:

$$q_i^R = \beta_0^R + \beta_p^R p_i^R + \beta_y^R y_i^R + \varepsilon_i^R; \tag{3.24}$$
$$q_i^S = \delta_0 \beta_0^R + \delta_p \beta_p^R p_i^S + \delta_y \beta_y^R y_i^S + \varepsilon_i^S; \tag{3.25}$$

where δ_j's measure the discrepancies between the RP and SP demand equation parameters and facilitate the consistency tests. Their hypothesis tests reject the consistency between RP and SP data regardless of whether or not the constraint of equal variances ($\sigma_R^2 = \sigma_S^2$) is imposed or not. They also find that the underlying variability of preferences revealed by the two data sets is similar, with $\delta_\sigma \equiv \sigma_S / \sigma_R$ estimated to be close to one, and the correlation between these two data sources is substantial and significant different from zero with ρ lying between 0.62 and 0.73. In a related study, Azevedo [5] links the RP and SP wetland data using a discrete choice model for SP data. His findings are similar to those of Azevedo, Herriges, and Kling [6].

3.3 DATA SOURCE

In this essay, I investigate the use of a Bayesian framework for combining revealed and stated preferences in the context of recreation demand. The data come from the same source as in Chapter 2; i.e., the Iowa Wetlands Survey Study. Two key differences exist in the data sources used here versus what was used in Chapter 2. First, while Chapter 2 used only RP data (i.e., actual trips), this chapter uses SP data as well (i.e., predicted trips reported by survey respondents). Specifically, as is commonly done in the literature, the counterparts of SP data are based on individuals' responds to hypothetical scenarios that relate individuals' usage of wetlands to changes of travel costs. In the Iowa Wetlands Survey, respondents were asked how their behavior would change if (1) the total cost per trip of each of their trips had been $15 more, and (2) the total cost per trip to visit the megazone where they lived had been $15 more.[7]

The second key difference is the level of zone aggregation. As mentioned in Chapter 2, the state is divided into 15 zones based roughly on the Iowa crop reporting districts. In order to keep the analysis tractable in this chapter, these zones are aggregated. Specifically, the 15 zones are further grouped into five megazones according to their geographic positions and similarities: the Missouri River Region (1,2,3), the Prairie Pothole Region (4,5,8), the Iowa River Corridor Region (9,10,11), the Mississippi Region (13,14,15), and the remainder of the state (6, 7, 12).

Three explanatory variables are considered in the analysis: income, out of pocket travel cost and the opportunity cost of time. The out of pocket travel cost to each zone is calculated

---

[7] The amount of money varied across surveys, with bid values $5, $10, and $15 each randomly assigned to 20% of the sample and bid values of $20, $30, $40, and $50 each assigned 10% of the sample.

based on travel distance (priced at 0.21 cents per mile) and the opportunity cost is based on the product of travel time and individual's wage rate. These two variables then are integrated into megazone level with a weighted averaging calculation (weighting zonal data by the proportion of total trips taken to the zone). The summary statistics for these data are provided in Table 3.1. A copy of Iowa Wetlands Survey is attached in Appendix for reference.

3.4 BAYESIAN APPROACH

The objective of this second essay is to develop a Bayesian framework for integrating RP and SP data on recreation demand of the form elicited in the Iowa wetlands survey. The analysis begins with a simple Tobit regression model of visits to the wetland as in Chapter 2; i.e.,

$$y_i^k = \begin{cases} (\beta^k)' x_i^k + \varepsilon_i^k & \varepsilon_i^k > -(\beta^k)' x_i \\ 0 & \varepsilon_i \leq -(\beta^k)' x_i \end{cases} \tag{3.26}$$

and

$$\varepsilon_i^k \sim iid\ N(0, \sigma_k^2), \qquad k = RP, SP. \tag{3.27}$$

where x_i^k is a vector of individual characteristic variables such as age, travel cost to the site and income and β^k denotes the unknown parameter vector of interest, with the superscript and subscript k to index the RP or SP model. As in Chapter 2, in order to carry out Bayesian analysis for this model, a data augmentation procedure is used to handle the censoring problem. The object of our Bayesian analysis (RP and SP trips) then becomes a standard system of two linear regression models.

I begin this section with a discussion of a linear regression model for combining RP and SP data based on the augmented data (i.e., abstracting from the censoring problem).

Subsection 3.4.2 then considers the alternative prior distributions regarding the relationship between RP and SP preferences. Methods for comparing the data's support for these alternative priors are outlined in section 3.4.3. Finally, subsection 3.4.4 outlines the Gibbs sampling and data augmentation procedures used to implement our model for the wetland data.

3.4.1 Model Specification

Consider a basic linear model for the augmented RP and SP data:

$$y_i^{RP} = (\beta^{RP})' x_i^{RP} + \varepsilon_i^{RP}; \tag{3.28}$$

$$y_i^{SP} = (\beta^{SP})' x_i^{SP} + \varepsilon_i^{SP}, \tag{3.29}$$

where the error terms are assumed to be i.i.d. normal distributed with zero means:

$$\varepsilon_i^{RP} \sim iid\, N(0,\, \sigma_{RP}^2);$$

$$\varepsilon_i^{SP} \sim iid\, N(0,\, \sigma_{SP}^2).$$

When these two models are considered separately, no assumption is made about the relationship between ε_i^{RP} and ε_i^{SP}. However, to link RP and SP model, it is more realistic to assume that ε_i^{RP} and ε_i^{SP} are correlated. Specifically, I assume that $(\varepsilon_i^{RP},\, \varepsilon_i^{SP})$ are drawn from an i.i.d bivariate normal distribution:

$$\begin{pmatrix} \varepsilon_i^{RP} \\ \varepsilon_i^{SP} \end{pmatrix} \sim \text{i.i.d. } N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix},\ \Sigma_\varepsilon \right),$$

where

$$\Sigma_\varepsilon \equiv \begin{pmatrix} \sigma^2_{RP} & \rho\sigma_{RP}\sigma_{SP} \\ \rho\sigma_{RP}\sigma_{SP} & \sigma^2_{SP} \end{pmatrix}.$$

To relate the RP and SP coefficients, I use a model structure analogous to the one suggested by Azevedo [5], defining $\beta^{RP} \equiv \beta$ and $\beta^{SP} \equiv \beta + \delta$; i.e.,

$$y_i^{RP} = \beta' x_i^{RP} + \varepsilon_i^{RP}, \tag{3.30}$$

$$y_i^{SP} = (\beta + \delta)' x_i^{SP} + \varepsilon_i^{SP}. \tag{3.31}$$

Here δ denotes the difference between β^{RP} and β^{SP}, measuring the discrepancy between RP and SP models. The restriction $\delta = 0_{K\times 1}$ corresponds to complete consistency between these two models, where K is the number of elements in the unknown parameter vector β. This specification provides an easy way to model and incorporate the prior belief about the parameters in the two models.

Note that in a classical framework, both this model structure (3.30-3.31) and the one in Azevedo [5] (3.24-3.25) are equivalent re-parameterizations of the same basic model (3.28-3.29). These models will result in the same set of maximum likelihood estimates when the parameters are treated as fixed. However, the above statement will not hold in a Bayesian framework where the parameters are assumed to be random instead of fixed. An additive, instead of a multiplicative, structure is used here. Because the posterior distributions of the slope parameters in a linear regression are generally normal and linear combinations of normal distributions are still normal, an additive structure provides more reliable convergence properties when the Monte Carlo simulation is used to approximate the posterior distributions.

3.4.2 Alternative Priors

In our analysis, the priors on β and Σ_ε are of less concern than the discrepancy parameters δ, which is a key parameter to link RP and SP models. Throughout this chapter, it is assumed that the prior on β is diffuse. Similarly, a conditionally conjugate prior is used for Σ_ε:

$$\Sigma_\varepsilon \sim \text{inv-W}((\rho_0 R_0)^{-1}, \rho_0), \tag{3.32}$$

with the values of ρ_0 and R_0 chosen to make this prior uninformative. I used $\rho_0 = 2$ and $R_0 = \text{diag}(0.001, 0.001)$.

As noted above, prior assumption regarding the δ is a key to linking the preference structures revealed by RP and SP data. By specifying different priors on δ, one can incorporate into the model various beliefs about the relationship between RP and SP data. In practice, a normal prior is used to for this purpose:

$$\delta \sim N(\mu_\delta^0, V_0), \tag{3.33}$$

By specifying the values of the location parameter μ_δ^0 and scale parameter V_0, various scenarios can be investigated.

In classical framework, the general consistency between RP and SP requires δ to be an vector of zeros and $\sigma_{RP}^2 = \sigma_{SP}^2$. The scale effects relax the assumption that the RP and SP survey responses stem from the exact same data generating process. In line with Morikawa [39], Adamowicz *et al.* [1] and Cameron *et al.* [15], I concentrate on the consistency on the coefficients rather than on the variance parameter here. This consistency can be expressed as the following hypothesis:

$$H_0 \ : \ \delta_{\beta_1} = \delta_{\beta_2} = ... = \delta_{\beta_K} = 0\ ; \tag{3.34}$$

$$H_1 \ : \ \text{at least one of above equation does not hold.}$$

To investigate this consistency, one can start with diffuse priors for δ and compute the posterior distributions of these parameters. All the posterior distributions being centered tightly on zero will suggest support to the hypothesis of general consistency.

On the other hand, by assigning different values to the prior of δ, we can incorporate alternative prior belief about the degree of agreement between the coefficient estimates from RP data and from SP data. For example, one can begin by assuming a high degree of consistency between RP and SP data. At an extreme, one can assume complete consistency *a priori* using $\mu_0 = 0$ and $V_0 = 0$. Alternatively, it is commonly believed that the cost coefficient estimated from RP data are more reliable than from SP data. In other words, the estimates of the price coefficient from these two data sets may not agree with each other. This corresponds to the situation where we have a diffuse prior on δ_{cost} and tight priors on other elements in δ.

3.4.3 Model comparison and Bayes Factor

The standard Bayesian approach to model comparison is to calculate the Bayes factor; i.e., the posterior odds of one model versus another when the prior probabilities of the two models are equal. Bayes factor was the centerpiece of the methodology developed by Jeffreys [27] for quantifying the evidence in favor of a scientific theory and has received considerable attention since. Kass and Raftery [29] give a comprehensive review for the Bayes factor.

Following Kass and Raftery [29], suppose under two models H_0 and H_1 the likelihood functions for the observed data D are $p(D|H_0)$ and $p(D|H_1)$ with priors $p(H_0)$ and $p(H_1) = 1 - p(H_0)$, according to the Bayes's rule, we have

$$p(H_j|D) = \frac{p(D|H_j)p(H_j)}{p(D|H_0)p(H_0) + p(D|H_1)p(H_1)}, \qquad (j = 0, 1);$$

so that the posterior odds of H_1 against H_0 is

$$\frac{p(H_1|D)}{p(H_0|D)} = \frac{p(D|H_1)p(H_1)}{p(D|H_0)p(H_0)},$$

where the $\frac{p(D|H_1)}{p(D|H_0)} \equiv B_{10}$ is the Bayes factor and $\frac{p(H_1)}{p(H_0)}$ is the prior odds. In words,

posterior odds = Bayes factor × prior odds.

When the models H_0 and H_1 are equally likely (i.e., $p(H_0) = p(H_1) = 0.5$) the Bayes factor B_{10} is equivalent to the posterior odds in favor of H_1 or against H_0. To interpret the Bayes factor, Jeffreys [27] suggests classifying B_{10} in half-units on the log_{10} scale, which is summarized by Kass and Raftery [29] in the following table. A measure on the log_e scale is also added for convenience here:

$log_{10}(B_{10})$	$log_e(B_{10})$	B_{10}	Evidence against H_0
$0 - 1/2$	$0 - 1.15$	$1 - 3.2$	Not worth more than a bare mention
$1/2 - 1$	$1.15 - 2.30$	$3.2 - 10$	Substantial
$1 - 2$	$2.30 - 4.61$	$10 - 100$	Strong
> 2	> 4.61	> 100	Decisive

To calculate the Bayes factor, one needs to calculate the marginal density $p(D|H_j)$. Suppressing the subscript j, we have

$$p(D|H) = \int p(D|\theta, H)p(\theta|H)d\theta \tag{3.35}$$

where θ is the parameter (vector) and $p(\theta|H)$ is its prior density under H, and $p(D|\theta, H)$ is the density of D given θ. In some simple models, it is possible to evaluate the integral (3.35) analytically. More often, this integral is intractable and must be calculated using numerical methods. A useful approximation to the marginal density (3.35) can be obtained through Laplace's method (Tierney and Kadane [48], Tierney *et al.* [49]):

$$p(D|H) \approx (2\pi)^{d/2}|\tilde{\Sigma}|^{1/2}p(D|\tilde{\theta}, H)p(\tilde{\theta}|H), \tag{3.36}$$

where d is the dimension of θ, $\tilde{\theta}$ is the posterior mode and $\tilde{\Sigma}$ is the inverse negative Hessian of $log(p(D|\theta, H)p(\theta|H))$ evaluated at the mode $\tilde{\theta}$. The relative error of this approximation as well as the resulting approximation to B_{10} is of order $O(n^{-1})$, where n is the sample size. As noted by Kass and Raftery [29] (p. 778): "In general, the method provides adequate approximations in well-behaved problems (those in which the likelihood functions are not grossly nonnormal) of modest dimensionality," They also suggests "...that samples of size $20d$ are large enough for the method to work well in most familiar problems provided that a reasonably good parameterization is used." In our application, the sample sizes are large enough and Laplace's approximation is used as a primary tool to calculate the Bayes factor.

3.4.4 Gibbs Sampling and Data Augmentation

As in Chapter 2, the joint posterior distribution for current model is complicated and numerical simulation methods are required to approximate this distribution. Again, the Gibbs sampling method is used for that purpose. The Gibbs sampling is straightforward in our application: all of the conditional distributions are of standard form, drawing from which is simple.

Let $y_i \equiv \begin{pmatrix} y^{RP} \\ y^{SP} \end{pmatrix}$, $x_i \equiv \begin{pmatrix} x_i^{RP} & 0_{n\times 1} \\ 0_{n\times 1} & x_i^{SP} \end{pmatrix}$, $\phi \equiv \begin{pmatrix} \beta \\ \beta+\delta \end{pmatrix}$ and $y = (y_1', ..., y_n')$.

The data likelihood can be expressed as the product of bivariate normal densities:

$$p(y|\beta, \delta, \Sigma_\varepsilon) \propto \prod_{i=1}^{n} N(y_i|x_i\phi, \Sigma_\varepsilon).$$

The joint posterior distribution for the parameters (β, δ, and Σ_ε) given the data is then

$$\begin{aligned} p(\beta, \delta, \Sigma_\varepsilon|y) &\propto p(\beta, \delta, \Sigma_\varepsilon)p(y|\beta, \delta, \Sigma_\varepsilon) \\ &\propto p(\beta, \delta, \Sigma_\varepsilon)\prod_{i=1}^{n} N(y_i|(x_i\phi, \Sigma_\varepsilon). \end{aligned} \tag{3.37}$$

The necessary conditional distributions used for Gibbs sampling are not difficult to derive. First, consider the conditional density of Σ_ε. With the prior (3.32), $\Sigma_\varepsilon|y, \beta, \delta$ follows a inverse Wishart distribution with degree of freedom $n + \rho_0$ and scale parameter $S = \sum_i (y_i - x_i\phi)'(y_i - x_i\phi) + \rho_0 R_0$:

$$\Sigma_\varepsilon|y, \beta, \delta \sim \text{inv-W}(n + \rho_0, S^{-1}). \tag{3.38}$$

To derive the conditional distributions of β and δ, we can factor the bivariate normal distribution of y_i as the product of the marginal distribution of y_i^{RP} and the conditional distribution of y_i^{SP} (conditional on y^{RP}) :

$$p(y_i|\beta, \delta, \Sigma_\varepsilon) \propto p(y_i^{RP}|\beta, \delta, \Sigma_\varepsilon)p(y_i^{SP}|\beta, \delta, \Sigma_\varepsilon, y_i^{RP}). \tag{3.39}$$

The joint posterior distribution is then

$$
\begin{aligned}
p(\beta,\delta,\Sigma_\varepsilon|y) &\propto p(\beta,\delta,\Sigma_\varepsilon)\prod_{i=1}^{n}\{p(y_i^{RP}|\beta,\delta,\Sigma_\varepsilon)p(y_i^{SP}|\beta,\delta,\Sigma_\varepsilon,y_i^{RP})\} \qquad (3.40)\\
&\propto p(\beta,\delta,\Sigma_\varepsilon)\prod_{i=1}^{n}\left\{N(y_i^{RP}|\beta' x_i^{RP},\sigma_{RP}^2)N(y_i^{SP}|\mu_{y_i^{SP}|y_i^{RP}},\sigma_{SP}^2(1-\rho^2))\right\};
\end{aligned}
$$

where

$$\mu_{y_i^{SP}|y_i^{RP}} = (\beta+\delta)' x_i^{SP} + \rho\frac{\sigma_{SP}}{\sigma_{RP}}(y_i^{RP} - \beta' x_i^{RP}).$$

With some algebra, the conditional distribution of β can be recognized as a multivariate normal distribution when a diffuse prior is used; i.e.,

$$p(\beta|y,\delta,\Sigma_\varepsilon) \sim N(\hat{\beta},\ \hat{V}); \qquad (3.41)$$

where

$$
\begin{aligned}
\hat{\beta} &= (X^{*\prime}X^*)^{-1}X^{*\prime}Y^*, \qquad (3.42)\\
\hat{V} &= (X^{*\prime}X^*)^{-1}, \qquad (3.43)
\end{aligned}
$$

with $X^* \equiv (x_1^{*\prime}, ..., x_{2n}^{*\prime})$ and $Y^* \equiv (y_1^*, ..., y_{2n}^*)$.

$$
y_i^* = \begin{cases} \dfrac{y_i^{RP}}{\sigma_{RP}} & \text{for } i = 1,2,...,n \\ \dfrac{y_{i-n}^{SP} - \delta' x_{i-n}^{SP} - \rho\frac{\sigma_{SP}}{\sigma_{RP}} y_{i-n}^{RP}}{\sigma_{SP}\sqrt{(1-\rho^2)}} & \text{for } i = n+1, n+2, ..., 2n \end{cases}
$$

$$
x_i^* = \begin{cases} \dfrac{x_i^{RP}}{\sigma_{RP}} & \text{for } i = 1,2,...,n \\ \dfrac{x_{i-n}^{SP} - \rho\frac{\sigma_{SP}}{\sigma_{RP}} x_{i-n}^{RP}}{\sigma_{SP}\sqrt{(1-\rho^2)}} & \text{for } i = n+1, n+2, ..., 2n. \end{cases}
$$

The conditional distribution of δ is normal as well. With a diffuse prior,

$$\delta|y,\beta,\Sigma_\varepsilon \sim N(\hat{\delta},\hat{V}_\delta), \qquad (3.44)$$

where

$$\hat{\delta} = (X^{SP\prime}X^{SP})^{-1}X^{SP\prime}Y^{**}, \tag{3.45}$$

$$\hat{V}_{\delta} = \sigma_{SP}^2(1-\rho^2)(X^{SP\prime}X^{SP})^{-1} \tag{3.46}$$

with $X^{SP} = (x_1^{SP}, ...x_n^{SP})$, $Y^{**} \equiv (y_1^{**}, ..., y_n^{**})$ and

$$y_i^{**} = y_i^{SP} - \beta' x_i^{SP} - \rho\frac{\sigma_{SP}}{\sigma_{RP}}(y_i^{RP} - \beta' x_i^{RP}) \qquad \text{for } i = 1, 2, ..., n.$$

For a specific normal prior of δ (3.33), the conditional posterior density of δ is still normal as distribution (3.44), with $\hat{\delta}$ being replaced with $(\hat{V}_{\delta}^{-1} + V_0^{-1})^{-1}(\hat{V}_{\delta}^{-1}\hat{\delta} + V_0^{-1}\mu_{\delta}^0)$ and $\hat{V}_{\delta}$ replaced with $(\hat{V}_{\delta}^{-1} + V_0^{-1})^{-1}$. These conditional distributions for β (3.41), δ (3.44) and Σ_{ε} (3.38) provide the basis of the Gibbs routine used in our analysis.

Turning to the data augmentation procedure, one must take into account the fact that a bivariate normal distribution, instead of independent normal distributions is assumed for the error terms in the RP and SP models. Utilizing the factorization of bivariate normal distribution (3.39), we introduce latent variables z_i^{RP1}, z_i^{SP1}, z_i^{RP2} and z_i^{SP2} and define a set of new observations:

$$y_i^{new} = \begin{cases} (y_i^{RP}, y_i^{SP}) & \text{if } y_i^{RP} > 0 \text{ and } y_i^{SP} > 0; \\ (z_i^{RP1}, y_i^{RP}) & \text{if } y_i^{RP} = 0 \text{ and } y_i^{SP} > 0; \\ (y_i^{RP}, z_i^{SP1}) & \text{if } y_i^{RP} > 0 \text{ and } y_i^{SP} = 0; \\ (z_i^{RP2}, z_i^{SP2}) & \text{if } y_i^{RP} = 0 \text{ and } y_i^{SP} = 0, \end{cases} \tag{3.47}$$

where

$z_i^{RP1} \sim N(y_i^{RP}|\mu_{y_i^{RP}|y_i^{SP}},\ \sigma_{RP}^2(1-\rho^2))$ truncated at the right by 0;

$z_i^{SP1} \sim N(y_i^{SP}|\mu_{y_i^{SP}|y_i^{RP}},\ \sigma_{SP}^2(1-\rho^2))$ truncated at the right by 0;

and

$$z_i^{RP2} \sim N(\beta' x_i^{RP}, \sigma_{RP}^2) \text{ truncated at the right by 0;}$$

$$z_i^{SP2} \sim N(y_i^{SP} | \mu_{y_i^{SP}|y_i^{RP}},\ \sigma_{SP}^2(1-\rho^2)) \text{ truncated at the right by 0}$$

with

$$\mu_{y_i^{RP}|y_i^{SP}} \equiv \beta' x_i^{RP} + \rho \frac{\sigma_{RP}}{\sigma_{SP}} (y_i^{SP} - (\beta+\delta)' x_i^{SP});$$

and

$$\mu_{y_i^{SP}|y_i^{RP}} \equiv (\beta+\delta)' x_i^{SP} + \rho \frac{\sigma_{SP}}{\sigma_{RP}} (y_i^{RP} - \beta' x_i^{RP}).$$

By construct, this set of new observations $\{y_i^{new}\}$ is bivariate normal and is used in place of $\{y_i\}$ in the above Gibbs sampling procedure.

