1999

Linking revealed and stated preference data in recreation demand modeling

Christopher Dean Azevedo
Iowa State University

Linking revealed and stated preference data in recreation demand modeling

by

Christopher Dean Azevedo

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Major Professors: Joseph A. Herriges and Catherine L. Kling

Iowa State University

Ames, Iowa

1999

UMI Number: 9950075

UMI®

UMI Microform 9950075

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Graduate College

Iowa State University

This is to certify that the Doctoral dissertation of

Christopher Dean Azevedo

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Graduate College

TABLE OF CONTENTS

CHAPTER 1. LINKING CONTINGENT VALUATION AND TRAVEL COST DATA: INTRODUCTION AND BACKGROUND

A large percentage of the land in the state of Iowa is used for agricultural production. This is the result of favorable terrain as well as the high productivity of the soil. Much of this land was made available for farming by draining wetlands. As a result, an important political issue related to agriculture is what to do with those areas of land that remain as wetlands. Should additional wetlands be drained to expand agricultural production, or should the existing wetlands be preserved? Should marginal land that is currently in agricultural production be restored to wetlands? In short, what is the optimal level of wetlands in the state?

In order to answer these questions it is necessary to conduct cost-benefit analysis. The cost of wetland preservation is the forgone opportunity to use the land in its most productive capacity, usually agricultural production in Iowa. To calculate the benefit of wetland preservation one must determine the recreational value as well as the non-use value. Wetland recreational activities include hunting, fishing, wildlife viewing, bird watching, hiking, camping, and biking. An example of non-use value would be someone who simply values the existence of wetlands.[1]

The problem is that policymakers have very little information to work with in the case of wetlands, primarily because of their public-good nature. When a market exists for a good, traditional economic techniques may be utilized to determine the value society associates

[1] This value might exist because people intend to visit wetlands in the future, or they might simply value ecological diversity, even though they may never intend to visit a wetland.

with the good. When that good is not traded in a market, as in the case of wetland recreation, the economist must find some way to compensate for the lack of market signals.

In deriving the economic value of nonmarket goods, economists have largely relied on either indirect or direct methods. Indirect methods use information that is indirectly revealed to the analyst. The most popular indirect method is the travel cost method (TCM). In the travel cost method, travel costs serve as the price of access. With this data, the researcher can estimate either a continuous choice model or a discrete choice model.

In a continuous choice model, which estimates the quantity of trips taken to a site in a particular time period, a demand curve for the recreation site is estimated and used to conduct welfare analysis. In a discrete choice model, which is used to explain the choice between sites for a particular choice occasion, the consumer is assumed to choose from among alternatives within a set of available sites. The random utility framework, which will be described in the next chapter, is used to model this choice, and logit or probit methods are used in the estimation. In the literature this type of data is often referred to as "revealed" preference data because the economic agent's actions reveal information that can be used by the analyst.

Direct methods use information that is directly stated to the analyst, typically through the means of a survey. The contingent valuation method (CVM) directly elicits information about the consumer's value. A hypothetical scenario is presented to the consumer in which he is asked to state his willingness to pay or accept a compensation for a change in an environmental good or amenity. One example of a contingent valuation question is, "Would you be willing to pay $X in order to have access to recreation site A?" The consumer's answer places an upper bound on willingness to pay in the case of a "no", and a lower bound in the case of a "yes". This type of data is often referred to as "stated" preference data because the agent is making a statement about their preferences.

Another type of stated data is referred to as contingent behavior data. Instead of asking the respondent a question about their willingness to pay or accept, the respondent is asked about their behavior under a hypothetical scenario. For example, the respondent might be asked how many trips they would take if their travel cost were $X. The response to this question can be used to discover information about their willingness to pay or accept.

There are numerous advantages and problems associated with each method. One advantage of the indirect methods is that they are based on observable behavior. For example, when using the TCM both the number of trips taken and the distance the person has actually traveled are observed. Among the problems with the TCM are the valuation of time costs, how to define a "site" and its substitutes, and how to deal with multiple day trips, as well as measurement issues related to the price of access.

The most obvious potential problem with the direct method is that since a hypothetical scenario is being presented, respondents may provide a hypothetical answer that is inaccurate. Many researchers have called into question the reliability of these hypothetical responses. Diamond and Hausman (1994) discuss several criticisms of the use of contingent valuation, including what has been called "the embedding effect." The embedding effect is the name given to the tendency for respondents to report similar willingness to pay values across surveys when theory predicts that they should be very different. An example of the embedding effect would be if the respondent reported a willingness to pay to restore one wetland site roughly equal to their willingness to pay to restore ten wetland sites, including the one asked about individually.

Other researchers have considered different hypotheses about how respondents answer contingent valuation questions. Andreoni (1989) explored the idea that respondent may receive a "warm glow" from exhibiting support for good causes. Kahneman and Ritov (1993) consider the hypothesis that respondents use the money metric to express an attitude toward the public good in question, simply because the researcher asks them to use the money metric.