3.5 RESULTS

The estimation results from the bivariate Tobit regression are shown in Table 3.2. In contrast to the standard Tobit estimation results for zone-level data in Chapter 2, most of the point estimates are significant. This is not surprising since more aggregated data with larger sample sizes are used here. For both RP and SP data, the resulting signs of the significant coefficients have the expected signs, with the number of trips increasing with income and decreasing with both direct travel cost and the opportunity cost of travel time. As found by Azevedo [5], there are noticeable and statistically significant differences between the parameter estimates for the RP and SP data. The SP price coefficient in particular is substantially smaller in absolute value ($\delta > 0$), where as the SP income coefficient is larger. Finally, the estimates of the correlation parameters range from 0.62 to 0.69, showing a strong correlation between RP and SP data.

Table 3.3 shows the posterior means and standard deviations from the Bayesian analysis when diffuse priors are used for β and Σ_ε, and a weak normal prior (3.33) is used for the discrepancy parameter vector δ with mean $\mu_\delta^0 = 0_{4\times1}$ and variance $V_0 = 100^2 \times \text{diag}(\hat{\sigma}_{\beta_1}^2, ..., \hat{\sigma}_{\beta_4}^2)$, where $\hat{\sigma}_{\beta_i}$'s are the standard errors for the estimates of β_i's in the bivariate Tobit regressions. These results are similar to their classical counterparts (the posterior means versus the point estimates and the posterior standard deviations versus the standard errors) in Table 3.2, which is expected given the use of non-informative priors. The posterior density plots for the discrepancy parameters are presented in Figures 3.1 – 3.5. These figures do not suggest support for the assumption that these parameters center on zeros with narrow spread.

While diffuse priors on δ are convenient, many practicers come to the task of combining RP and SP data with more informative prior beliefs regarding their compatibility. We consider two broad classes of alternative priors:

- Class 1: $\delta|\lambda \sim N(0_{K\times1}, V_1(\lambda))$, where $V_1(\lambda) \equiv \lambda^2 \text{diag}(\hat{\sigma}_{\beta_1}^2, ..., \hat{\sigma}_{\beta_4}^2)$. This prior class allows for varying degrees of consistency between the RP and SP parameters. The distribution of δ is centered around zero, with prior uncertainty about consistency increasing as λ increases. At the extreme, complete consistency results if $\lambda = 0$.
- Class 2: $\delta|\lambda \sim N(0_{K\times1}, V_2(\lambda))$, where $V_2(\lambda) \equiv \text{diag}(\lambda^2\hat{\sigma}_{\beta_1}^2, \lambda^2\hat{\sigma}_{\beta_2}^2, 100^2\hat{\sigma}_{\beta_3}^2, 100^2\hat{\sigma}_{\beta_4}^2)$. This prior corresponds to the belief that SP respondents do not consider price carefully in responding to hypothetical cost changes, but do provide information of the impact of other factors.

Table 3.4 provides results given a class 1 prior with $\lambda = 0.01$. As we can see, such tight a prior results in almost zero posterior means and small standard deviations for δ. In contrast, Table 3.5 shows the Bayesian analysis results given a class 2 prior with $\lambda = 0.01$. In this case, while the RP and SP constant and income coefficients are tightly tied together *a priori*, the cost and opportunity cost coefficients are loosely centered. As a result, differences do emerge, but greater similarity exists than when completely diffuse priors are used for δ.

From the point of view of a neutral researcher, it may be interesting to compare these competing models with different priors and to see how they are supported by the data. Since the logarithm of the Bayes factor is just the difference between logarithm of the corresponding marginal likelihoods; i.e., $\log(B_{10}) = \log(p(D|H_1)) - \log(p(D|H_0))$, we construct the plots of log marginal likelihood to access Bayes factor analysis for these two sets of models. Specifically, we consider models with λ taking a sequence of the values from 0.2 to 20 with increment of 0.1. Utilizing the Maxlik routine in Gauss, one can conveniently calculate the posterior modes and the inverse negative Hessian matrices for these models. Then using the formula (3.36) one can approximate the marginal likelihoods of y. The results are summarized graphically with log marginal likelihood of y being plotted against λ: Figures 3.6 – 3.10 present the results for models with priors in class 1 and Figure 3.11 - 3.15 for models with priors in class 2 for the 5 megazones respectively.[8] As we can see in the figures, the models with λ close to zero have the smallest marginal likelihoods while models with λ in the range three to four have the largest marginal likelihoods. Thus, the data provide little support for either class of models relying on a high degree of consistency between the RP and SP data. These plots in Figures 3.6 through 3.15 can be used further to identify the range of values of

---

[8] Note that there are a few points where the maximization routine can not calculate the inverse of Hessians, thus the log likelihood values are omitted for those instances. Also, these curves are smoothed using a procedure (command "loess") provided in Gauss which is based on Cleveland, William S. "Robust Locally Weighted Regression and Smoothing Scatterplots." JASA . Vol. 74, 1979, 829-36.

λ supported by the data within each class. To see this more clearly, two lines are also drawn based on the classification for Bayes factor suggested by Jeffreys [27]:

line1: log marginal likelihood = the maximum log marginal likelihood - 2.30;

line2: log marginal likelihood = the maximum log marginal likelihood - 4.61.

Thus based on the position of a model's log marginal likelihood in these figures, one can determine how well the data support the model. If its log marginal likelihoods are below line 1 (line 2), then we conclude that the data provide strong (decisive) evidence against the model. Since the main issue here is the consistency between RP and SP data, we are more concerned about the answer to the following question: to what extend do the data support (or provide strong evidence against) a model with λ close to zero? The following tables show the cutoff values for λ with two classifications:

cutoff points of "*strong* evidence" for λ

MegaZone	1	2	3	4	5
Class 1	1.4	1.6	2.7	4.3	2.2
Class 2	1.1	1.2	2.9	3.6	1.8

cutoff points of "*Decisive* evidence" for λ

MegaZone	1	2	3	4	5
Class 1	1.1	1.0	2.2	3.5	1.7
Class 2	0.9	0.8	2.1	2.7	1.4

This suggests that the data do not support the models with priors on δ implying consistency with too narrow a spread for both sets of models. Take megazone 1, for example, if one specifies the model with prior $\delta \sim N(0_{K\times1}, V_1(1.5))$, the data do not provide strong evidence against this specification. However, this is not the case if one assume $\delta \sim N(0_{K\times1}, V_1(1.3))$. For all 5 megazones we can conclude that, any null hypothesis that implies the consistency with spread parameter $\lambda < 1.4$ for class 1 and $\lambda < 1.1$ for class 2 are not supported by the data.

3.6 CONCLUSION

This essay illustrates how a Bayesian framework can be used to link RP and SP data. A bivariate Tobit regression model is considered and the Gibbs sampling and data augmentation method is used to carry out the Bayesian analysis. By modeling with additive discrepancy parameters, one can easily incorporate their prior beliefs about the consistency between RP and SP data. The model structure also facilitates the comparison of models with different priors and Bayes factor analysis is used for that purpose. In the application to the Iowa wetlands survey data, two series of models relating to the consistency between RP and SP data are investigated. We reach the same conclusions as Azevedo [5] does in his research: the data do not support tight consistency between RP and SP coefficients. However, we are able to further identify the range of model structures within a class of models that are supported by the data. In the analysis, two classes of models are considered with class 1 focusing on general consistency and class 2 on relaxing the consistency on the cost and opportunity cost coefficients. Classes of models that focus on other coefficients such as income and age coefficients could also be investigated in a similar manner.

Table 3.1 Mean Statistics for Mega Zone Data
(Standard Deviation is shown in parenthesis)

MegaZone	1	2	3	4	5	Total
#Obs	214	273	364	1420	355	2626
#Trips Taken	6.21	8.20	7.84	7.23	7.71	7.40
	(9.87)	(10.99)	(11.10)	(10.36)	(10.62)	(10.53)
#Trips Planed to Take	1.86	2.66	2.50	2.11	2.36	2.23
	(4.93)	(6.33)	(6.96)	(5.77)	(6.97)	(6.12)
Out of Pocket Cost ($)	26.65	22.81	30.69	25.49	25.30	26.00
	(13.80)	(10.37)	(13.72)	(9.87)	(14.76)	(11.79)
Out of Pocket Cost if Travel Cost Increased ($)	54.85	48.21	59.66	50.83	55.90	52.79
	(19.68)	(16.96)	(20.89)	(18.10)	(20.64)	(19.20)
Travel Time (hours)	1.93	1.80	2.42	1.80	2.23	1.96
	(0.37)	(0.33)	(0.48)	(0.33)	(0.56)	(0.46)
Income ($1000)	44.52	41.03	37.73	44.78	45.22	43.45
	(31.79)	(25.45)	(24.81)	(29.92)	(30.05)	(29.09)
License[1]	0.73	0.77	0.79	0.68	0.69	0.71
	(0.45)	(0.42)	(0.41)	(0.47)	(0.46)	(0.45)
Gender[2]	0.74	0.75	0.75	0.76	0.77	0.76
	(0.44)	(0.44)	(0.43)	(0.43)	(0.42)	(0.43)
Age (years)	49.67	48.28	46.85	48.35	48.08	48.21
	(16.16)	(15.61)	(15.05)	(15.97)	(15.42)	(15.75)

[1] License=1 if individual owns a hunting or fishing license, =0 otherwise;
[2] Gender=1 if respondent is male, =0 if female.

Table 3.2 Bivariate Tobit Estimation
(Standard Error is shown in parenthesis)

MegaZone	**1**	**2**	**3**	**4**	**5**
β_{Constant}	16.950**	25.288**	26.090**	19.572**	19.862**
	(2.735)	(3.315)	(2.416)	(1.329)	(2.193)
β_{Income}	0.028	0.032	0.052	0.074**	0.048
	(0.045)	(0.064)	(0.039)	(0.019)	(0.036)
β_{Cost}	-0.705**	-1.051**	-0.803**	-0.728**	-0.734**
	(0.091)	(0.151)	(0.079)	(0.049)	(0.088)
$\beta_{\text{Opp.Cost}}$	0.010	0.007	-0.020	-0.040**	-0.020
	(0.028)	(0.062)	(0.027)	(0.015)	(0.031)
δ_{Contant}	-8.801*	-13.432**	-3.178	-8.524**	-4.305
	(3.482)	(4.039)	(3.096)	(1.714)	(3.217)
δ_{Income}	0.115*	0.123	0.083	0.108**	0.077
	(0.047)	(0.068)	(0.051)	(0.023)	(0.039)
δ_{Cost}	0.360**	0.650**	0.300**	0.291**	0.283**
	(0.099)	(0.156)	(0.088)	(0.054)	(0.097)
$\delta_{\text{Opp.Cost}}$	-0.091**	-0.200**	-0.188**	-0.137**	-0.127**
	(0.028)	(0.069)	(0.047)	(0.021)	(0.037)
σ_{RP}	12.806**	13.656**	13.336**	13.712**	12.470**
	(1.071)	(1.058)	(1.048)	(1.026)	(1.049)
σ_{SP}	10.244**	13.598**	12.954**	13.061**	13.754**
	(1.107)	(1.089)	(1.072)	(1.042)	(1.077)
ρ	0.646**	0.642**	0.695**	0.621**	0.654**
	(0.076)	(0.059)	(0.048)	(0.030)	(0.056)
Log likelihood	-729.25	-1089.99	-1376.65	-5297.38	-1340.88

** Denotes significance at $\alpha = 1\%$ level
* Denotes significance at $\alpha = 5\%$ level

Table 3.3 Bayesian Analysis with Diffuse Priors
-- Posterior Mean and Standard Deviation (in parenthesis)

MegaZone	1	2	3	4	5
β_{Constant}	16.798	24.077	25.759	19.271	19.442
	(2.688)	(2.989)	(2.384)	(1.289)	(2.169)
β_{Income}	0.029	0.047	0.052	0.073	0.049
	(0.044)	(0.052)	(0.039)	(0.019)	(0.036)
β_{Cost}	-0.681	-0.974	-0.776	-0.701	-0.693
	(0.087)	(0.127)	(0.076)	(0.046)	(0.083)
$\beta_{\text{Opp.Cost}}$	0.007	-0.010	-0.025	-0.041	-0.025
	(0.027)	(0.047)	(0.027)	(0.015)	(0.031)
δ_{Contant}	-8.565	-11.989	-2.466	-8.216	-3.816
	(3.618)	(3.994)	(3.220)	(1.702)	(3.302)
δ_{Income}	0.122	0.113	0.087	0.110	0.082
	(0.049)	(0.061)	(0.053)	(0.023)	(0.041)
δ_{Cost}	0.317	0.558	0.262	0.262	0.232
	(0.099)	(0.139)	(0.087)	(0.052)	(0.093)
$\delta_{\text{Opp.Cost}}$	-0.094	-0.195	-0.193	-0.138	-0.130
	(0.029)	(0.061)	(0.049)	(0.021)	(0.038)
σ_{RP}	12.722	13.527	13.303	13.495	12.396
	(0.862)	(0.746)	(0.624)	(0.335)	(0.594)
σ_{SP}	10.994	14.297	13.414	13.159	14.271
	(1.227)	(1.293)	(0.986)	(0.537)	(1.107)
ρ	0.632	0.633	0.686	0.614	0.646
	(0.066)	(0.050)	(0.042)	(0.025)	(0.048)

Table 3.4 Bayesian Analysis with Tight Prior on δ
-- Posterior Mean and Standard Deviation (in parenthesis)

MegaZone	1	2	3	4	5
β_{Constant}	10.685	16.712	22.426	15.962	15.746
	(2.292)	(2.188)	(1.850)	(0.981)	(1.686)
β_{Income}	0.115	0.146	0.102	0.126	0.112
	(0.039)	(0.043)	(0.036)	(0.016)	(0.031)
β_{Cost}	-0.431	-0.582	-0.595	-0.563	-0.493
	(0.056)	(0.061)	(0.049)	(0.027)	(0.046)
$\beta_{\text{Opp.Cost}}$	-0.064	-0.131	-0.095	-0.097	-0.102
	(0.022)	(0.034)	(0.025)	(0.012)	(0.025)
δ_{Contant}	0.001	-0.001	0.000	-0.001	0.000
	(0.027)	(0.031)	(0.024)	(0.014)	(0.022)
δ_{Income}	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
δ_{Cost}	0.000	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\delta_{\text{Opp.Cost}}$	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
σ_{RP}	12.977	13.801	13.291	13.732	12.730
	(0.909)	(0.782)	(0.619)	(0.345)	(0.624)
σ_{SP}	11.843	14.632	13.938	12.945	13.737
	(1.228)	(1.223)	(0.958)	(0.484)	(0.918)
ρ	0.664	0.651	0.683	0.626	0.668
	(0.061)	(0.048)	(0.041)	(0.024)	(0.046)

Table 3.5 Bayesian Analysis with Diffuse Prior on δ_{Cost} and $\delta_{\text{Opp.Cost}}$, and Tight Priors on other δ's -- Posterior Mean and Standard Deviation (in parenthesis)

MegaZone	1	2	3	4	5
β_{Constant}	13.121	20.160	24.665	15.841	17.862
	(2.468)	(2.694)	(2.182)	(1.154)	(2.027)
β_{Income}	0.089	0.098	0.077	0.121	0.082
	(0.039)	(0.046)	(0.037)	(0.017)	(0.033)
β_{Cost}	-0.583	-0.808	-0.742	-0.588	-0.630
	(0.079)	(0.111)	(0.070)	(0.041)	(0.078)
$\beta_{\text{Opp.Cost}}$	-0.024	-0.052	-0.041	-0.073	-0.054
	(0.025)	(0.041)	(0.026)	(0.014)	(0.029)
δ_{Contant}	0.000	-0.001	0.000	-0.001	0.000
	(0.027)	(0.031)	(0.024)	(0.014)	(0.022)
δ_{Income}	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
δ_{Cost}	0.153	0.263	0.215	0.078	0.156
	(0.059)	(0.090)	(0.051)	(0.030)	(0.060)
$\delta_{\text{Opp.Cost}}$	-0.044	-0.159	-0.137	-0.080	-0.078
	(0.019)	(0.049)	(0.032)	(0.014)	(0.027)
σ_{RP}	12.902	13.553	13.302	13.561	12.516
	(0.888)	(0.761)	(0.621)	(0.338)	(0.610)
σ_{SP}	11.189	14.928	13.213	13.204	13.824
	(1.176)	(1.316)	(0.904)	(0.522)	(1.000)
ρ	0.654	0.640	0.692	0.622	0.659
	(0.062)	(0.049)	(0.041)	(0.024)	(0.047)

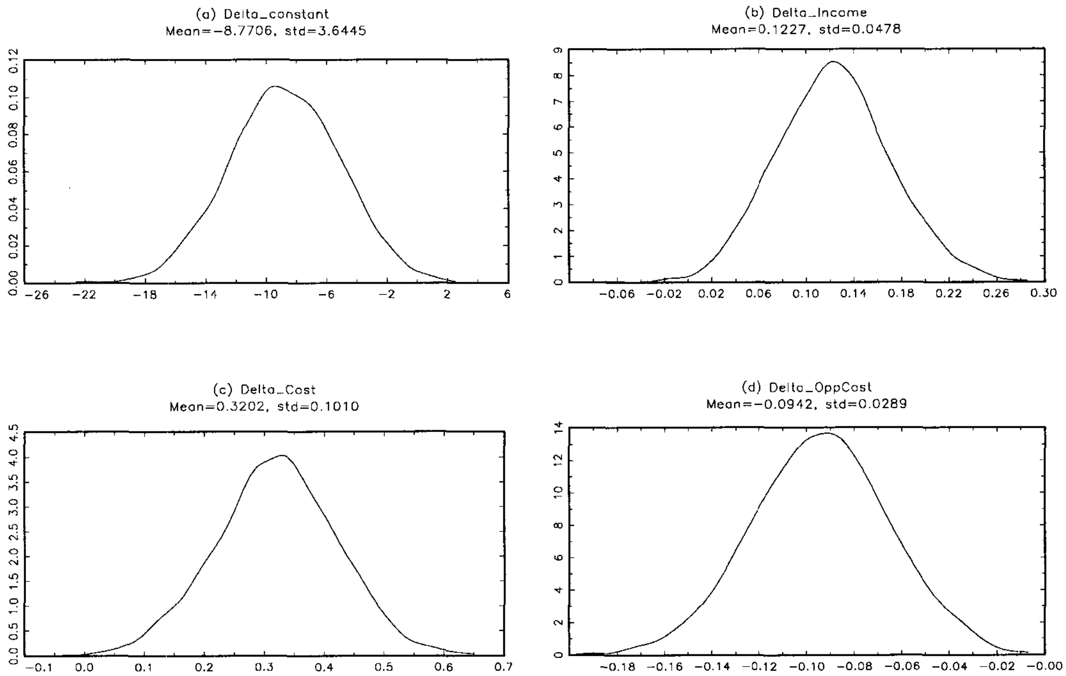


Figure 3.1. Kernel Density Plots for Discrepancy Parameters – MegaZone 1

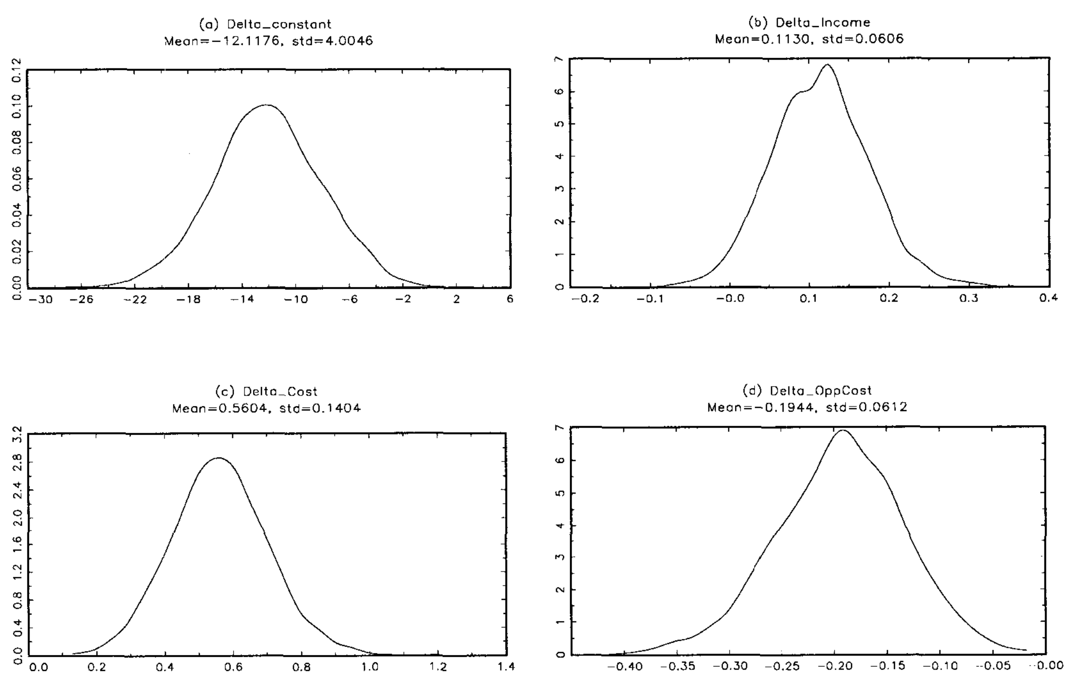


Figure 3.2. Kernel Density Plots for Discrepancy Parameters – MegaZone 2

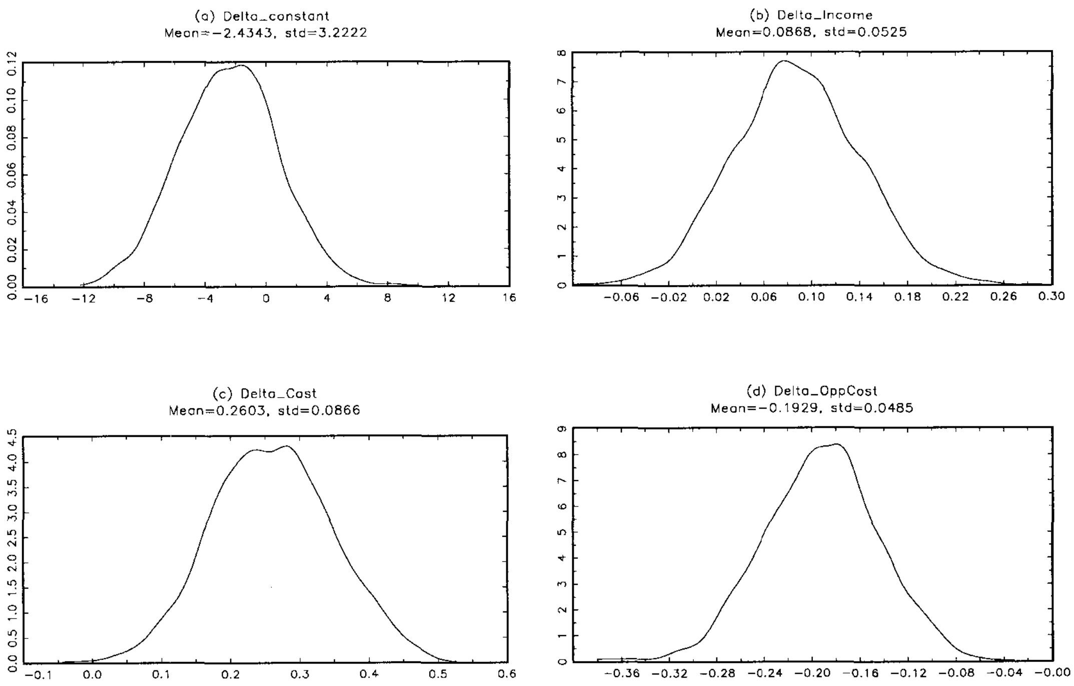


Figure 3.3. Kernel Density Plots for Discrepancy Parameters – MegaZone 3

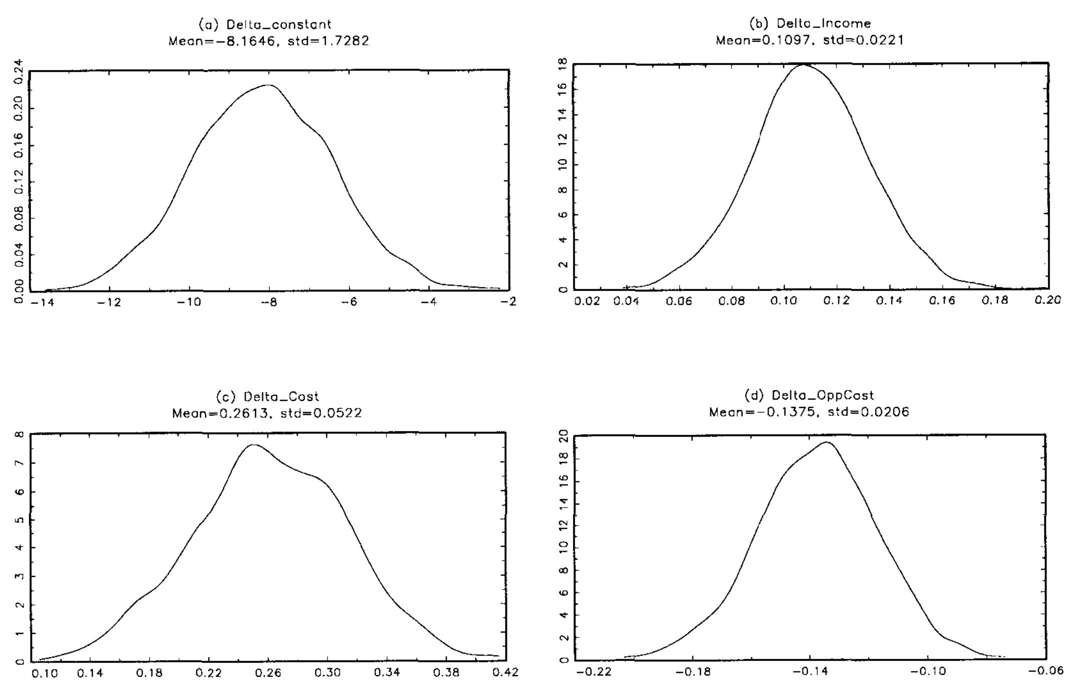


Figure 3.4. Kernel Density Plots for Discrepancy Parameters – MegaZone 4

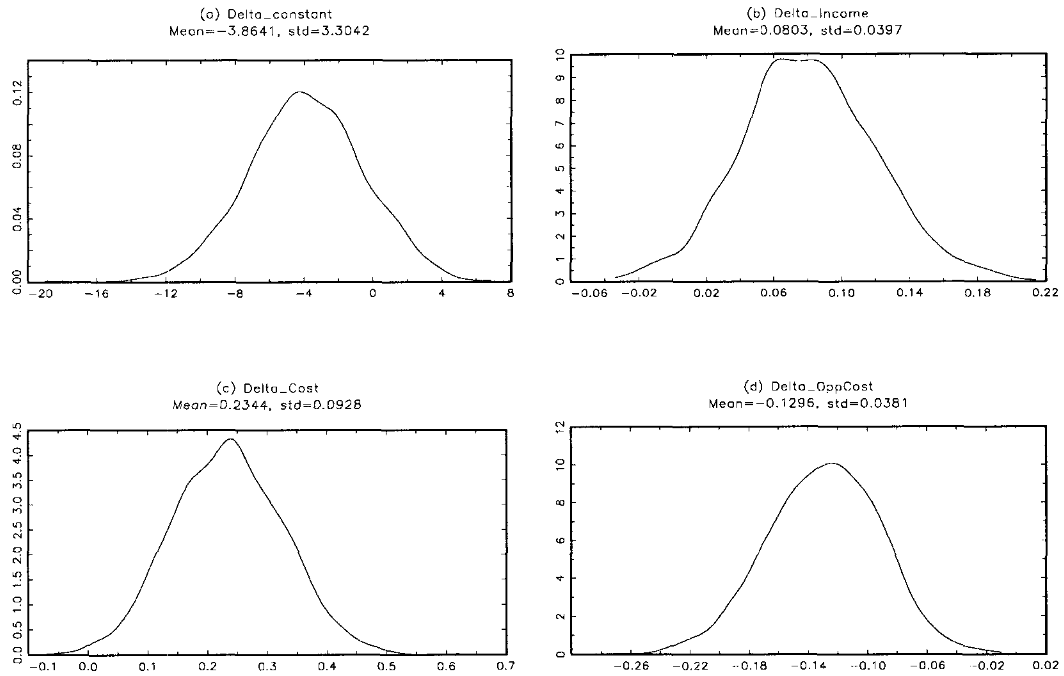


Figure 3.5. Kernel Density Plots for Discrepancy Parameters – MegaZone 5

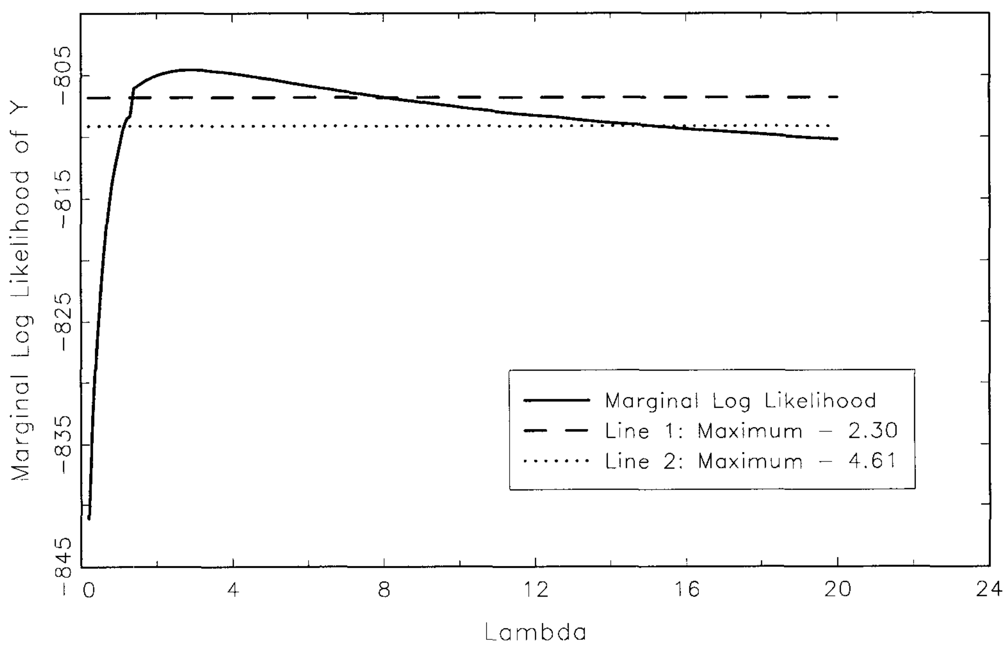


Figure 3.6. Log Marginal Likelihood Plot for Class 1 – MegaZone 1

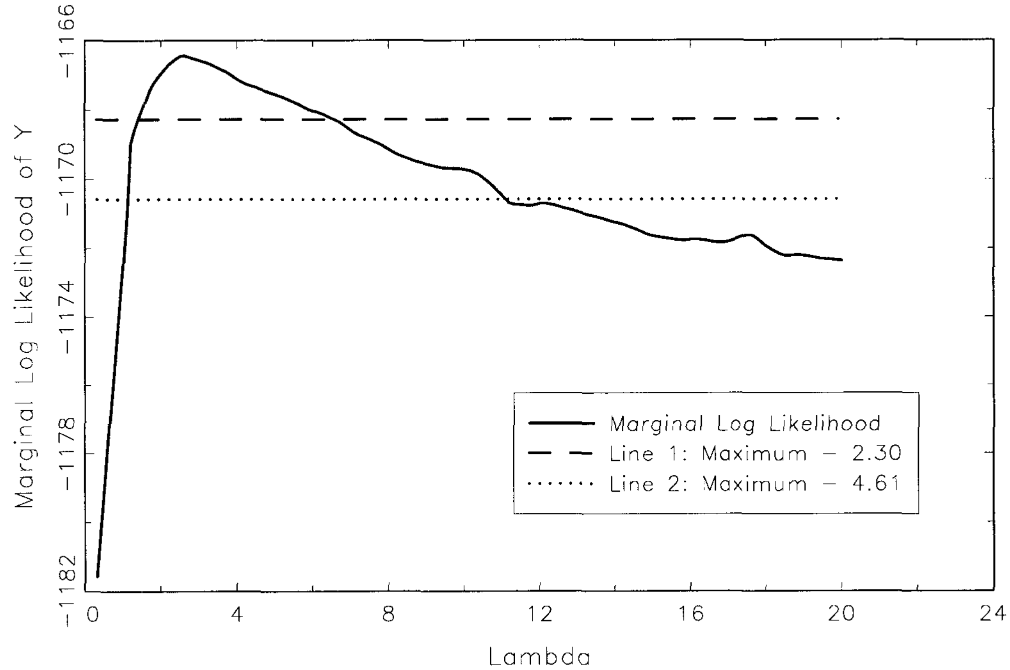


Figure 3.7. Log Marginal Likelihood Plot for Class 1 – MegaZone 2

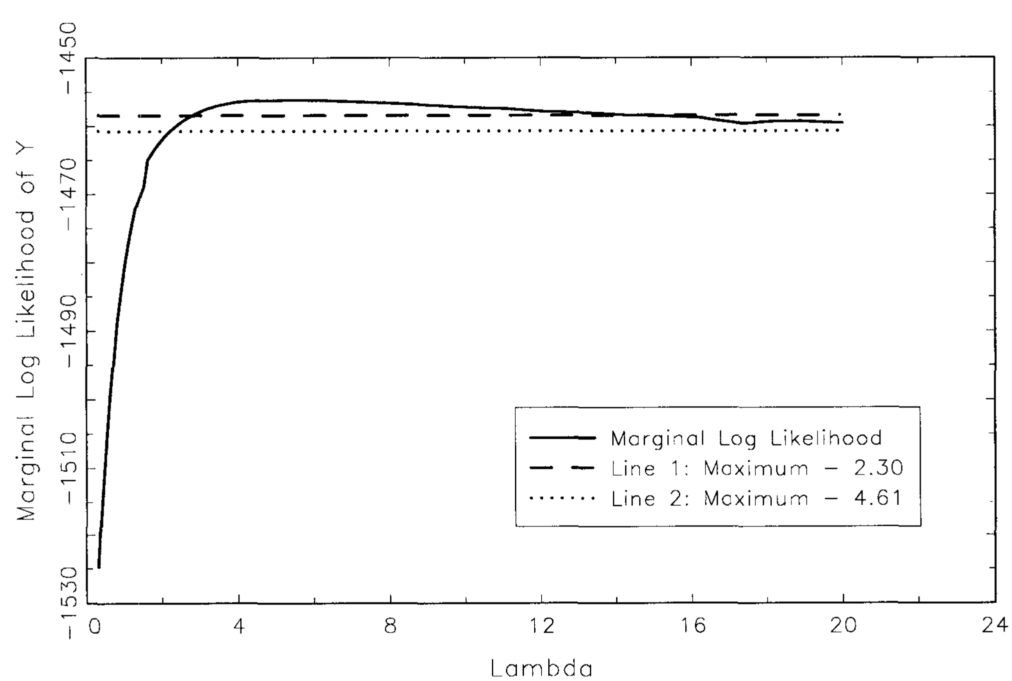


Figure 3.8. Log Marginal Likelihood Plot for Class 1 – MegaZone 3

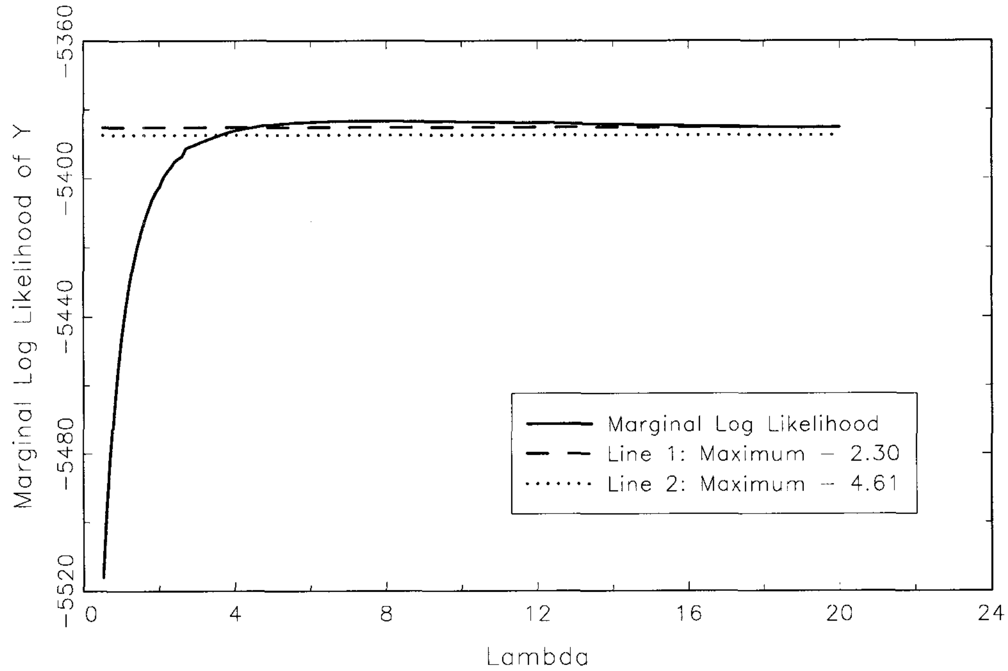


Figure 3.9. Log Marginal Likelihood Plot for Class 1 – MegaZone 4

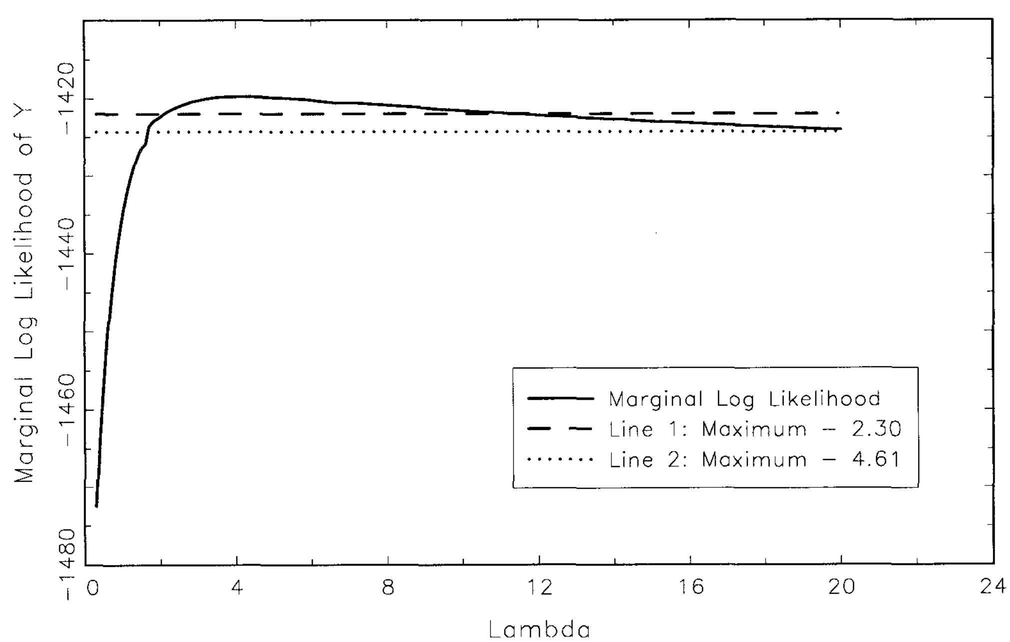


Figure 3.10. Log Marginal Likelihood Plot for Class 1 – MegaZone 5

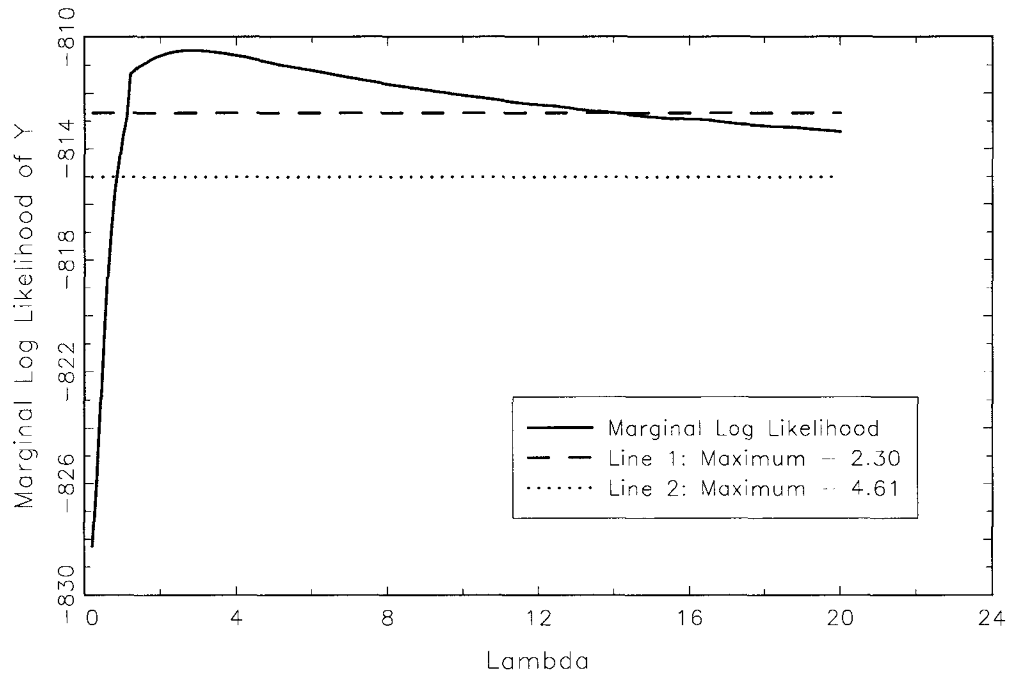


Figure 3.11. Log Marginal Likelihood Plot for Class 2 – MegaZone 1

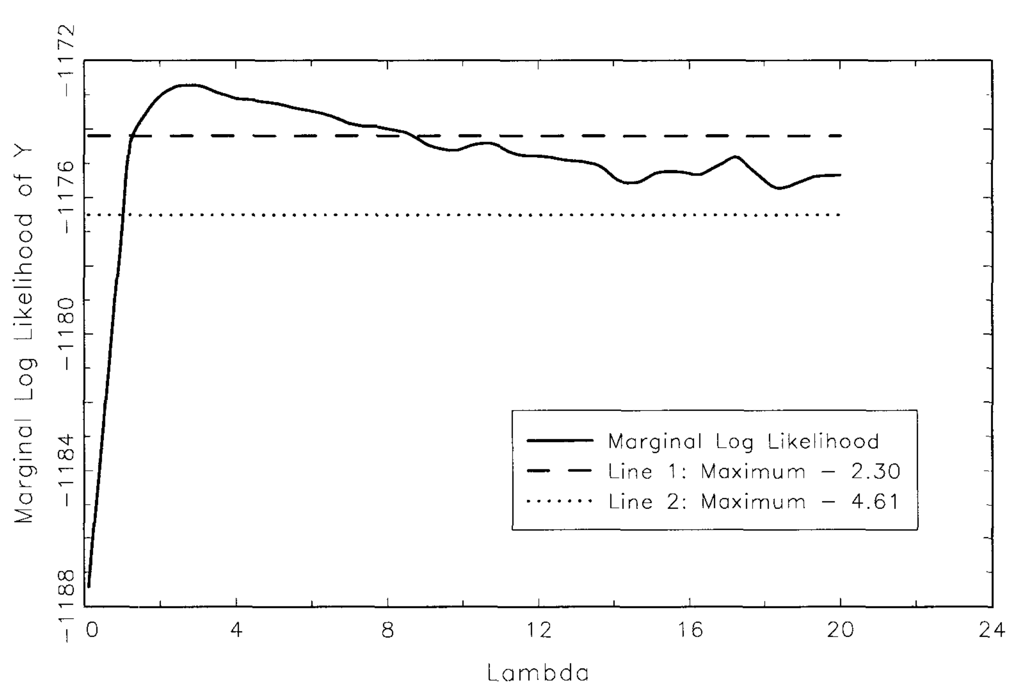


Figure 3.12. Log Marginal Likelihood Plot for Class 2 – MegaZone 2

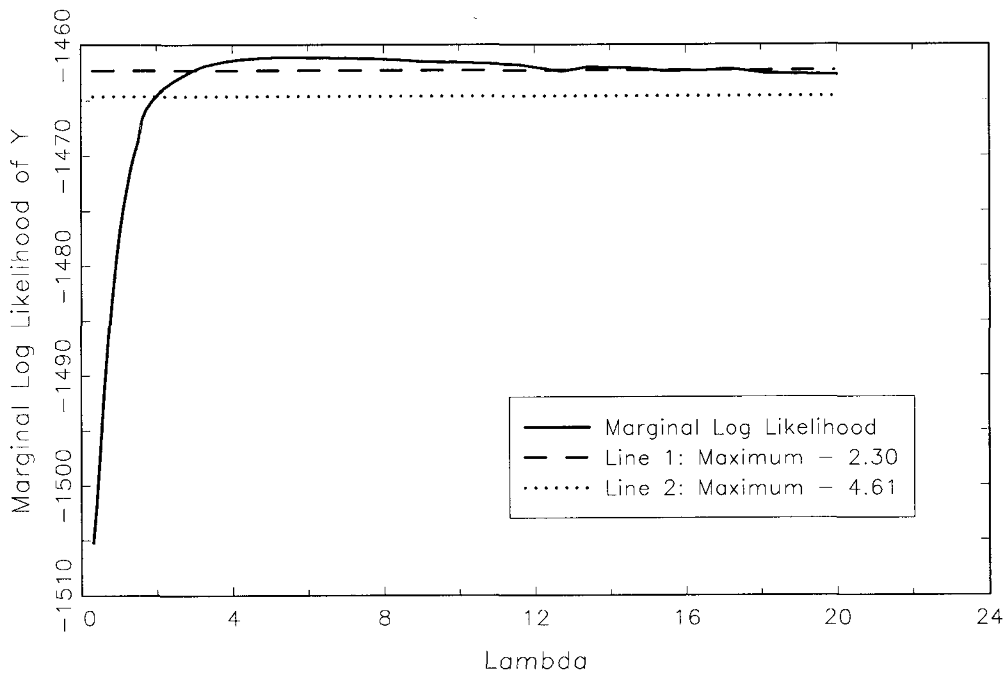


Figure 3.13. Log Marginal Likelihood Plot for Class 2 – MegaZone 3

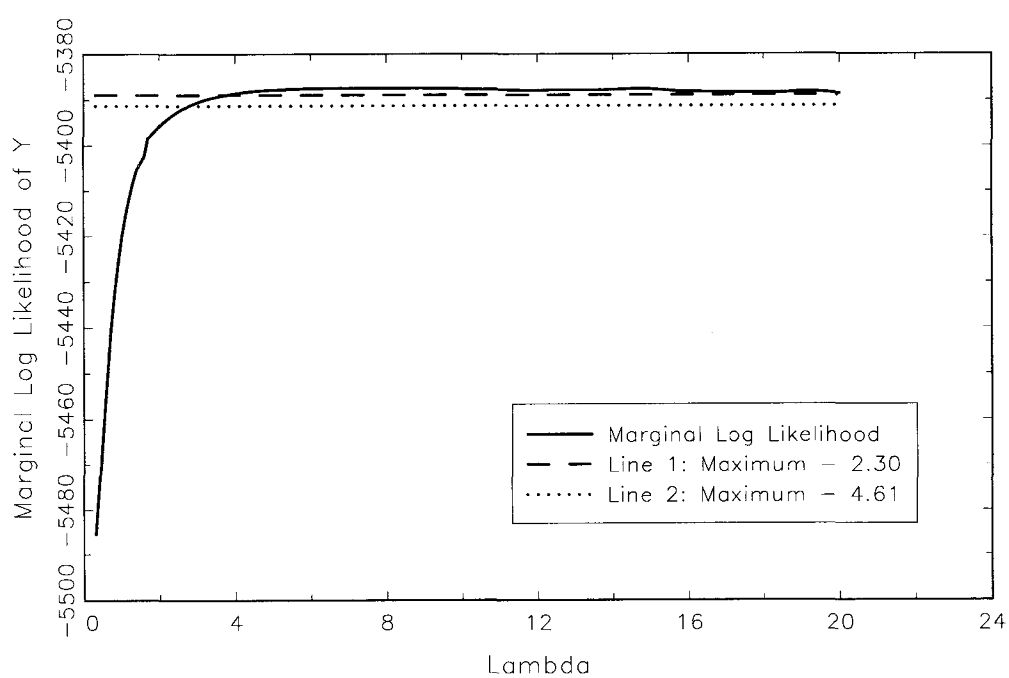


Figure 3.14. Log Marginal Likelihood Plot for Class 2 – MegaZone 4

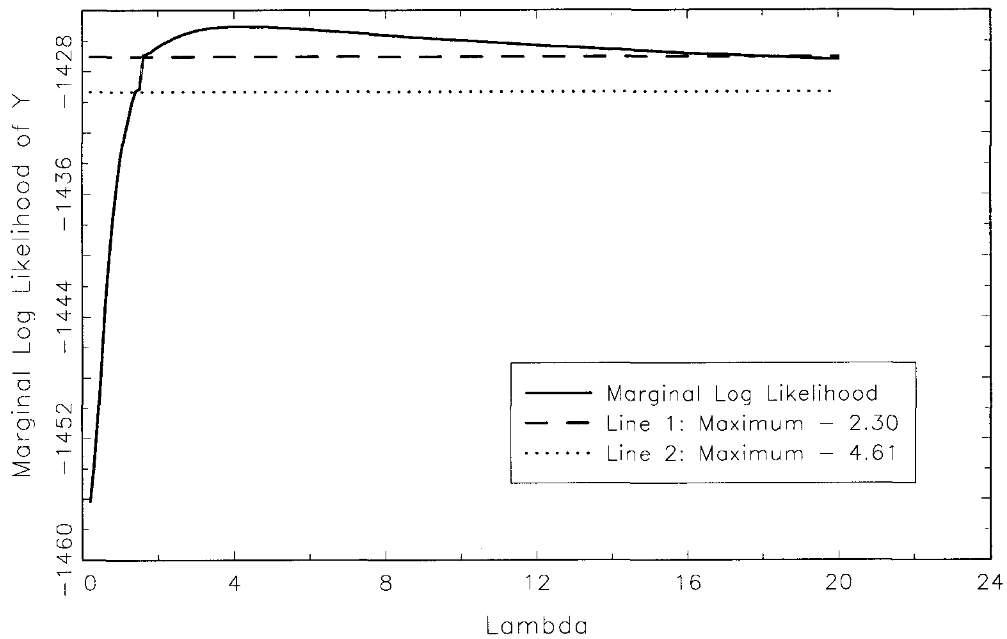


Figure 3.15. Log Marginal Likelihood Plot for Class 2 – MegaZone 5

Chapter 4
Mixed Logit Models for RP and SP

4.1 INTRODUCTION

In my third essay I develop a mixed logit model as an alternative approach to combining revealed and stated preference data. As noted above, most efforts to date at combining stated and revealed data sources have focused on testing the consistency in the underlying preferences. Indeed, most studies rely upon a simple test of consistency, (i.e., H_0: consistent and H_A: inconsistent), with little attention paid to the form or sources of inconsistency. Recently, Azevedo, Herriges and Kling [6] have attempted to isolate the sources of the inconsistency (e.g., respondents ignoring their budget constraint or measurement errors in RP prices). In this essay, we take the analysis further by trying to explain both the sources of the discrepancy between stated and revealed preferences and to capture heterogeneity in the degree of consistency among individuals. Specifically, I propose to model both SP and RP data from the wetland data described in Chapter 3 using a mixed logit model of the demand for wetland visits. In this framework, the discrepancies between SP and RP responses can be modeled as having a distribution in the population; a discrepancy whose means and variances can also depend on observed attributes of the survey respondents. Understanding the sources of these discrepancies can help to better design SP surveys.

The remainder of this chapter is divided into three sections. Section 2 provides a brief overview of the logit model, its advantages and limitations. Section 3 then introduces the mixed logit (MXL) framework of McFadden and Train [35], describing how it relaxes the restrictive nature of either logit or nested logit models. The proposed MXL model, combining RP and SP data from the Iowa wetlands studies, is then outlined in section 4.

4.2 THE LOGIT SPECIFICATION

The logit model in the context of recreation demand typically begins by assuming that the utility that individual i receives from visiting site j on choice occasion t is given by

$$\begin{aligned} U_{ijt} &= V_{ijt} + \varepsilon_{ijt}, \\ \text{for } i &= 1, ..., N; \;\; j = 1, ..., J; \;\; t = 1, ...T, \end{aligned} \tag{4.1}$$

where V_{ijt} is the systematic portion of the utility and ε_{ijt} is the error term. Typically $j = 0$ is used to denote the "stay at home" option. Often called the representative utility function, V_{ijt} is specified by the researcher to capture the statistical relationship between the decision maker's utility and the observed factors (Train [53]) Assuming that the ε_{ijt} are iid extreme value variates yields the logit model for the probability of visiting site j on choice occasion t:

$$p_{ijt} = \frac{e^{V_{ijt}}}{\sum_k e^{V_{ikt}}}. \tag{4.2}$$

V_{ijt} is usually assumed to be linear in parameters; i.e.,

$$V_{ijt} = \beta' x_{ijt} + \varepsilon_{ijt},$$

where x_{ijt} is a vector of observed variables including individual and site specific attributes. With this specification, the logit probability becomes

$$p_{ijt} = \frac{e^{\beta' x_{ijt}}}{\sum_k e^{\beta' x_{ikt}}}. \tag{4.3}$$

Treating decisions on each choice occasion as independent, as is done in most applications, yields the repeated logit model of Morey, Rowe and Watson [38], with choice probabilities that are simply products of the respective choice occasion probabilities.

The nested logit model generalizes the standard logit specification by allowing for patterns of correlation among the choice alternatives on a given choice occasion. Rather than assuming that the ε_{ijt}'s are iid extreme value, the error vector $\varepsilon_{i.t} = (\varepsilon_{i1t}, ..., \varepsilon_{iJt})'$ is assumed to be drawn from a generalized extreme value distribution. Sites are grouped (or nested) with greater correlation among the errors in within the same nest than across different nests (see Herriges and Phaneuf [25]). However, nested logit requires the analyst to specify *a priori* the nesting structure. With repeated choices available to the consumer, a repeated nested logit model is constructed by again assuming that the decisions on different choice occasions are independent and using product of choice occasion specific probabilities to construct the contribution to the likelihood function.

Logit, including its generalization nested logit, remains the most widely used discrete choice model in recreation demand analysis. Its popularity is due to its many desirable properties, perhaps the most attractive of which is that the formulas for the choice probabilities have simple closed forms. This convenience, however, is offset by the additional structure it imposes on individual preferences. Specifically, as pointed out by many authors, logit and nested logit models impose a number of important restrictions (e.g., Mcfadden [34], Train [50], [51] and [53]), two of which are of particular concern with regard to the investigation of the issue of consistency between RP and SP data. First, the coefficients of variables that enter the model are assumed to be the same for all individuals. This assumption implies that different people with the same observed characteristics have the same preference for each factor entering the model. This may not be true in reality. People's tastes vary across individuals on the basis of unobserved as well as observed characteristics. In the context of comparing RP and SP data sources, this heterogeneity of preferences may be particularly important. It would be useful to know, for example, if for some portion of the population the discrepancy between RP and SP data is particularly large, so that the source of this discrepancy can be more readily isolated, or particularly small, so that more confidence can be placed in the combined results for this group. Furthermore, it would be helpful to know if the discrepancies that do exist take the form of a uniform shift in parameters or an increased variability of parameters in the case of SP data. Some have suggested that the latter might be expected in cases where the individual faces an unfamiliar option in a stated preference question.

Second, logit exhibits the "independence from irrelevant alternatives" (IIA) for all alternatives. In a typical logit model setting, the IIA property is apparent by noting the structure of relative choice probabilities

$$\frac{p_{ijt}}{p_{ikt}} = e^{V_{ijt}-V_{ikt}}. \tag{4.4}$$