Another important difference between the two methods is that they may measure different types of values. The TCM measures use value. The price that is calculated is the price associated with use. In contrast, it is theoretically possible to use CVM models to obtain both use and non-use values. Comparisons across the two methods must be careful to insure that the same types of values (i.e., use versus non-use) are being measured.

Traditionally, researchers have approached a problem by designing a survey instrument to collect revealed data, stated data, and, in some cases, both types of data. The researcher then conducted welfare analysis using either the TCM or the CVM. The welfare estimates were then compared to those obtained from other methods. If the estimates were similar, then the second method was said to "validate" the first. If not, then the possible reasons for divergence were discussed. Recently, some researchers have attempted to link revealed preference and stated preference methods in order to take advantage of the strengths of each type of data. Examples of papers that have taken this approach are Cameron (1992), Larson (1990), and Huang, Haab, and Whitehead (1997).

This is a welcome shift in focus from an emphasis on pitting the methods against each other, to an emphasis on understanding the conditions that must exist for the methods to complement one another. A number of reasons have been proposed for linking the methods.

One reason is that you are adding information to the likelihood function, which will result in more precise parameter estimates. Another reason is that it is often the case that the researcher is interested in a scenario that has never existed. This means that there are no revealed data to work with, and stated data must be relied upon.

The problem of linking revealed and stated preference information will be the main focus of this research. Within this area, I will examine the three issues of conducting and interpreting various hypothesis tests, valuing travel time in the recreation demand model, and transferring benefit estimates from one geographical region to another.

Theoretically, both revealed and stated methods should elicit information about the same set of preferences, but in a different manner. Much of the new research has concentrated on measuring the degree of correlation between the estimates generated by the two models. The process of testing the "consistency" of the two data sets takes the form of hypothesis tests of identical parameter estimates between the models. Other research has focused on measuring the degree of bias in the parameter estimates.

Along these lines, various hypothesis tests can be conducted to investigate the external validity of each type of data. This would be a fairly straightforward procedure if one of the sources of data were known to be accurate. However, there is a large literature that discusses the various sources of error and bias that may exist with each source of data. Because of this, it is only possible to test one of the data sources by making the maintained hypothesis that the other source of data is "correct."[2] In this dissertation I will look directly

[2] Ideally, both sources of data would benefit from external validation.

at the issue of hypothesis testing with regard both to consistency between the two types of data as well as sources of bias within each type of data.

An important aspect of hypothesis testing not explored in the literature is how the results of the tests should be interpreted. On the surface this appears to be a rather benign issue, but it is actually a very interesting question. In addition to the many hypothesis tests that will be considered, I will also examine the alternative ways of interpreting the conclusions.

The second issue (i.e., how to value time) was briefly mentioned as a problem with the travel cost method. This is actually a much more general problem in recreation demand modeling. When calculating the travel cost for a respondent, one is confronted with the problem of how to value the respondent's travel time i.e., what is the opportunity cost of travel time for the respondent? The answer to this question depends on whether or not the respondent has the option to choose how many hours to work. If the respondent can choose between working and engaging in leisure activities, then the respondent's wage rate can give us some information about their opportunity cost of time.

In practice, most researchers simply choose a fraction of the wage rate at which they will value travel time. In his work with urban commuter data, Cesario (1976) suggested that it might be reasonable to value travel time at one-third the wage rate. Most researchers who wish to include travel time have adopted the approach of using a fixed fraction of the wage rate, with one-third being a popular choice. This is a rather ad-hoc approach, and alternatives have been proposed in the literature (McConnell and Strand (1981), Bockstael, Strand, and Hanemann (1987)).

From a more general perspective, policy makers are often interested in calculating benefit estimates for a previously unstudied recreational resource. The third issue of whether benefit estimates from a previously studied recreational resource can be transferred to an unstudied recreational resource has a long history. In 1981, President Reagan signed Executive Order 12291 requiring new regulations to be subject to benefit-cost analysis. In response, the Environmental Protection Agency formed guidelines to be followed in conducting benefit-cost analysis. They stated that because of limited research and time budgets, "off-the-shelf" studies could be used to serve as the basis for the benefit-cost analysis. This use of existing studies has come to be known as "benefit transfer."

As mentioned before, one main theme will characterize the research contained in this dissertation: linking revealed and stated preference data. Within this general topic I will explore three sub-themes: (1) conducting and interpreting specific hypothesis tests concerning consistency between revealed and stated preference data, as well as hypothesis concerning specific sources of bias in each of the data sets, (2) exploring benefit transfer, and (3) incorporating time into the model.

In particular, I will develop a model that can be used to link revealed and stated preference data. This model will allow for the linking of revealed preference (RP) data with two types of stated preference (SP) data, each of which differ in their information content. The models will account for the respondent's opportunity cost of time, and will be used to test RP/SP consistency, as well as several hypotheses concerning sources of bias in each type of data. Additionally, my data allows for examination of the possibility of benefit transfer between different geographical regions.