Thus, the ratio between p_{ijt} and p_{ikt} is independent from alternatives other than j and k. While nested logit does not exhibit this property across alternatives from different nests, it does within each nest. Furthermore, nested logit suffers from an analogous problem across nests. Its so-called independence of irrelevant nests (IIN) property implies that the relative probabilities of any two alternatives (say j and k) are independent of alternatives outside of the nests to which j and k belong. As a result of IIN, these models exhibit proportionate substitution patterns. In other words, if some attributes of alternative l change (where l is not in the same nest as j or k), the ratio (4.4) must stay constant. Herriges and Phaneuf [25] detail the implied restriction on the substitution patterns in both logit and nested logit models. These substitution patterns may be unrealistic. In the context of attempting to combine SP and RP data sources, discrepancies found between the two data sources may be due to unrealistic structure imposed by the logit or nested logit model, rather than actual preference discrepancies.

To overcome these limitations, a more flexible model is needed. In this essay, I will employ the mixed logit framework of McFadden and Train [35], which generalizes standard logit by allowing the coefficients vary randomly. This generalization addresses the first limitation directly and also allows for more general patterns of correlation and substitution than either logit or nested logit.

4.3 BASIC IDEA OF MIXED LOGIT MODEL

The mixed logit model is defined as a model whose choice probabilities are the integrals of standard logit probabilities over a density of parameters (McFadden and Train [35], Train [53]). Following Train [53], the probability for individual i choosing alternative j on choice occasion t can be expressed as:

$$p_{ijt} = \int p_{ijt}(\beta) f(\beta) d\beta, \tag{4.5}$$

where $p_{ijt}(\beta)$ is the logit probability conditioned on the parameter vector β :

$$p_{ijt}(\beta) = \frac{e^{V_{ijt}(\beta)}}{\sum_{k=1}^{J} e^{V_{ikt}(\beta)}}; \tag{4.6}$$

$f(\beta)$ is the density function for β; and $V_{ijt}(\beta)$ is the observed portion of utility whose value depends on parameter vector β as well as observed attributes (denoted by vector x_{ijt}).

Still assuming that V_{ijt} is linear in parameters, i.e., $V_{ijt}(\beta) = \beta' x_{ijt}$, the conditional logit probability (4.6) becomes:

$$p_{ijt}(\beta) = \frac{e^{\beta' x_{ijt}}}{\sum_{k=1}^{J} e^{\beta' x_{ikt}}}. \tag{4.7}$$

Thus, the mixed logit probability is a weighted average of the of the logit probabilities $p_{ijt}(\beta)$ evaluated at different coefficient values β, with the weight given by density $f(\beta)$. In other words, mixed logit is just a mixture of the logit probabilities. Standard logit is then a special case in which the density of β is degenerate at a fixed point b; i.e.,

$$f(\beta) = \begin{cases} 1 & \text{if } \beta = b \\ 0 & \text{otherwise} \end{cases}.$$

McFadden and Train ([35]) show that under mild regularity conditions, any discrete choice model derived from random-utility maximization has choice probabilities that can be approximated as accurately as desired by a mixed logit. This can be done by choosing appropriate explanatory variables and assigning appropriate distributions for the random parameters in a mixed logit model. As noted by Herriges and Phaneuf ([25]), the mixed model can be used to capture complex correlation patterns across alternatives and/or choice occasions by introducing additional error components.

The primary interest of mixed model analysis is the distribution of the parameters.[9] The usual technique is to assume that β has some distributions (e.g., normal, uniform, lognormal) which can be characterized by some hyper-parameters, say θ. Then the model becomes

$$p_{ijt}(\theta) = \int p_{ijt}(\beta) f(\beta|\theta) d\beta, \tag{4.8}$$

[9] In Bayes framework, it is straight forward to model this mixed structure with a hierarchical model and use Bayesian tools to analyze the posterior distribution of parameters.

where $p_{ijt}(\beta)$ is defined as in (4.7). If θ is given, the density function of β is known and the integral probability (4.5) can be calculated. In other words, the probability (4.8) is a function of θ, with the parameter β being integrated out. In this setting, estimating θ is of interest by researchers. In general, the integral (4.8) does not have a closed form. Thus it is difficult to calculate the exact (log) likelihood function and the analytical maximum likelihood estimates for the parameters. Nevertheless, as computing technology advances, numerical solutions for the estimation problem are feasible. To date, three simulation methods have been well developed to estimate the model: maximum simulated likelihood (MSL), method of simulated moments (MSM) and method of simulated scores (MSS), which are addressed in detail in Train [53].

In our application, the MSL estimation method will be used for the analysis. Given the specification (4.8), and assuming independency for the decisions across choice occasions, the log-likelihood function can be expressed as

$$LL(\theta) = \sum_{i=1}^{N} \sum_{j=1}^{J} \sum_{t=1}^{T} I_{ijt} \ln \left(p_{ijt}(\theta) \right), \tag{4.9}$$

where dummy variable $I_{ijt} = 1$ if individual i chooses alternative j on choice occasion t and takes the value 0 otherwise. For any given value of θ, this log-likelihood function $LL(\theta)$ can be approximated through simulation:

(1) Draw a large set of values of β from $f(\beta|\theta)$, say, $\beta^{(1)}, ..., \beta^{(S)}$ with the superscript referring to the s-th draw, $s = 1, ...S$ and S is total number of draws;

(2) Calculate $p_{ijt}(\beta)$ for each drawn value of β : $p_{ijt}(\beta^{(1)}), ..., p_{ijt}(\beta^{(S)})$;

(3) Take the average value of $p_{ijt}(\beta^{(1)}), ..., p_{ijt}(\beta^{(S)})$ as the simulated probability of $p_{ijt}(\beta)$:

$$\tilde{p}_{ijt}(\theta) = \frac{1}{S}\sum_{s=1}^{S} p_{ijt}(\beta^{(s)});$$

(4) Repeat step 1-3 for each observation;

(5) Substitute the simulated probabilities calculated from step 1-4 into the log-likelihood function (4.9), get the simulated likelihood function:

$$SLL(\theta) = \sum_{i=1}^{N}\sum_{j=1}^{J}\sum_{t=1}^{T} I_{ijt}\ln\left(\tilde{p}_{ijt}(\theta)\right).$$

The maximum simulated likelihood estimator (MSLE) is the value of θ that maximizes SLL. As noted by Train [51] and [53], even though $\tilde{p}_{ijt}(\theta)$ is unbiased for $p_{ijt}(\theta)$, SLL is not unbiased for LL because of the log transformation. The bias decreases as the number of draws S increases. When S rises faster than the square root of the number of observations, MSLE is asymptotically equivalent to the maximum likelihood estimator.

The above simulation process involves drawing from a density function. Most statistical packages provide random number generators for densities of standard distributions such as normal and uniform. The random draws can be based on these generators. As we all know, draws from those routines are actually pseudo-random numbers, which are calculated through a certain algorithm to mimic the randomness of distributions. In regards to MSLE, McFadden and Train [35] point out, “It is also possible to allow dependence across the different simulation draws, provided there is sufficient mixing for them to satisfy a central limit property.” Through Monte Carlo studies, Train [52] found that patterned pseudo-random numbers such as Halton sequences give estimators lower mean square errors than independent random draws. Further more, Halton sequences are proved to be more efficient than the independent pseudo-random draws (Bhat [9], Train [52]). A robust asymptotic covariance matrix estimator is also recommended by McFadden and Train [35], which is

$$\Gamma(\theta)^{-1}\Delta(\theta)\Gamma(\theta)^{-1};$$

where

$$\Gamma(\theta) = -\mathrm{E}_S\left[\nabla_{\theta\theta'}LL(\theta)\right];$$

$$\Delta(\theta) = \mathrm{E}_S\left[\{\nabla_{\theta}LL(\theta)\}\{\nabla_{\theta'}LL(\theta)\}'\right]$$

with E_S denoting empirical expectation for a random sample of size S, $\nabla_{\theta}LL(\theta)$ and $\nabla_{\theta\theta'}LL(\theta)$ denoting the first and second derivative of $LL(\theta)$ with respect to vector θ respectively. In the mixed logit analysis presented below, I employ Halton sequences, as recommended by Train [52] and the robust covariance matrix described above.

4.4 DATA SOURCE

The (Repeated) Mixed Logit models specified in previous section are applied to the revealed and stated trip data from Iowa Wetlands Survey described in Chapter 3. Unlike Chapter 3, where wetland visitations were modeled on a megazone by megazone basis, the statewide sample is modeled and trips to wetlands throughout the state are modeled simultaneously using the repeated mixed logit model (RXL). Specifically, I consider a random utility model where individuals choose on 52 choice occasions to stay at home or take trips to Iowa wetlands in one of the 5 megazones described in Chapter 3. For each individual, the decisions made in those 52 occasions are assumed to be independent. A separate model is developed for the RP and SP data.

The following individual characteristic variables are common for both revealed preference data and stated preference data: age (years), income (thousand dollars), license (a dummy variable that indicates whether a individual owns hunting license or fishing license) and gender. The summary statistics (means and standard deviations) for these variables are shown in Table 4.1(a). Besides the individual characteristic variables, RP data consist of the numbers of trips taken to 5 megazones as well as the numbers staying home in 52 (weekly) occasions in 1996, and the average travel costs to each megazone. Here the travel cost to each zone is calculated based on travel distance (priced at 0.21 cents per mile) and the cost of travel time (priced at one-third the individual's wage rate). And the average cost to each megazone is the weighted average (weighted by the number of trips to each zone) of the costs to the zones within that megazone. As commonly done in the literature, the counterparts of SP data are based on individuals' responds to the hypothetical scenarios that relate individuals' usage of wetlands to changes of travel costs. For example, respondents were asked how their behavior would change if (1) the total cost per trip of each of their trips had been $15 more, and (2) the total cost per trip to visit the megazone where they lived had been $15 more.[10] A copy of Iowa Wetlands Survey is attached in Appendix for reference. The summary statistics for these variables are also shown in Table 4.1(b).

4.5 MODEL SPECIFICATION

I focus on modeling RP data on trips to the five megazones and SP data regarding planned trips in the face of an increased travel cost (i.e., an entrance fee). Two sets of

[10] The amount of money varied across surveys, with bid values $5, $10, and $15 each randomly assigned to 20% of the sample and bid values of $20, $30, $40, and $50 each assigned 10% of the sample.

variables are included in the analysis: a set of individual characteristic variables (income, age, owning fishing or hunting license or not, and gender) and a site specific variable (price). It would be ideal if we could allow all of the parameters in the choice probabilities to vary by individual. However, as Rudd [43] and Revelt and Train [41] point out, mixed logit models have a tendency to be unstable when all coefficients are allowed to vary. In addition, most of our explanatory variables are individual specific characteristic so that, for any individual, these variables are the same across alternatives. For these reasons, the models will be based on a mixture of individual level taste's parameters (random coefficients) and representative preference's parameters (fixed coefficients). Specifically, we assume that only the price coefficient (i.e., the marginal utility of income) varies across individual and other coefficients are fixed. Three model specifications will be investigated. Since the price coefficient is necessarily negative, it is assumed to have a lognormal distribution in the first two specifications.[11] For simplicity and computational convenience, the price coefficient is assumed to have a normal distribution in the third model, where the parameters of mixing distribution are allowed to depend upon individual characteristics and are correlated between the RP and SP data.

4.5.1 Basic Model

First I will consider a simple model where SP model is assumed to be independent of RP model. For both SP and RP model, the systematic components of the utility functions are assumed of linear form, and the following function is specified for the utility associated with

[11] Actually, the price coefficient for individual i is set to be $-\beta_{price,i}$, where $\beta_{price,i}$ is Lognormally distributed; i.e., the negative of the price enters the model. Note that $\beta_{price,i}$ is always positive.

individual i choosing site j or staying at home $(j = 0)$ at period t :

$$U_{ijt} = \begin{cases} \beta_{00} + \beta_A \text{Age}_i + \beta_I \text{Income}_i + \varphi_{0i} + \varepsilon_{i0t} & j = 0; \\ \beta_{0j} - \beta_{pi} \text{Price}_{ij} + \varphi_{1i} + \varepsilon_{ijt} & j = 1, ..., 5, \end{cases} \tag{4.10}$$

where

$$\varphi_{0i} \sim N\left(0, \sigma^2_{\varphi 0}\right),$$

$$\varphi_{1i} \sim N\left(0, \sigma^2_{\varphi 1}\right),$$

ε_{ijt} are i.i.d distributed extreme value errors with variance σ^2_{ε}; φ_{0i}, φ_{1i} and ε_{ijt} are independently distributed. In this utility function, β_{0j}'s are alternative specific constants, β_A is the age coefficient, β_I is income coefficient, $-\beta_{pi}$ is individual specific price coefficient.[12] The price coefficient is expected to be negative for each individual with only the magnitude differing, thus a lognormal distribution is assigned to β_{pi}. The mean and standard deviation of the $ln(\beta_{pi})$ are estimated, and the mean and standard deviation of the estimated β_{pi} are then calculated. Since the lognormal distribution is defined over the positive range, the negative of the price enters the model.

Note here φ_{0i} and φ_{1i} are error components set to capture unobservable factors that underlie the nesting correlation patterns. φ_{0i} is the error component associated with the nest of "staying at home", while φ_{1i} associated with "taking trips". For example, an individual who owns a boat or a sport utility vehicle will have large φ_{1i} which is associated with $U_{i1t}, ..., U_{i5t}$. Such positive factor values make them more likely to visit one of the sites rather than staying at home. On the other hand, an individual who has a serious illness tends to stay at home as opposed to going on a trip, thus he will have a large positive φ_{0i} which is associated with

[12] Note that not all of the parameters are identified in model (4.10). To estimate the model, parameter normalization is necessary and this is discussed below.

U_{i0t} (or alternatively a large negative φ_{1i}). These factors are known to each individual but are not observed by analyst.[13] Thus they enter the model as error terms. To see how this works, consider the covariance structure of the utility function. For individual i and choice occasion t, the variance-covariance matrix for utilities across alternatives (conditioning on the β coefficients) is

$$Var(U_{i.t}) = \begin{bmatrix} \sigma_{00} & 0 & \cdots & 0 \\ 0 & \sigma_{11} & \cdots & \sigma_{15} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \sigma_{51} & \cdots & \sigma_{55} \end{bmatrix} = \begin{bmatrix} \sigma_{00} & 0_{1x5} \\ 0_{5x1} & \Sigma_{trip} \end{bmatrix},$$

where

$$\sigma_{00} = \sigma_{\varphi 0}^2 + \sigma_{\varepsilon}^2;$$

$$\sigma_{jj} = \sigma_{\varphi 1}^2 + \sigma_{\varepsilon}^2, \qquad \text{for } j = 1, ..., 5;$$

$$\sigma_{jk} = \sigma_{\varphi 1}^2, \qquad \text{for } j, k = 1, ..., 5 \text{ and } j \neq k;$$

and

$$\Sigma_{trip} = (\sigma_{\varphi 1}^2 + \sigma_{\varepsilon}^2) \begin{bmatrix} 1 & \rho & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho & \rho \\ \rho & \rho & 1 & \rho & \rho \\ \rho & \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & \rho & 1 \end{bmatrix},$$

13 One can certainly incorporate the factor of owning a boat or a sport utility vehicle into the survey when design the questionnaire. However, it is unlikely for a research to cover every factor affecting individual's utility from a given choice alternative (e.g., health, past experiences, etc.).

where $\rho = \frac{\sigma_{\varphi 1}^2}{\sigma_{\varphi 1}^2 + \sigma_\varepsilon^2} \in (0,1)$ measures the correlation among the utilities associated with the trip alternatives. Clearly, this mixed logit model captures a correlation pattern the same way as a nested logit model does (Herriges and Phaneuf [25]).

However, not all of the parameters of the model in (4.10) are identified. Thus, it is necessary to normalize some of the parameters in the model in order to estimate the model. The mixed logit model consists two sets of relevant parameters, a set for the fixed part of the model ($\beta_{00}, \beta_{01}, .., \beta_{05}, \beta_A$, and β_I) and a set for the random part ($\beta_{pi}, \varphi_{0i}, \varphi_{01}$ and ε_{ijt}). Identification and normalization for the parameters of the fixed part is exactly the same as that for a multinomial logit model, which are well understood. With regards to the identification of the parameters of the random part in the mixed logit model, Ben-Akiva *et al.* [7] discuss several issues in detail. They suggest that three conditions be examined: order condition, rank condition and positive definiteness condition. In our model the variance of ε_{ijt}, σ_ε^2, is normalized to $\pi^2/6$, the variance of a standard extreme value random variable. For the error components φ_{0i} and φ_{01}, only the sum of their variance ($\sigma_{\varphi 0}^2 + \sigma_{\varphi 1}^2$) is identifiable given the constraint on the variance of ε_{ijt}. For computation convenience, $\sigma_{\varphi 1}^2$ is normalized to be zero. The alternative specific constant β_{05} is also normalized to be zero.

The utility function (4.10) can be used to examine RP or SP data separately. However, assuming the ε_{ijt}'s are independent across the RP and SP sources, one can also stack the RP and SP model with separate coefficient sets to carry out tests for consistency between the two data sets. This forms our first, or basic model specification. The assumption of independence is relaxed in subsequent models.

4.5.2 Extended Models

The basic model provides an initial view of the preferences in a mixed logit framework. However, it ignores the obvious relationship between RP and SP survey responses since it assumes independence between the two sets preferences elicited from the same individuals. To further investigate the RP and SP, I extend the basic model in two directions: one is to allow the correlation between RP and SP survey responses and the other is to explore the discrepancy between RP and SP specific parameters by allowing the parameters of the mixing distributions to depend upon individual characteristics.

4.5.2.1 *Including Correlation between RP and SP*

The assumption of independence between the RP and SP data sources is likely to be unrealistic. There are always some factors (e.g., knowledge of a specific lake, a bad experience in visiting a site) that are known by individuals but not observed by analyst that affect decisions of individuals to visit certain sites or not. The same factors are also likely to affect individuals' future decisions about choosing sites. These unobservable factors create a positive correlation between RP and SP. This will be incorporated into the model in our second model specification. Specifically, the utility function associated with individual i choosing site j or staying at home ($j = 0$) at period t is assumed to be

$$U_{ijt}^{k} = \begin{cases} \beta_{00}^{k} + \beta_{A}^{k}\text{Age}_i + \beta_{I}^{k}\text{Income}_i + \varphi_{0i}^{k} + \xi_{i0} + \varepsilon_{i0t}^{k} & j = 0; \\ \beta_{0j}^{k} - \beta_{pi}^{k}\text{Price}_{ij}^{k} + \varphi_{1i}^{k} + \xi_{ij} + \varepsilon_{ijt}^{k} & j = 1, ..., 5, \end{cases} \tag{4.11}$$

where

$$\varphi_{0i}^{k} \sim N\left(0, \sigma_{\varphi_0^k}^2\right);$$
$$\varphi_{1i}^{k} \sim N\left(0, \sigma_{\varphi_1^k}^2\right);$$

$$\xi_{ij} \sim N\left(0, \sigma^2_{\xi_j}\right), \quad \text{for } j = 0, 1, ..., 5;$$

and ε^k_{ijt}'s are i.i.d distributed extreme value errors. Note here the superscript k refers to the parameters of RP or SP model. The basic setup for separate RP or SP model is essentially the same as the basic model (4.10) with a single additional error component, ξ_{ij}. This error component is introduced to capture the positive correlation between revealed preference and stated preference. Here, ξ_{ij} is associated with individual i choosing alternative j and assumed to be the same in across preferences (RP and SP). Intuitively, an individual who, according to their RP data, chose to visit site j because of unobservable factors (i.e., a large positive ξ_{ij}) is likely to anticipate visiting site j in the future since ξ_{ij} also increases U^{SP}_{ijt}. The identification and normalization for the basic model (4.10) still apply here. For simplicity, it is assumed that the variances of ξ_{ij}'s are the same across alternatives; i.e.

$$\sigma^2_{\xi_j} = \sigma^2_{\xi} \qquad \forall j.$$

Note, for the separate models, σ^2_{ξ} cannot be identified given the normalization setup for other parameters as in the basic model. However, when combining RP and SP model together, σ^2_{ξ} can be identified. See the appendix to this chapter (4.A) for details.

4.5.2.2 *Conditioning Factors*

In the third model specification, I consider allowing the parameters of the mixing distributions to depend upon individual characteristics, such as age or gender. In this way, it may be possible to identify population subgroups that exhibit more or less consistency between their RP and SP responses. More specifically, starting from the basic random utility function

$$U_{ijt}^{k} = \begin{cases} \beta_{00}^{k} + \beta_{A}^{k}\text{Age}_i + \beta_{I}^{k}\text{Income}_i + \varphi_{0i}^{k} + \varepsilon_{i0t}^{k} & j = 0; \\ \beta_{0j}^{k} - \beta_{pi}^{k}\text{Price}_{ij}^{k} + \varphi_{1i}^{k} + \varepsilon_{ijt}^{k} & j = 1, ..., 5; \end{cases} \quad (4.12)$$
$$\text{for } k = RP,\ SP,$$

The basic model is generalized by assuming that there is linear dependency between β_{pi}^{RP} and β_{pi}^{SP}. Specifically, the model involves linear combination of mixing distributions. To make the model tractable, I assume normality for the price coefficients.[14] The main reason for doing so is that the linear combinations of normal distributions are still normal, which makes the interpretation straightforward and simplifies the computations. Except for the distributional assumption of the price coefficients, the setup as well as the normalization for this model is the same as that of our basic model (4.10).

For the RP price coefficient, we assume

$$\beta_{pi}^{RP} \sim N\left(\bar{\beta}_{p}^{RP}, \sigma_{\beta^{RP}}^{2}\right).$$

The SP price coefficient is assumed to have some departure from the RP price coefficient:

$$\beta_{pi}^{SP} = \beta_{pi}^{RP} + \delta_i.$$

Further, this departure is assumed to depend on the gender and license variables:

$$\delta_i = \alpha_{0i} + \alpha_g \text{Gender}_i + \alpha_l \text{License}_i;$$

where

$$\alpha_{0i} \sim N\left(\bar{\alpha}_0, \sigma_{\alpha_0}^{2}\right);$$

and β_{pi}^{RP} and α_{0i} are independently distributed. Note this structure imposes a restriction on the randomness of β_{pi}^{RP} and β_{pi}^{SP}, or $Var(\beta_{pi}^{SP}) = \sigma_{\beta^{RP}}^{2} + \sigma_{\alpha_0}^{2} > \sigma_{\beta^{RP}}^{2} = Var(\beta_{pi}^{RP})$.

[14] Notice that the terms ξ_{ij} introduced in equation (4.11) are not used here. However, correlation between the RP and SP response will now arise through the price coefficient.

This restriction is reasonable, since there may be additional variability in the underlying SP preferences due to the forward looking nature of the SP question and lack of familiarity with the new conditions. Indeed, in the results reported for the basic model below, where no correlations are allowed, the estimated variance of β_{pi}^{SP} is bigger than that of β_{pi}^{RP}. Another consequence is the positive correlation between β_{pi}^{RP} and β_{pi}^{SP}.

4.6 MODEL ESTIMATION AND RESULTS

Maximum Simulated Likelihood Estimation method was used to estimate the mixed logit models. For each individual, 1000 Halton draws was used during the simulation. The Gauss programs used throughout the estimation procedure were those developed by Train, Revelt and Rudd, and modified by Herriges.

4.6.1 Model 1: Basic Results

The final point estimates for the basic model are shown in column 3 of Table 4.2.[15] For comparison, the results for standard (repeated logit) version of the basic model, in which the price coefficients are assumed to be fixed instead of random, are also shown in column 2 of the same table. The detailed results (including point estimates, asymptotic and robust standard errors, and p-values) are provided in the appendix to this chapter (4.B). There is substantial

[15] Unfortunately, the log-likelihood associated with the mixed logit model is not globally concave, so that one cannot be assured a local maximum is indeed a global maximum. I did encounter cases in which different starting values yielded different converged parameter estimates. Throughout this chapter, the results reported are those yielding the largest log-likelihood. I also provide in the appendix (4.C) to this chapter results derived from different starting values. These alternative parameter estimates, while different, do not yield qualitatively different results. The problem appears to lie primarily with the RP estimates rather than the SP estimates. This may be due to the greater price variability in the SP data.

and significant improvement in the log likelihood function by allowing the price coefficient to vary across individuals (−58790.29 for the mixed model vs. −79877.02 for standard model).

Given the sample size (with 1558 individuals and 52 observations for each individual) it is not surprising that, the point estimates for all the parameters associated with age, income and price are highly significant with p-value's less than 0.01. The signs of coefficients of age, income and price are consistent with expectations in both the basic repeated logit model (column 2) and its mixed logit counterpart (column 3). Within each model, the parameters are also similar across the RP and SP data sources. As expected, on average the marginal utility of staying at home is positive with respect to age and negative with respect to income. In other words, younger people and people with higher income tend to visit wetlands more often as they have higher utilities associated with taking trips to wetlands. Also as expected, higher travel costs, including gas price, travel time, entrance fees and so on, have a negative impact on individual's utility associates with taking trips. Comparing column 2 with column 1, we can see that the results from mixed model are comparable with the results from the standard logit model in magnitude, though some differences do exist. Nevertheless, one can observe that the means of the price coefficients in the mixed model are very close to their counterpart in the standard logit model results.

Two structural differences exist between the basic repeated logit model of column 2 and the mixed logit model of column 3. First, the mixed logit model allows for a nesting structure (grouping wetland trips as distinct from the staying at home option). The magnitude and significance of the estimates for $\sigma_{\varphi_0^{RP}}$ and $\sigma_{\varphi_0^{SP}}$ provide strong evidence for the existence of the correlation patterns within the nests (staying home and taking trips).

Second, the mixed logit model allows the price coefficient (i.e., the marginal utility of income) to vary across individuals. The estimates of the standard deviation of the price coefficients suggest that there is significant variation of taste across individuals in both RP and SP data, with more variability observed in SP data. Figure 4.1 also provide a visual look at variability of the price coefficients. As we can see from the density plots, the means of the price coefficients are close for the RP and SP, with most of the density on the range between 0 and 0.08.

At a glance, we find that except for the alternative specific constants, all the point estimates are very close in magnitude for RP and SP data. Three null hypotheses are considered to explicitly test consistency between the RP and SP data:

- H_0^1: Age, income coefficients are equal, mean and standard deviation of price coefficient are the same for RP and SP;
- H_0^2: Age, income coefficients are equal for RP and SP;
- H_0^3: Mean and standard deviation of price coefficient are the same for RP and SP.

Wald test statistics were computed using both asymptotic and robust standard errors. The results are shown in Table 4.3.[16] The first and the third hypotheses were uniformly rejected at a 1% significant level with either error computation method. For test 2, we failed to reject the hypothesis at a 5% significant level with the basic model. These tests suggest, although the estimates for RP and SP data look very similar in magnitude, statistically there exist discrepancies between these two sets of data.

---

[16] As with the basic mixed logit model, different MSLE estimates were obtained starting from different initial values. Table 4.3 shows the test results based on the estimates with the largest likelihood function. A second set of results based on different starting values are shown in the appendix to this chapter (4.C).

4.6.2 Model 2: Allowing for RP/SP Correlation

The point estimates results for our second specification are shown in column 4 of Table 4.2,[17] with detailed results provided in the appendix to this chapter (4.B). This model is again a significant improvement over the standard logit version of the basic model. The coefficients for the age, income and price are generally similar in size, sign and significance to those reported for the basic mixed logit model. The biggest shifts appear in the alternative specific constants. We still see a great deal of similarity between the RP and SP parameters associated with age income and price.

The major change introduced in this second mixed logit model is that it allows for correlation between the RP and SP responses. Indeed, the significance and magnitude of the estimate of σ_ξ does suggest strong (positive) correlation between RP and SP. Indeed, σ_ξ is larger in magnitude than either $\sigma_{\varphi_0^{RP}}$ and $\sigma_{\varphi_0^{SP}}$, though the latter terms remain statistically significant.

Turning to the various hypothesis test regarding consistency, corresponding Wald test statistics are shown in Table 4.3.[18] As in model 1, the first (age, income coefficients are equal, mean and standard deviation of price coefficient are the same for RP and SP) and the third (mean and standard deviation of price coefficient are the same for RP and SP) hypotheses were uniformly rejected at a 1% significant level with either error computation method. For test 2 (mean and standard deviation of price coefficient are equal for RP and SP), the outcomes

[17] These results are based on the MSLE with largest converged likelihood function. A second set of results derived from different starting values is shown in the appendix to this chapter (4.C). See footnote [15].

[18] These testing results are based on the MSLE with the largest likelihood function I find in the study. Another set of testing results based on different starting values are shown in the appendix to this chapter (4.C), see footnote [12].

are somewhat mixed. If using the standard asymptotic standard error, we will reject the hypothesis at a 5% significant level; but fail to do so if the robust standard errors are used.

4.6.3 Model 3: Conditioning Factors

The results for the model with conditioning factors are shown in Table 4.4, where the point estimates and asymptotic standard errors are reported. The results for the corresponding standard logit model are also shown. As we do in the first two models, we observe a significant improvement in log likelihood function for mixed logit model (-57699.26) over the repeated logit model (-79415.20). The point estimates are all highly significant with in the model. Overall, the third model has similar results as the basic model does for their common parameters (constants, coefficients for age and income, and the nested structure parameters).

The key difference between model 3 and the previous two mixed models is the way in which the price coefficient (and its variability in the population) is modeled. Specifically, we now have

$$\beta_{pi}^{RP} \sim N\left(\bar{\beta}_p^{RP}, \sigma_{\beta^{RP}}^2\right);$$
$$\beta_{pi}^{SP} = \beta_{pi}^{RP} + \alpha_{0i} + \alpha_g \text{Gender}_i + \alpha_l \text{License}_i,$$

where

$$\alpha_{0i} \sim N\left(\bar{\alpha}_0, \sigma_{\alpha_0}^2\right).$$

Thus, the discrepancy between the RP and SP price coefficient is

$$\begin{aligned}\delta_i &= \alpha_{0i} + \alpha_g \text{Gender} + \alpha_l \text{License} \\ &\sim N(\bar{\alpha}_0 + \alpha_g \text{Gender} + \alpha_l \text{License}, \sigma_{\alpha_0}^2)\end{aligned}$$

and depends on the characteristics of the individual. The coefficient estimates in Table 4.4 indicate that price coefficients are, similar at the mean characteristics.[19]

	RP Model		SP Model	
	Mean	Std. Dev.	Mean	Std. Dev.
Basic Model (1)	-0.0391	0.0417	-0.0377	0.0482
Conditioning Factors	-0.0330	0.0442	-0.0391	0.0482

More importantly, this model provides additional information about the discrepancy between RP and SP over the first two models. The magnitude of the standard deviation of α_0 confirms our earlier assumption that there exists additional variation in the tastes of SP compared with the RP. Furthermore, we found that there exist gender and license effects associated with the discrepancy between the tastes of RP and SP. These effects are statistically significant and not small at all compared with the average discrepancy between RP and SP. The gender effect is $\alpha_1 = -0.0230$, where as license effect $\alpha_2 = -0.0129$. The discrepancy between RP and SP price coefficients on average is $\bar{\delta} = \bar{\beta}_p^{SP} - \bar{\beta}_p^{RP} = -0.0061$. However, males and individuals who own licenses tend to overstate their price response, with $\bar{\delta} = -0.017$.

Estimation of δ

δ	Male	Female
License	-0.017	0.006
No License	-0.004	0.019

[19] In the model with conditioning factors, the price coefficients reported are calculated according the following formulas

$$\bar{\beta}_p^{sp} = \bar{\beta}_p^{rp} + \bar{\alpha}_0 + \alpha_g \text{Mean(Gender)} + \alpha_l \text{Mean(License)};$$
$$\sigma^2_{\beta^{sp}} = \sigma^2_{\beta^{rp}} + \sigma^2_{\alpha_0}.$$

4.6.4 Welfare Analysis

We will apply the welfare analysis towards the first two specifications. As in Chapter 2, the following simple form is used to calculate consumer surplus

$$CS = \frac{q^2}{-2\beta_{price}}. \tag{4.13}$$

We consider the CS for an individual who took only one trip to the wetland during the year of survey. Since lognormal distributions are used to describe the randomness of the price coefficients, numerical simulations are easily to be carried out to perform the analysis. Figure 4.2 provides a graphical view of distributions of the normalized CS. In both models, we observe higher mean value for CS for SP model comparing with RP model. This suggests that on average, CS tends to be overstated in SP than RP. Also, bigger variations are associated with SP rather than RP. Another important point is that if correlation between RP and SP is not counted, the CS will be over estimated in general.

4.6.5 Prediction

Prediction exercises were also carried out using the first two models in our study. Three scenarios were considered with regard to the total number of choices for 6 alternatives (staying home, taking trips to wetlands of 5 megazones) over a one year period with 52 choice occasions:

- An individual with age, income and costs to each zone equal to the average level of these variables;
- An individual with age, income equal to the average level of these variables, while costs to each zone are 10 dollars above the average level;

- An individual with age, income equal to the average level of these variables, while costs to each zone are 10 dollars below the average level.

The prediction results are shown in Table 4.5. The results illustrate that as expected people tend to take more trips to wetlands when the costs to visit wetlands decrease; SP models tend to predict higher probabilities for taking trips; and once the correlation between RP and SP is included, the variation for predicted numbers increase dramatically.

4.7 DISCUSSION AND CONCLUSION

The mixed logit framework provides considerable flexibility over standard logit models. This allows us to investigate heterogeneity in individual level preferences as well as to explore underlying error structures of the model. In the application to the Iowa Wetlands Survey data, we are able to use this framework to link the RP and SP data and investigate the discrepancy issues associated with these two types of data. We found several key results. First, the mixed logit models provide better fit over the standard logit model. Second, though the general pattern of effects (nesting structure, demographic effects, and RP/SP differences) did not change substantially, the correlation between RP and SP may not be ignored. Third, our model 3 suggest that there exist gender and license differences in the discrepancies between RP and SP responses. Males and those holding licenses exhibited greater differences in their RP and SP responses. However, our analysis of these effects were somewhat limited by the available sociodemographic and site-specific characteristics. Further research can be done when more characteristic variables are available.

Table 4.1. Summary Statistics for Data Used in Chapter 2 (Number of Observation: 1558)

(a) Means of Individual Characteristic Variables

	Mean	Std.
Age (years)	49.90	16.63
Income($1000)	41.56	29.14
License[1]	0.66	0.48
Gender[2]	0.73	0.44

(b) Means of Alternative Specific Variables (Standard Deviation is shown in parenthesis)

Alternative[3]		0	1	2	3	4	5
RP	# Choices	44.91	0.74	0.89	1.45	2.78	1.22
		(11.47)	(3.54)	(3.51)	(5.22)	(6.72)	(4.67)
	Cost	0.00	144.78	111.66	98.49	72.24	110.78
		(0.00)	(65.28)	(53.85)	(47.03)	(48.28)	(70.54)
SP	# Choices	45.75	0.70	0.82	1.53	1.86	1.34
		(10.40)	(3.25)	(3.11)	(4.76)	(5.02)	(4.71)
	Cost	0.00	146.69	113.71	101.38	83.88	113.31
		(0.00)	(62.67)	(51.69)	(45.47)	(44.89)	(68.85)

[1] License=1 if individual owns a hunting or fishing license, =0 otherwise;
[2] Gender=1 if respondent is male, =0 if female.
[3] Alternative 0: Staying at home; 1-5: Taking trips to wetlands in Megazone 1-5 respectively.

Table 4.2. Estimation Results for Models 1 & 2

Parameter	Repeated Logit	Mixed Logit	
		Basic Model	RP/SP Correlation
β_{00}^{rp}	4.949	7.675	8.458
β_{00}^{sp}	4.395	7.394	7.805
β_{01}^{rp}	-1.029	-0.108(2)	0.273(3)
β_{01}^{sp}	-0.872	-0.246	-0.295
β_{02}^{rp}	-0.057(1)	0.237	0.192(4)
β_{02}^{sp}	-0.256	0.077	0.060(5)
β_{03}^{rp}	0.706	0.703	1.205
β_{03}^{sp}	0.476	0.655	0.601
β_{04}^{rp}	1.465	1.631	2.067
β_{04}^{sp}	0.743	1.068	1.006
β_{Age}^{rp}	0.027	0.071	0.053
β_{Age}^{sp}	0.028	0.065	0.051
β_{Income}^{rp}	-0.017	-0.024	-0.016
β_{Income}^{sp}	-0.016	-0.029	-0.021
$\bar{\beta}_{p}^{rp}$	0.035	0.039	0.034
$\bar{\beta}_{p}^{sp}$	0.028	0.038	0.036
$\sigma_{\beta_p^{rp}}$		0.042	0.024
$\sigma_{\beta_p^{sp}}$		0.048	0.034
$\sigma_{\varphi_0^{rp}}$		2.980	1.139
$\sigma_{\varphi_0^{sp}}$		3.275	1.404
σ_{ξ}			2.295
Log likelihood	-79877.02	-58790.29	-53857.33

Calculated with asymptotic standard errors, p_value's for all the point estimates are less than 0.01, except for those with superscripts: (1) 0.119; (2) 0.113; (3) 0.014; (4)0.043; (5) 0.531.

The detailed results for these models are presented in Appendix 4.B. The (-) price coefficients are assumed to be lognormal in the mixed models and the estimates for the lognormal parameters are provided in Appendix 4.B.

Table 4.3. Results for Hypothesis Tests

TEST 1	H_0:	β_{Age}, β_{Incom}, Mean and Std. Dev.of $\ln(\beta_{price})$. are equal for RP and SP		
	D.F.	4		
		Error Comp. Method	Wald Stat.	P_value
	Model 1*	Asymptotic	27.98	0.000
		Robust	15.49	0.004
	Model 2	Asymptotic	100.51	0.000
		Robust	43.73	0.000
TEST 2	H_0:	β_{Age}, β_{Incom} are equal for RP and SP		
	D.F.	2		
		Error Comp. Method	Wald Stat.	P_value
	Model 1	Asymptotic	2.69	0.261
		Robust	1.82	0.403
	Model 2	Asymptotic	7.18	0.028
		Robust	3.28	0.194
TEST 3	H_0:	Mean and Std. Dev.of $\ln(\beta_{price})$. are equal for RP and SP:		
	D.F.	2		
		Error Comp. Method	Wald Stat.	P_value
	Model 1	Asymptotic	21.12	0.000
		Robust	9.83	0.007
	Model 2	Asymptotic	92.82	0.000
		Robust	42.28	0.000

* Model 1: Basic Mixed Model
Model 2: Mixed Model with Correlation between RP and SP

Table 4.4. Estimation for Model with Conditioning Factors

Parameter	Repeated Logit		Mixed Logit	
	Estimate	Std. Err.	Estimate	Std. err.
β_{00}^{rp}	4.949	0.032	7.814	0.109
β_{00}^{sp}	4.373	0.028	7.685	0.106
β_{01}^{rp}	-1.029	0.041	-0.319	0.074
β_{01}^{sp}	-0.913	0.040	-0.359	0.065
β_{02}^{rp}	-0.057(1)	0.037	0.152(2)	0.065
β_{02}^{sp}	-0.278	0.037	0.103(3)	0.058
β_{03}^{rp}	0.706	0.033	0.833	0.056
β_{03}^{sp}	0.485	0.031	0.816	0.049
β_{04}^{rp}	1.465	0.030	1.795	0.047
β_{04}^{sp}	0.744	0.030	1.270	0.043
β_{Age}^{rp}	0.027	0.001	0.056	0.005
β_{Age}^{sp}	0.026	0.001	0.078	0.004
β_{Income}^{rp}	-0.017	0.000	-0.016	0.003
β_{Income}^{sp}	-0.015	0.000	-0.027	0.002
$\bar{\beta}_{p}^{rp}$	-0.035	0.000	-0.033	0.001
$\sigma_{\beta^{rp}}$			0.044	0.001
$\bar{\alpha}_0$	0.024	0.001	0.019	0.004
σ_{α_0}			0.019	0.001
α_1	-0.004	0.001	-0.023	0.002
α_2	-0.017	0.001	-0.013	0.003
$\sigma_{\varphi_0^{rp}}$			3.189	0.080
$\sigma_{\varphi_0^{sp}}$			3.281	0.066
Log likelihood	-79415.20		-57699.26	

Note: Only asymptotic standard error estimates are provided here; p_value's for all the point estimates are less than 0.01, except for the three with superscripts: (1) 0.119; (2) 0.019; (3) 0.074.

Table 4.5. Prediction with Basic Model

Scenario	Alternative	Mean number of choices			
		Basic Model		Model with Correlation	
		RP	SP	RP	SP
1	0	50.606 (6.872)	50.207 (8.011)	48.686 (10.789)	48.360 (11.343)
	1	0.021 (0.176)	0.053 (0.326)	0.357 (3.382)	0.423 (3.736)
	2	0.046 (0.311)	0.112 (0.592)	0.318 (3.19)	0.561 (4.326)
	3	0.137 (0.766)	0.429 (1.998)	0.791 (5.196)	0.889 (5.544)
	4	1.163 (5.745)	1.105 (4.981)	1.566 (7.406)	1.223 (6.524)
	5	0.027 (0.211)	0.095 (0.518)	0.282 (3.002)	0.544 (4.246)
2	0	50.839 (6.261)	50.505 (7.286)	49.074 (10.148)	48.798 (10.643)
	1	0.017 (0.159)	0.046 (0.311)	0.312 (3.189)	0.374 (3.521)
	2	0.039 (0.284)	0.094 (0.541)	0.283 (2.980)	0.493 (4.036)
	3	0.114 (0.689)	0.356 (1.813)	0.682 (4.804)	0.783 (5.204)
	4	0.968 (5.238)	0.921 (4.527)	1.404 (7.035)	1.077 (6.107)
	5	0.022 (0.190)	0.078 (0.471)	0.244 (2.793)	0.476 (3.951)
3	0	50.331 (7.550)	49.842 (8.781)	48.319 (11.361)	47.875 (12.038)
	1	0.026 (0.202)	0.064 (0.364)	0.382 (3.508)	0.461 (3.835)
	2	0.057 (0.356)	0.134 (0.649)	0.360 (3.417)	0.646 (4.649)
	3	0.165 (0.837)	0.517 (2.197)	0.881 (5.455)	1.000 (5.853)
	4	1.388 (6.301)	1.329 (5.450)	1.743 (7.855)	1.401 (7.020)
	5	0.032 (0.229)	0.114 (0.569)	0.314 (3.180)	0.617 (4.547)

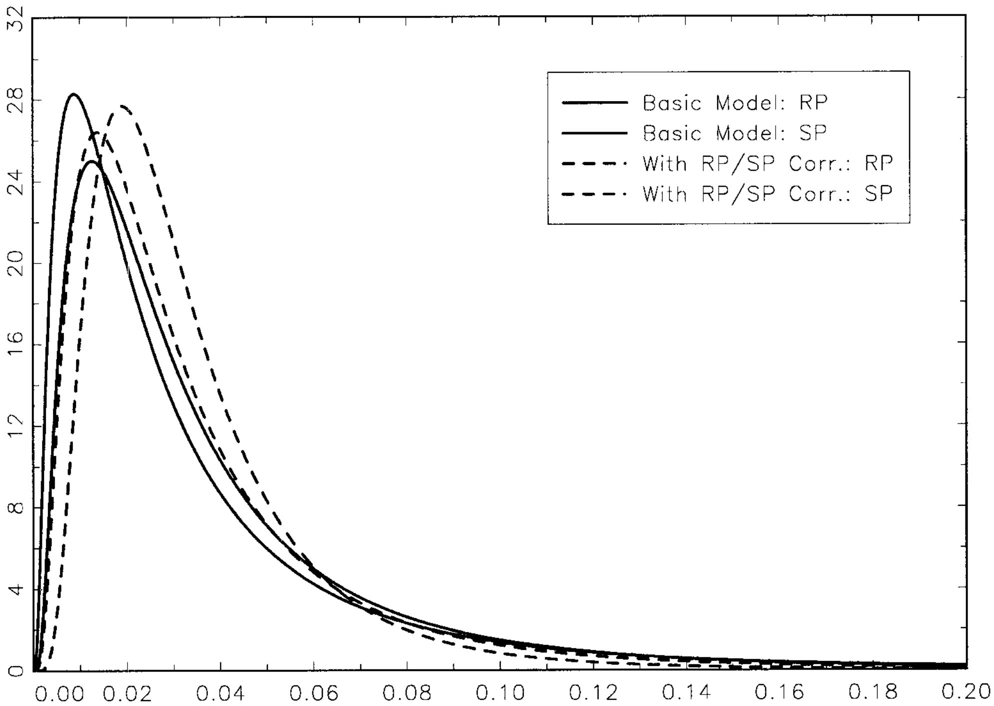


Figure 4.1. Density Plots for (−) Price Coefficients

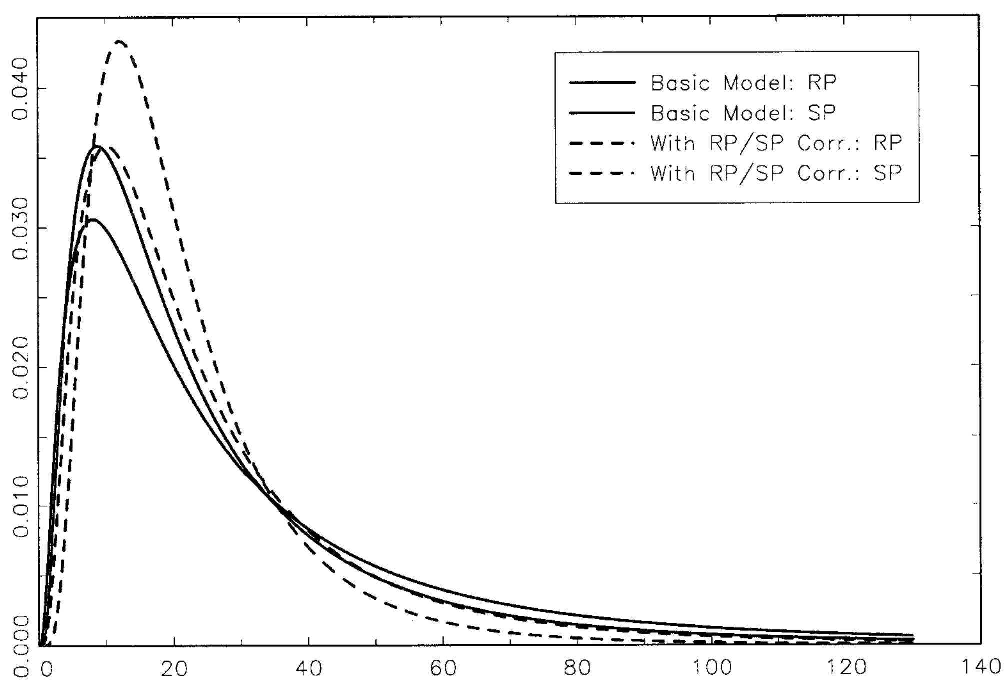


Figure 4.2. Density Plots for Consumer Surplus

Appendix 4.A IDENTIFICATION OF THE ERROR COMPONENTS IN THE MIXED LOGIT MODEL

Since in the random utility model, only the difference of the utility matters, consider the utility difference (base=alternative 0) of model (4.11):

$$\begin{aligned} U^k_{i1t} - U^k_{i0t} &= \ldots + \varphi^k_{1i} + \xi_{i1} + \varepsilon^k_{i1t} - (\varphi^k_{0i} + \xi_{i0} + \varepsilon^k_{i0t}); \\ &\vdots \\ U^k_{i5t} - U^k_{i0t} &= \ldots + \varphi^k_{1i} + \xi_{i5} + \varepsilon^k_{i5t} - (\varphi^k_{0i} + \xi_{i0} + \varepsilon^k_{i0t}). \end{aligned}$$

Noted that the systematic parts are irrelevant and emitted here. The variance-covariance matrix of the error parts for this system is

$$\Omega^k = \begin{bmatrix} \omega_{11} & & & & \\ \omega_{21} & \omega_{22} & & & \\ \omega_{31} & \omega_{32} & \omega_{33} & & \\ \omega_{41} & \omega_{42} & \omega_{43} & \omega_{44} & \\ \omega_{51} & \omega_{52} & \omega_{53} & \omega_{54} & \omega_{55} \end{bmatrix}$$

where

$$\omega_{ii} = \sigma^2_{\varphi^k_1} + \sigma^2_{\varphi^k_0} + 2\sigma^2_\xi + 2\sigma^2_\varepsilon \quad \forall i;$$

$$\omega_{ij} = \sigma^2_{\varphi^k_1} + \sigma^2_{\varphi^k_0} + \sigma^2_\varepsilon \quad \forall i \neq j.$$

For a separate RP or SP model, we have only one degree of freedom due to the need to set the scale, and 4 unknowns. Given normalization of $\sigma^2_{\varphi^k_0} = 0$ and $\sigma^2_\varepsilon = \pi^2/6$, we still can not identified both σ^2_ξ and $\sigma^2_{\varphi^k_1}$. However, for the combined model, σ^2_ξ, $\sigma^2_{\varphi^{RP}_1}$ and $\sigma^2_{\varphi^{RP}_1}$will be identified because of the correlation terms between RP and SP. We can see this from its variance-covariance matrix:

$$\Omega^{all} = \begin{bmatrix} \Omega^{RP} & 2\sigma_{\xi}^{2} I_5 \\ 2\sigma_{\xi}^{2} I_5 & \Omega^{RP} \end{bmatrix}.$$

Appendix 4.B DETAILED ESTIMATION RESULTS FOR MODELS

(a) Repeated Logit Estimation

Param.	Estimates	Asymptotic Error Est.		Robust Error Est.	
		Std. err.	P_value	Std. err.	P_value
β_{00}^{rp}	4.949	0.032	0.000	0.162	0.000
β_{00}^{sp}	4.395	0.028	0.000	0.126	0.000
β_{01}^{rp}	-1.029	0.041	0.000	0.196	0.000
β_{01}^{sp}	-0.872	0.039	0.000	0.168	0.000
β_{02}^{rp}	-0.057	0.037	0.119	0.129	0.658
β_{02}^{sp}	-0.256	0.036	0.000	0.125	0.041
β_{03}^{rp}	0.706	0.033	0.000	0.137	0.000
β_{03}^{sp}	0.476	0.031	0.000	0.119	0.000
β_{04}^{rp}	1.465	0.030	0.000	0.124	0.000
β_{04}^{sp}	0.743	0.030	0.000	0.114	0.000
β_{Age}^{rp}	0.027	0.001	0.000	0.003	0.000
β_{Age}^{sp}	0.028	0.001	0.000	0.003	0.000
β_{Income}^{rp}	-0.017	0.000	0.000	0.001	0.000
β_{Income}^{sp}	-0.016	0.000	0.000	0.002	0.000
$-\beta_{price}^{rp}$	-0.035	0.000	0.000	0.002	0.000
$-\beta_{price}^{sp}$	-0.028	0.000	0.000	0.002	0.000
Log likelihood	-79877.02				

(b) Mixed Logit Estimation (Basic Model)

Parameter	Estimate	Asymptotic Error Est.		Robust Error Est.	
		Std. err.	P_value	Std. err.	P_value
β_{00}^{rp}	7.675	0.125	0.000	0.207	0.000
β_{00}^{sp}	7.394	0.133	0.000	0.211	0.000
β_{01}^{rp}	-0.108	0.068	0.113	0.319	0.734
β_{01}^{sp}	-0.246	0.062	0.000	0.271	0.364
β_{02}^{rp}	0.237	0.061	0.000	0.203	0.241
β_{02}^{sp}	0.077	0.056	0.169	0.186	0.678
β_{03}^{rp}	0.703	0.053	0.000	0.177	0.000
β_{03}^{sp}	0.655	0.047	0.000	0.146	0.000
β_{04}^{rp}	1.631	0.045	0.000	0.162	0.000
β_{04}^{sp}	1.068	0.041	0.000	0.137	0.000
β_{Age}^{rp}	0.071	0.004	0.000	0.005	0.000
β_{Age}^{sp}	0.065	0.005	0.000	0.006	0.000
β_{Income}^{rp}	-0.024	0.003	0.000	0.003	0.000
β_{Income}^{sp}	-0.029	0.002	0.000	0.002	0.000
Mean of $\ln\left(\beta_{price}^{rp}\right)$	-3.620	0.032	0.000	0.050	0.000
Mean of $\ln\left(\beta_{price}^{sp}\right)$	-3.762	0.043	0.000	0.076	0.000
Std. Dev. of $\ln\left(\beta_{price}^{rp}\right)$	0.871	0.016	0.000	0.023	0.000
Std. Dev. of $\ln\left(\beta_{price}^{sp}\right)$	0.984	0.021	0.000	0.035	0.000
$\sigma_{\varphi_0^{rp}}$	2.980	0.080	0.000	0.163	0.000
$\sigma_{\varphi_0^{sp}}$	3.275	0.085	0.000	0.119	0.000
Log likelihood	-58790.29				

(c) Results for Model with RP/SP Correlation

Parameter	Estimate	Asymptotic Error Est.		Robust Error Est.	
		Std. err.	P_value	Std. err.	P_value
β_{00}^{rp}	8.458	0.118	0.000	0.215	0.000
β_{00}^{sp}	7.805	0.099	0.000	0.193	0.000
β_{01}^{rp}	0.273	0.111	0.014	0.198	0.167
β_{01}^{sp}	-0.295	0.108	0.006	0.288	0.307
β_{02}^{rp}	0.192	0.095	0.043	0.150	0.200
β_{02}^{sp}	0.060	0.095	0.531	0.255	0.815
β_{03}^{rp}	1.205	0.091	0.000	0.135	0.000
β_{03}^{sp}	0.601	0.087	0.000	0.192	0.002
β_{04}^{rp}	2.067	0.091	0.000	0.157	0.000
β_{04}^{sp}	1.006	0.086	0.000	0.239	0.000
β_{Age}^{rp}	0.053	0.002	0.000	0.004	0.000
β_{Age}^{sp}	0.051	0.003	0.000	0.007	0.000
β_{Income}^{rp}	-0.016	0.001	0.000	0.002	0.000
β_{Income}^{sp}	-0.021	0.001	0.000	0.002	0.000
Mean of $\ln\left(\beta_{price}^{rp}\right)$	-3.576	0.024	0.000	0.049	0.000
Mean of $\ln\left(\beta_{price}^{sp}\right)$	-3.648	0.031	0.000	0.086	0.000
Std. Dev. of $\ln\left(\beta_{price}^{rp}\right)$	0.627	0.016	0.000	0.028	0.000
Std. Dev. of $\ln\left(\beta_{price}^{sp}\right)$	0.796	0.015	0.000	0.036	0.000
$\sigma_{\varphi_0^{rp}}$	1.139	0.052	0.000	0.111	0.000
$\sigma_{\varphi_0^{sp}}$	1.404	0.068	0.000	0.235	0.000
σ_{ξ}	2.295	0.026	0.000	0.045	0.000
Log likelihood	-53857.33				

Appendix 4.C ALTERNATIVE ESTIMATION RESULTS FOR MODELS

(a) Mixed Logit Estimation (Basic Model)

Parameter	Estimate	Asymptotic Error Est.		Robust Error Est.	
		Std. err.	P_value	Std. err.	P_value
β_{00}^{rp}	7.746	0.133	0.000	0.227	0.000
β_{00}^{sp}	7.394	0.133	0.000	0.211	0.000
β_{01}^{rp}	-0.136	0.068	0.048	0.320	0.672
β_{01}^{sp}	-0.246	0.062	0.000	0.271	0.364
β_{02}^{rp}	0.207	0.062	0.001	0.205	0.314
β_{02}^{sp}	0.077	0.056	0.169	0.186	0.678
β_{03}^{rp}	0.693	0.054	0.000	0.177	0.000
β_{03}^{sp}	0.655	0.047	0.000	0.146	0.000
β_{04}^{rp}	1.616	0.045	0.000	0.161	0.000
β_{04}^{sp}	1.068	0.041	0.000	0.137	0.000
β_{Age}^{rp}	0.059	0.004	0.000	0.005	0.000
β_{Age}^{sp}	0.065	0.005	0.000	0.006	0.000
β_{Income}^{rp}	-0.029	0.003	0.000	0.004	0.000
β_{Income}^{sp}	-0.029	0.002	0.000	0.002	0.000
Mean of $\ln\left(\beta_{price}^{rp}\right)$	-3.666	0.036	0.000	0.062	0.000
Mean of $\ln\left(\beta_{price}^{sp}\right)$	-3.762	0.043	0.000	0.076	0.000
Std. Dev. of $\ln\left(\beta_{price}^{rp}\right)$	0.884	0.017	0.000	0.031	0.000
Std. Dev. of $\ln\left(\beta_{price}^{sp}\right)$	0.984	0.021	0.000	0.035	0.000
$\sigma_{\varphi_0^{rp}}$	3.073	0.084	0.000	0.134	0.000
$\sigma_{\varphi_0^{sp}}$	3.275	0.085	0.000	0.119	0.000
Log likelihood	-58791.27				

(b) Results for Model with RP/SP Correlation

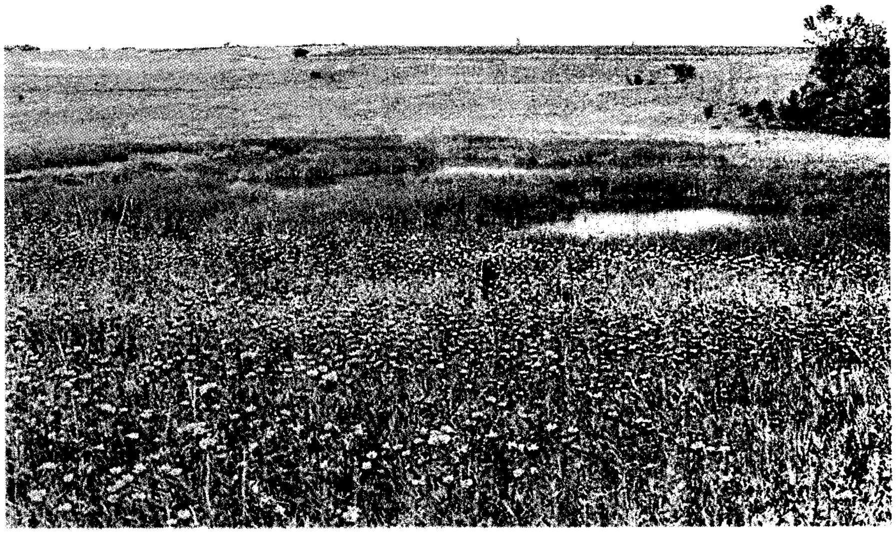
Parameter	Estimate	Asymptotic Error Est.		Robust Error Est.	
		Std. err.	P_value	Std. err.	P_value
β_{00}^{rp}	8.194	0.118	0.000	0.223	0.000
β_{00}^{sp}	8.386	0.122	0.000	0.226	0.000
β_{01}^{rp}	0.303	0.112	0.007	0.219	0.166
β_{01}^{sp}	-0.255	0.136	0.062	0.495	0.607
β_{02}^{rp}	0.192	0.091	0.034	0.141	0.171
β_{02}^{sp}	0.164	0.136	0.230	0.389	0.674
β_{03}^{rp}	1.148	0.088	0.000	0.119	0.000
β_{03}^{sp}	0.579	0.155	0.000	0.585	0.322
β_{04}^{rp}	1.998	0.090	0.000	0.155	0.000
β_{04}^{sp}	0.996	0.146	0.000	0.528	0.059
β_{Age}^{rp}	0.052	0.002	0.000	0.004	0.000
β_{Age}^{sp}	0.075	0.004	0.000	0.014	0.000
β_{Income}^{rp}	-0.016	0.001	0.000	0.002	0.000
β_{Income}^{sp}	-0.027	0.003	0.000	0.015	0.066
Mean of $\ln\left(\beta_{price}^{rp}\right)$	-3.601	0.024	0.000	0.053	0.000
Mean of $\ln\left(\beta_{price}^{sp}\right)$	-3.532	0.042	0.000	0.180	0.000
Std. Dev. of $\ln\left(\beta_{price}^{rp}\right)$	0.628	0.016	0.000	0.032	0.000
Std. Dev. of $\ln\left(\beta_{price}^{sp}\right)$	0.764	0.021	0.000	0.100	0.000
$\sigma_{\varphi_0^{rp}}$	1.120	0.072	0.000	0.224	0.000
$\sigma_{\varphi_0^{sp}}$	2.631	0.109	0.000	0.429	0.000
σ_{ξ}	2.189	0.028	0.000	0.055	0.000
Log likelihood	-53891.00				

(c) Results for Hypothesis Tests

TEST 1	H_0:	β_{Age}, β_{Incom}, Mean and Std. Dev.of $\ln(\beta_{price})$. are equal for RP and SP		
	D.F.	4		
		Error Comp. Method	Wald Stat.	P_value
	Model 1.	Asymptotic	23.03	0.000
		Robust	12.88	0.011
	Model 2.	Asymptotic	91.69	0.000
		Robust	15.91	0.000
TEST 2	H_0:	β_{Age}, β_{Incom} are equal for RP and SP		
	D.F.	2		
		Error Comp. Method	Wald Stat.	P_value
	Model 1.	Asymptotic	0.92	0.631
		Robust	0.56	0.754
	Model 2.	Asymptotic	32.79	0.000
		Robust	2.41	0.300
TEST 3	H_0:	Mean and Std. Dev.of $\ln(\beta_{price})$. are equal for RP and SP:		
	D.F.	2		
		Error Comp. Method	Wald Stat.	P_value
	Model 1.	Asymptotic	19.08	0.000
		Robust	8.20	0.017
	Model 2.	Asymptotic	85.67	0.000
		Robust	14.70	0.001

Appendix. Iowa Wetlands Survey

Iowa Wetlands Survey



Lowell Washburn

IOWA STATE UNIVERSITY

In order to make intelligent decisions concerning the future of wetland areas in Iowa, it is important to understand the benefits and costs associated with wetlands. The answers you give to the questions in this survey are very important in this process. Please try to answer each of the questions below. When an arrow follows the answer you select, please continue to the second part of the question.

What are wetlands?

Wetlands are transition areas between dry land and open waters. While this sounds like a simple enough idea, where one draws the line between a wetland and dry land is not always clear. Wetlands are not always wet, changing over time with the seasons and with changing weather patterns. Most scientists, in fact, define wetlands not only in terms of the amount of standing water, but also in terms of the types of soil and plants found in the region. One commonly used definition of wetlands describes them as

"...low areas where water stands or flows continuously or periodically. Usually wetlands contain plant-life characteristic of such areas. Water-saturated soils in these low areas are normally without oxygen and are described as anaerobic. Anaerobic soils and the presence of one or more members of a small group of plants able to tolerate and grow in such soils are universal features of all wetlands."[1]

Some of the plants found in wetlands include duckweed, water lilies, cattails, pondweed, reeds, sedges, and bulrushes.

In Iowa, two of the most common types of wetlands are the prairie pothole and riverine wetlands. Prairie pothole wetlands are typically found in the northcentral region of the state and are characterized by depressions in the land, mostly less than two feet deep, that are filled with water at least part of the year. Riverine wetlands refer to areas of marshy land that are near rivers and streams. Other names for these areas are marshes, sloughs, side channels, floodplains, backwaters, and old oxbows.

When you answer the questions we pose in this survey, we want you to think of wetlands as including both prairie pothole wetlands and riverine wetlands. This includes the following types of areas: floodplains, streams and creeks, lowlands, ponds and marshes. We **do not** want you to include the large lakes themselves or the main flow of major rivers (e.g., the Mississippi, the Missouri, the Des Moines River, etc.), but we **do** want you to include the uplands in the vicinity of lakes and rivers.

---

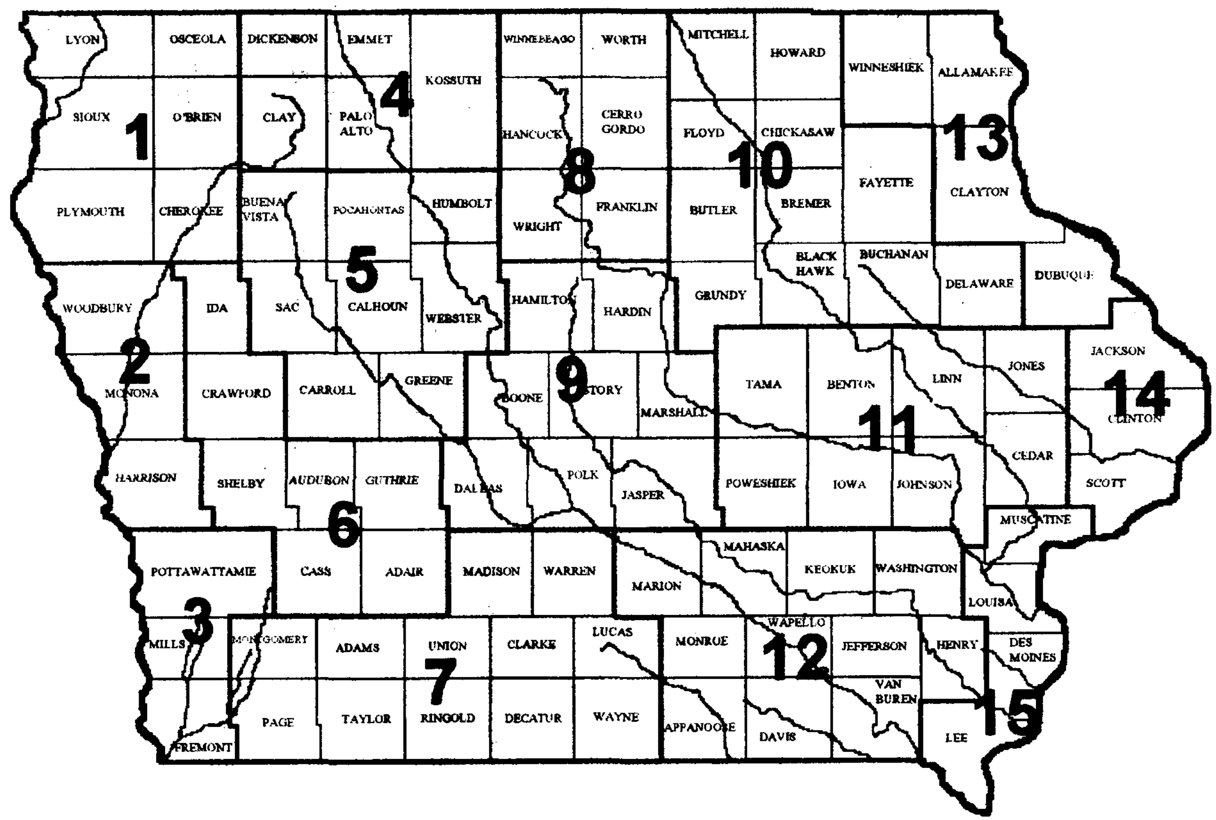
[1] Bishop, R. A., and A. van der Valk. 1982. *Wetlands*. In Cooper, T., *Iowa's Natural Heritage*. Iowa Natural Heritage Foundation and Iowa Academy of Science, Des Moines, pp. 208-29.

In this section, we would like to ask you about visits you and/or your family may have made to wetland areas for any reason during the past year. Please keep in mind the above description of wetlands.

1. On the opposite page is a map of Iowa, with the state divided into 15 areas (outlined in red). Please complete the following table. It is important that you report the number of trips you made to each area to the best of your memory. If you did not visit an area, you can simply leave that line blank.

Wetland Area	**Number of trips to wetlands in this area in 1997**	Also, please indicate the activities that you and/or your household typically engaged in while visiting wetlands in these areas (check all that apply)					
		Upland hunting	waterfowl hunting	Biking or hiking	fishing	wildlife viewing	other
1		❑	❑	❑	❑	❑	❑
2		❑	❑	❑	❑	❑	❑
3		❑	❑	❑	❑	❑	❑
4		❑	❑	❑	❑	❑	❑
5		❑	❑	❑	❑	❑	❑
6		❑	❑	❑	❑	❑	❑
7		❑	❑	❑	❑	❑	❑
8		❑	❑	❑	❑	❑	❑
9		❑	❑	❑	❑	❑	❑
10		❑	❑	❑	❑	❑	❑
11		❑	❑	❑	❑	❑	❑
12		❑	❑	❑	❑	❑	❑
13		❑	❑	❑	❑	❑	❑
14		❑	❑	❑	❑	❑	❑
15		❑	❑	❑	❑	❑	❑

2. Please mark the map below with an "X" indicating the specific location within the county of your most recent visit to a wetland.



147

*If you **did not** visit any wetland areas in 1997, please check here ❑, skip the next five questions, and proceed to question 8.*

3a. How many of the visits that you reported in question #1 were to areas within 5 miles of your home? ______________

3b. If you visited a wetland in your own wetland area (area # 10), how far did you typically travel to reach it? ________ miles one-way

4. During your typical visit to a wetland, how long do you stay? (Please choose only one)

❑ Under 1 hour
❑ 1 to 2 hours
❑ 2 to 4 hours
❑ 4 to 8 hours
❑ The entire day
❑ More than one day

*If you **did not** visit wetland areas 9, 10, or 11 in 1997, please skip the next section and proceed to question 8.*

In question #1, you indicated how much you visited various wetlands in Iowa. Next, we will be asking you questions to help us understand the economic value of all of your recreation trips to wetlands in Iowa this year. Depending on your particular situation, the dollar amounts written below may seem high or low. Regardless, please answer the question as carefully as you can, as your answer will help us represent a wide range of views.

5 Consider all of the recreation trips you made to wetlands areas #9, 10, and 11 in Iowa in 1997. Suppose that the **total cost per trip of each of your trips** to these areas had been $15 more (for example, suppose landowners charged a fee of this amount to use their land or that public areas charged this amount as an access fee). Would you have taken **any** recreation trips to the areas 9, 10 or 11 in 1997?

❑ No → If no, please skip to question 6b.
❑ Yes

6a. With this **additional** cost of $15 per trip of visiting areas 9, 10, 11, would this affect the number of trips you made to any of the 15 areas identified in question #1.

- ❑ No → If no, please proceed to question 8.
- ❑ Yes → If yes, how many **fewer** trips would you have taken to areas 9, 10 or 11 in 1997?

Area 9 ________ Area 10 ________ Area 11 ________

6b. With this **additional** cost of $15 per trip of visiting areas 9, 10, 11, would you have taken any additional trips to the remaining areas (whose costs have not changed)?

- ❑ No
- ❑ Yes→ If yes, how many **more** trips would you have taken to the following areas in 1997?

Area 1	________	Area 5	________	Area 12	________
Area 2	________	Area 6	________	Area 13	________
Area 3	________	Area 7	________	Area 14	________
Area 4	________	Area 8	________	Area 15	________

7. If you were no longer willing to visit areas 9, 10 or 11, please tell us why (Please check only the single most important reason):

- ❑ I cannot afford to pay the higher trip cost
- ❑ It's not worth the extra money
- ❑ It is wrong for landowners or public agencies to charge for access to land for recreational use
- ❑ The question is unclear or inappropriate
- ❑ Other: ______________________________

In this section, we would like to ask you some questions concerning what you may have read or known about wetlands ***before*** *receiving this survey. Our goal is to better understand the general public's knowledge about and attitudes towards wetlands. Please complete this section of the survey* ***before*** *proceeding onto later sections of the survey.*

8. What benefits, if any, do you associate with wetlands? (Please check all that apply)

- ☐ flood control
- ☐ wildlife habitat
- ☐ water quality purification
- ☐ recreation
- ☐ aesthetic enjoyment
- ☐ maintaining fisheries
- ☐ groundwater recharge
- ☐ protection of plant and animal biodiversity
- ☐ stabilizing shorelines and helping to prevent streambank erosion
- ☐ other:____
- ☐ don't know

9. What drawbacks, if any, do you associate with wetlands? (Please check all that apply)

- ☐ difficult to farm
- ☐ crop losses
- ☐ unproductive lands
- ☐ obstacle to development
- ☐ disease
- ☐ mosquitoes
- ☐ other:____
- ☐ don't know

10. When you visit wetland areas in Iowa, generally how important is each of the following when deciding where to go?

	Not Important	Somewhat Important	Important	Very Important
Ease of Access	❑	❑	❑	❑
Size of Wetland Area	❑	❑	❑	❑
Lack of Congestion	❑	❑	❑	❑
Variety of Wildlife	❑	❑	❑	❑
Public (*not private*) land ownership	❑	❑	❑	❑
Likely Hunting Success	❑	❑	❑	❑
Likely Fishing Success	❑	❑	❑	❑
Bird Viewing Opportunities	❑	❑	❑	❑
Water Quality	❑	❑	❑	❑
Facilities (e.g., picnic areas, playgrounds, restrooms, etc.)	❑	❑	❑	❑

11. Which of the following do you believe best describes what has been happening to the number of acres of wetlands in Iowa over the past ten years?

❑ declining ❑ stable ❑ increasing ❑ don't know

12. Do you support or oppose efforts to protect and/or restore wetlands in Iowa?

- ❑ Strongly support
- ❑ Somewhat support
- ❑ Indifferent
- ❑ Somewhat oppose
- ❑ Strongly oppose
- ❑ no opinion

13. There are a variety of programs currently being used to help restore and/or protect wetlands. How do you feel about each of the following programs?

	Strongly Support	Somewhat Support	Indifferent	Somewhat Oppose	Strongly Oppose
Outright public purchase of wetlands areas from willing sellers	❑	❑	❑	❑	❑
Private efforts to purchase and restore wetlands, including efforts by such groups as Ducks Unlimited, Pheasants Forever, and The Nature Conservancy	❑	❑	❑	❑	❑
Federal restoration of wetlands, with federal leasing of wetlands (CRP) or long term easements (WRP) to keep the lands out of crop production	❑	❑	❑	❑	❑
State and federal regulations prohibiting the further draining and conversion of wetlands to other uses	❑	❑	❑	❑	❑
Tying federal farm support funds to compliance with wetland protection	❑	❑	❑	❑	❑

14. To protect and/or restore wetlands often costs money. How do you think wetland conservation efforts should be paid for? (Please check all that you think apply)

- ❑ voluntary donations
- ❑ redistribute state revenues
- ❑ increase state taxes
- ❑ increase local taxes
- ❑ user fees
- ❑ increase fishing/hunting license fees
- ❑ private restoration efforts
- ❑ federal taxes
- ❑ lottery revenues
- ❑ other:______________________
- ❑ don't know

15. Who do you think should be primarily responsible for protecting wetlands in Iowa? (Please check only one)

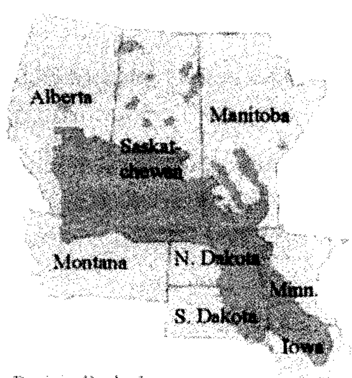
- ❑ federal government
- ❑ state government
- ❑ county government
- ❑ municipalities
- ❑ private conservation groups
- ❑ private landowners
- ❑ everyone
- ❑ other:________________
- ❑ don't know

In this section, we want to focus your attention on the prairie pothole wetlands and possible changes to the extent of these wetlands. Please do not go back to change your responses to earlier questions once you have read this section.

Ducks Unlimited

Prairie potholes are not always wet, changing in shape and size during the course of a year

As we mentioned earlier, prairie potholes are one of the major types of wetlands found here in Iowa. This kind of wetland consists of natural depressions in the landscape that are filled with water for at least part of the year and may range in size from a fraction of an acre to over 500 acres. In fact, as the picture above indicates, this type of wetland need not always be wet, but will often change in shape and size from year to year due to local flood or drought conditions.



Prairie Pothole Region of North America

The prairie potholes of Iowa are part of a larger collection of these wetlands in the United States and Canada known as the Prairie Pothole Region. The larger region, and the portion of Iowa that is contained in it, is highlighted on the map on the left. Although once quite numerous, the prairie pothole region has lost over half of its original wetland acreage and Iowa specifically has lost over 98% of its pothole acreage.

Prairie pothole wetlands provide a wide variety of benefits to both the local and regional environment:

- **Wildlife.** One of the benefits of prairie potholes is the habitat they provide for a variety of waterfowl and other species. About 175 species of birds breed in the prairie pothole region. Of this total, about 20 species of waterfowl breed in this region and 70% of the continent's waterfowl population is produced there. Up to half of the bird species in the region depends upon wetlands at some time in their lives.

Scott Nelson/Ducks Unlimited

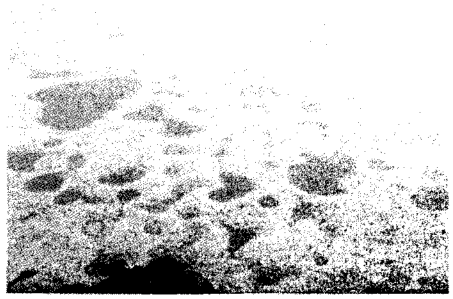
- **Water Quality.** Prairie potholes help to improve the water quality of local lakes and rivers by absorbing excess nutrients and chemicals that runoff from both farms and communities in the region.

- **Flood Control.** All types of wetlands, including prairie potholes, help by temporarily storing large quantities of water. This in turn reduces the severity of floods when they occur. In addition, by providing an area for storing excess rainfall, prairie potholes reduce water runoff from land, which in turn helps to control soil erosion.

- **Recreation.** Finally, one of the most obvious benefits provided by wetlands is the recreational opportunities they provide to hunters, anglers, hikers, bird watchers, and other wildlife and outdoor enthusiasts.

Some of these benefits are local (e.g., recreational opportunities, wildlife habitat, flood control, improved water quality) and some provide benefits to the entire region and elsewhere (e.g., preservation of endangered species, nesting grounds for migratory waterfowl, reduced soil erosion that would affect lakes and rivers elsewhere, flood reductions that occur elsewhere).

The dramatic declines in Iowa's prairie pothole wetlands have stopped and some wetlands have recently been restored. In 1986, the North American Waterfowl Management Plan was developed. As part of this plan, the Prairie Pothole Joint Venture was developed. In Iowa, about 27,000 acres have been placed under public protection.



U.S. Fish & Wildlife Service

The Prairie Pothole Joint Venture program has restored wetlands both by purchasing land outright from willing sellers and by developing a variety of easements where landowners retain the ownership of these lands, but agree to restore the land to its original prairie pothole wetland state. When the land is purchased and put under state or local control, the land is public and can be used by recreationists. When easements are used, the land remains private and can only be accessed with the landowner's permission.

In response to these programs, as well as recent increases in annual rainfall, populations of many species of birds and plants have shown notable increases. Waterfowl populations, which had hit their low during the mid-1980's, are now recovering. Populations of mallard and blue-winged teal ducks have shown promising increases.

Although biologists do not know exactly how populations of birds and other species will respond as more wetlands are reclaimed, it is likely that these gains will be maintained or even improved upon. Likewise, it is expected that significant additional gains in flood control and water quality will occur if more wetland acres are reclaimed.

As part of the Prairie Pothole Joint Venture, there is a goal for Iowa to acquire a total of 40,000 acres of land at a rate of about 2,000-3,000 acres per year for the next 15 years. These lands would be purchased from willing sellers, restored and held as public wetland reserves. Previous land has been acquired at a cost of about $940/acre.

One objective of this survey is to determine how valuable the Prairie Pothole Wetland Restoration Project is to Iowans. In the next question, we will be asking you about how much you would be willing to contribute to such a project. While you will not actually be contributing to the program at this time, we would like you to respond as if you were pledging to contribute to the project. In particular, please keep in mind any limits your budget would place on such contributions, as well as what you would have to give up to contribute.

16. Would you be willing to contribute an additional $100 on a one time basis (payable in annual installments of $20 over five years) to an Iowa Prairie Pothole Management trust fund? This fund would be used to acquire about 2500 acres of land annually for the next 15 years from willing landowners that would then be restored to prairie potholes.

    - ❑ Yes
    - ❑ No

17. To help us better understand your answers, please indicate the single most important reason for your response to the preceding question:

    - ❑ In general, the plan is <u>not</u> a good use for my money
    - ❑ In general, the plan is a good use for my money
    - ❑ I cannot afford to contribute to the program
    - ❑ The land acquisition plan is not realistic
    - ❑ The land acquisitions should be paid for by the government, not by individuals
    - ❑ I already contribute as much as I can afford
    - ❑ The question is unclear
    - ❑ Other: ____________________________

Information on you and other members of your household will help us better understand how household characteristics affect individuals' use of wetlands and their attitudes towards changes to them. It will also help us to determine how representative our sample is of the state of Iowa. All of your answers are strictly confidential. The information will only be used to report comparisons among groups of people. We will never identify individuals or households with their responses. Please be as complete as possible. Thank you.

18. Are you

☐ male ☐ female

19. How many years have you lived in the state of Iowa?

20. What is your age?

☐ Under 18
☐ 18-25
☐ 26-34
☐ 35-49
☐ 50-59
☐ 60-75
☐ 76+

21. What is the **highest** level of schooling that you have completed? (Please check only one)

☐ eight years or less
☐ some high school or less
☐ high school graduate
☐ some college or trade/vocational school
☐ two years of college or trade/vocational school
☐ college graduate
☐ some graduate school
☐ advanced degree

22. How many adults live in your household (over the age of 18)?______

23. How many children live in your household (18 or under)?_______

24. Please check the appropriate boxes if you or someone in your household has held any of the following licenses during the past 3 years:

- ❑ Iowa fishing license
- ❑ Iowa hunting license
- ❑ Iowa Duck Stamp

25 Do you own more than 40 acres of land in Iowa?

- ❑ No
- ❑ Yes ⇨ Are there any wetlands on your land?
  - ❑ No ❑ Yes

26. If you are currently employed, how many hours a week do you typically work? _____

27. If you are currently employed, do you have the option of working additional hours to increase your total income?

- ❑ Yes, If yes, what would your hourly wage be? $__________ per hour
- ❑ no

28. If you are currently employed, how much paid vacation do you receive per year? _______ days or ________weeks

29. How much "free time" do you typically have in a week? By "free-time" we mean time not spent on household chores, work, or other personal obligations.

Weekdays:	❑ 0 to 2 hours/day	❑ 2 to 4 hours/day	❑ 4 to 6 hours/day	❑ over 6 hours/day
Weekends:	❑ 0 to 3 hours/day	❑ 3 to 6 hours/day	❑ 6 to 9 hours/day	❑ over 9 hours/day

30. What was your total household income (before taxes) in 1997?

❑ under $10,000	❑ $40,000-$49,999
❑ $10,000-$14,999	❑ $50,000-$59,999
❑ $15,000-$19,999	❑ $60,000-$74,999
❑ $20,000-$24,999	❑ $75,000-$99,999
❑ $25,000-$29,999	❑ $100,000-$124,999
❑ $30,000-$34,999	❑ $125,000-$149,999
❑ $35,000-$39,999	❑ over $150,000

31. Approximately what percentage of your total household income did you spend last year on all of your leisure activities? (For example: movies, vacations, ball games, recreation trips, cable TV, dining out, etc.)

❑ 0 to 5%	❑ 5 to 10%	❑ 10 to 15%
❑ 15 to 20%	❑ over 20%	

Thank you for your time and effort in completing this survey. Please place the survey in the return envelope included with the survey and mail it. ***Do not*** *put your name anywhere on the survey or the return envelope. If for some reason the return envelope is missing, please send the survey to:*

JOSEPH A. HERRIGES
DEPARTMENT OF ECONOMICS
MAILSTOP C195
IOWA STATE UNIVERSITY
AMES, IA 50011-1070

References

[1] Adamowicz, W., Louviere, J., and Williams, M. "Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities," *Journal of Environmental Economics and Management* 10, 271-92. (1994)

[2] Adamowicz, W., Swait, J., Boxall, P., Louviere, J., Williams, M. "Perceptions Versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation," *Journal of Environmental Economics and Management* 32, 65-84. (1997)

[3] Aigner, D. J. and Leamer, E. E. "Estimation of Time-Of-Use Pricing Response in the Absence of Experimental Data," *Journal of Econometrics* 26, 205-27. (1984)

[4] Atkinson, S. E.; Crocker, T. D.; Shogren, J. F., "Bayesian Exchangeability, Benefit Transfer, and Research Efficiency," *Water Resources Research* 28(3), 715-22. (1992)

[5] Azevedo, C. Linking Revealed and Stated Preference Data in recreation Demand Modeling, Ph.D. Dissertation, Department of Economics, Iowa State University. (1999)

[6] Azevedo, C.; Herriges, J. and Kling, C., "Combining Revealed and Stated Preferences: Consistency Tests and Their Interpretations," *American Journal of Agricultural Economics,* forthcoming 2002.

[7] Ben-Akiva, M.; Bolduc, D. and Walker, J. "Specification, Identification, & Estimation of the Logit Kernel (or Continuous Mixed Logit) Model," Working paper, MIT. (2001)

[8] Bergstrom, J. C. and Civita, P. D. " Status of Benefits Transfer in the United States and Canada: A Review," *Canadian Journal of Agricultural Economics* 47, 79-87. (1999)

[9] Bhat, C., "Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model," working paper, Department of Civil Engineering, University of Texas, Austin. (1999)

[10] Bishop, R.; Heberlin, T. "Measuring Values of Extra Market Goods: Are Indirect Measures Biased?," *American Journal of Agricultural Economics* 61, 926-30. (1979)

[11] Bishop, R.; Heberlin, T.; and Kealy, M. "Contingent Valuation of Environmental Assets: Comparisons with a Simulated Market," *Natural Resources Journal* 23, 619-33. (1983)

[12] Boyle, K. J. and Bergstrom, J. C. "Benefit Transfer Studies: Myths, Pragmatism, and Idealism," *Water Resources Research*, 28(3), 657-63. (1992)

[13] Brookshire, D. S. and Neill, H. R. "Benefit Transfers: Conceptual and Empirical Issues," *Water Resources Research*, 28(3), 651-55. (1992)

[14] Cameron, T. A. "Combining Contingent Valuation and Travel Cost Data for the Valuation of Non-market Goods," *Land Economics* 68, 302-17. (1992)

[15] Cameron, T., Poe, G., Ethier, R., and Schulze, W. "Alternative Non-market Value-elicitation Methods: Are the Underlying Preferences the Same?" *Journal of Environmental Economics and Management* 44(3), 391-425. (2002)

[16] Caves, D.W., Herriges, J. A., Train, K. E. and Windle, R. J. "A Bayesian Approach to Combining Conditional Demand and Engineering Models of Electricity Usage," *Review of Economics and Statistics* 69, 438-48. (1987)

[17] Chib, S., "Bayes Inference in the Tobit Censored Regression Model," *Journal of Econometrics* 51, 79-99. (1992)

[18] Cummings, R. G.; Brookshire, D. S.; and Schulze, W.D. (Eds), "Valuing Environmental goods: A State of the Arts Assessment of Contingent Method," Rowman and Allanheld, Totowa, NJ. (1986)

[19] Desvousges, W. H.; Naughton, M. C.; Parsons, G. R., "Benefit Transfer: Conceptual Problems in Estimating Water Quality Benefits Using Existing Studies," *Water Resources Research* 28(3), 675-83. (1992)

[20] Garrod, G.; Willis, K. G., *Economic valuation of the environment: Methods and case studies,* Cheltenham, U.K. and Northampton, Mass.: Elgar; distributed by American International Distribution Corporation, Williston, Vt. (1999)

[21] Gelfand, A. E., Hills, S. E., Racine-Poon, A., and Smith, A. "Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling," *Journal of the American Statistical Association* 85, 972-85.

[22] Gelman, A., Carlin, J., Stern, H. and Rubin, D. *Bayesian Data Analysis*, Chapma & Hall/CRC. (1995)

[23] Geweke, J. "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication"(with discussion and rejoinder), *Econometric Reviews* 18, 1-126. (1999)

[24] Herriges, Joseph and Kling, Catherine "Updating Prior Methods for Non-Market Valuation: A Bayesian Approach to Combining Disparate Sources of Environmental Values," Project Proposal to the National Science Foundation, Department of Economics, Iowa State University. (1996)

[25] Herriges, Joseph and Phaneuf, Daniel, " Inducing Patterns of Correlation and Substitution in Repeated Logit Models of Recreation Demand," *American Journal of Agricultural Economics* 84(4), 1076-1090. (2002)

[26] Huang, J., Haab, T. and Whitehead, J. "Willingness to pay for Quality Improvements: Should Revealed and Stated Preference Data be Combined?" Working Paper, Department of Economics, East Carolina University. (1997)

[27] Jeffreys, H., *Theory of probability* (3rd ed.), Oxford, U.K.: Oxford University Press. (1961)

[28] Kask, S. B.; Shogren, J. F., "Benefit Transfer Protocol for Long-Term Health Risk Valuation: A Case of Surface Water Contamination," *Water Resources Research* 30(10), 2813-2823. (1994)

[29] Kass, R. E.; Raftery, A. E., "Bayes Factors," *Journal of the American Statistical Association*, Vol. 90, No. 430, 773-95. (1995)

[30] Kirchhoff, S.; Colby, B. G.; LaFrance, J. T., "Evaluating the Performance of Benefit Transfer: An Empirical Inquiry," *Journal of Environmental Economics and Management* 33(1), 75-93. (1997)

[31] Lindley, D. V. and Smith, A.F.M. "Bayes Estimation for the Linear Model," *Journal of the Royal Statistical Society B* 34, 1-41. (1972)

[32] Loomis, J. B. "The Evolution of a More Rigorous Approach to Benefit Transfer: Benefit Function Transfer," *Water Resources Research* 28(3), 701-05. (1992)

[33] Loomis, J. B.; Roach, B.; Ward, F.; Ready, R. "Testing Transferability of Recreation Demand Models Across Regions: A Study of Crops of Engineer Reservoirs," *Water Resources Research* 31(3), 721–30. (1995)

[34] McFadden, D. "Modeling the Choice of Residential Location," *Spatial Interaction Theory and Planning Models*, A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull (eds), New York, North Holland. (1978)

[35] McFadden, D. and Train, K., "Mixed MNL Models for Discrete Response," *Journal of Applied Econometrics* 15, 447-70. (2000)

[36] McConnell, K. E. "Model Building and Judgment: Implications for Benefit Transfers With Travel Cost Models," *Water Resources Research* 28(3), 695-700. (1992)

[37] Mitshell, R. C. and Carson, R. T., " Using surveys to Value Public Goods: The Contingent Valuation Method," Johns Hopkins Univ. Press for Resources for the Future, Baltimore, MD. (1989)

[38] Morey, E. R.; Rowe, R. D.; and Watson, M. "A Repeated Nested-Logit Model of Atlantic Salmon Fishing," *American Journal of Agricultural Economics* 75(3), 578-92. (1993)

[39] Morikawa, T. *Incorporating Stated preference Data in Travel Demand Analysis.* Ph.D. Dissertation, Department of Civil Engineering, M.I.T. (1989)

[40] Parsons, G. and Kealy, M. "Benefits Transfer in a Random Utility Model of Recreation," *Water Resources Research* 30(8), 2477-84. (1994)

[41] Revelt, D. and Train, K., "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level," *Review of Economics and Statistics* 80(4), 647-57. (1998)

[42] Rosenberger, R. S. and Loomis, J. B. "Using Meta-Analysis for Benefit Transfer: In-Sample Convergent Validity Tests of an Outdoor Recreation Database," *Water Resources Research* 36(4), 1097-07. (2000)

[43] Ruud, P., "Approximation and Simulation of the Multinomial Probit Model: An Analysis of Covariance Matrix Estimation," working paper,Department of Economics, University of California, Berkeley. (1996)

[44] Sellar, C.; Stoll, J.; and Chavas, J. "Validation of Empirical measures of Welfare Change: A Comparison of Nonmarket Techniques," *Land Economics* 61, 156-75. (1985)

[45] Smith, K. V. "On Separating Defensible Benefit Transfers From 'Smoke and Mirrors'," *Water Resources Research* 28(3), 685-94. (1992)

[46] Smith, V. K. and Kaoru, Y. "Signals or Noise? Explaining the Variation in Recreation Benefit Estimates," *American journal of Agricultural Economics* 1990, 419-33. (1990)

[47] Swait, J. and Louviere, J. "The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models," *Journal of Marketing research* 30, 305-14. (1993)

[48] Tierney, L.; and Kadane, J. B., "Accurate Approximations for Posterior Moments and Marginal Densities," *Journal of the American Statistical Association* 81, 82-86. (1986)

[49] Tierney, L.; Kass, R. B.; and Kadane, J. B., "Fully Exponential Laplace Approximations to Expectations and Variances of Nonpositive Functions," *Journal of the American Statistical Association* 84, 710-16. (1989)

[50] Train, K. E., Qualitative Choice Analysis, MIT Press. (1986)

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

[52] Train, K., "Halton Sequences for Mixed Logit," working paper, Department of Economics, University of California, Berkeley. (1999)

[53] Train, K. E., Discrete Choice Models with Simulation, Cambridge University Press. (2002)

[54] Wakefield, J., Smith, A., Racine-Poon, A., and Gelfand, A. E., " Bayesian Analysis of Linear and Non-linear Population Models by Using the Gibbs Sampler, " *Applied Statistics* 43, 201-21. (1994)

[55] Walsh, R. G.; Johnson, D. M.; McKean, J. R. "Benefit Transfer of Outdoor Recreation Demand Studies, 1968-1988," *Water Resources Research* 28(3), 707-13. (1992)

[56] Whitehead, J.; Haab, T., Huang, J. "Measuring Recreation Benefits of Quality Improvements with Revealed and Stated Behavior Data, " *Resource and Energy Economics* 22, 339-54. (2000)