A brief outline of the dissertation is as follows. In Chapter 2, I will discuss the relevant research in the areas of linking revealed and stated preference data, modeling time in the recreation demand framework, and benefit transfer. Chapter 3 contains discussion of a survey that gathered revealed and stated preference data from a sample of Iowa residents. In Chapter 4, I will develop the model that will be used to link the two forms of data, discuss the results of applying the model using the Iowa data set, and discuss various hypothesis tests. Chapter 5 will be used to examine benefit transfer. Chapter 6 will contain an examination of the issue of incorporating time into the model. Finally, Chapter 7 will discuss the conclusions and implications of this research.

CHAPTER 2. LITERATURE REVIEW

The existing literature of linking revealed and stated preference data can be feasibly categorized in terms of the underlying RP and SP formats. The first approach links two data sets that have the same form, but are gathered from different data generating processes. These two data sets have the same information content, conditional on equal reliability of the two data sources. For example, the researcher might have information concerning past quantity of trips and a corresponding travel cost, and link it with information concerning stated quantity of trips at a proposed cost. In this case, both types of data take the form of quantity of trips taken at a corresponding cost. The first set of data is the revealed quantity of trips taken in the past, while the second set of data is the quantity of trips that the respondent states they will take in the future. Dickie, Fisher, and Gerking (1987) call this "pooling" of the data, a name that will be used throughout this dissertation to refer to this type of model.

The second approach to linking data is to combine two types of data that are not only gathered from different data generating processes, but are of a different form. In this case the two data sets have a different information content. An example of this might be linking revealed travel information with the responses to a hypothetical yes/no question. In this case continuous travel cost information is linked with dichotomous choice information. This is often a more complicated modeling process than simply pooling data and may involve significant changes to the likelihood function.

In this dissertation I will discuss linking RP and SP data using these two methods. When discussing these methods of linking data, I will refer to the first group as "pooling" models and the second group as "combining" models. While this terminology has the potential to be confusing, I will attempt to make it clear by the context of the discussion. Though this grouping is a convenient way to categorize the models, it does ignore the chronological progression of the literature. Cameron (1992), discussed in the combining section, was the paper that sparked the research into linking RP and SP data.

It should be pointed out that the most common type of RP data used in linking models is travel cost data. Though this is the most popular type, other types of revealed data can be used. For instance, we might have information concerning a respondent's past donations to a wetland preservation program. In that case our revealed information would consist of individual's responses to an open-ended request for donations.[3]

In this section I will first discuss the various papers dealing with linking stated and revealed preference data by pooling similar data. Papers that deal with linking different forms of preference data, combining models, will then be discussed. It will be useful to further subdivide the pooling papers into those papers that develop discrete pooling models and those that develop continuous pooling models.

Linking Revealed and Stated Preference Data

Pooling Models: Discrete

Within this category of models there are at least two approaches that can be taken. All of the papers in the literature develop models that utilize the random utility model (RUM)

[3] One of the reasons that travel cost data is so widely used is that other forms of revealed data are often hard or impossible to get. Another example of revealed information is actual voting behavior. For example, if voters were considering a tax increase for a wetland preservation program, then their yes/no vote would be a piece of revealed information. However, voting records, other than whether the person voted, are often not kept or are not accessible.

developed by McFadden (1978). A random utility model is developed by first specifying individual i's utility function for trips to site j

$$u_{ij} = v_{ij}(x_i) + \varepsilon_{ij} \tag{1}$$

where $v_{ij}(\cdot)$ is the nonstochastic portion of the utility function, x_i represents observable characteristics of respondent i, and ε_{ij} represents the stochastic portion of the utility function that is assumed to be an i.i.d. random variable with zero mean. This stochastic term contains components that are unobservable to the analyst, although individual i knows their utility function with certainty.

Individual i prefers site A to site B if the utility from going to site A is greater than the utility of going to site B (i.e., $u_{iA} > u_{iB}$). This relationship can be used to form an expression for the probability that respondent i will choose site A over site B. This expression takes the form

$$\Pr(A) = \Pr(v_{iA} + \varepsilon_{iA} > v_{iB} + \varepsilon_{iB}). \tag{2}$$

By defining $\eta_i = \varepsilon_{iB} - \varepsilon_{iA}$, the following statement can be derived

$$\Pr(A) = F_\eta(\Delta v_i) \tag{3}$$

where F_η is the cumulative distribution function of η and $\Delta v_i = v_{iA} - v_{iB}$. Δv is called the utility difference function. By specifying the functional form of v_{ij} as well as the distribution of ε_{ij}, the analyst can estimate the parameters of the utility function.

Conceptually, it is also possible to develop a discrete pooling model using the bid function approach, though I have found no papers that take this approach. The bid function approach, outlined in Cameron (1988), is very similar to the random utility model. Consider the stylized question, "Would you pay $t to obtain access to recreation at site z for a year?" The respondent will answer "yes" if their willingness to pay, represented by the bid function, is greater than $t, the amount they are being asked to pay. Cameron specifies the bid function as

$$Y_i = B_i(x_i) + \omega_i \tag{4}$$

where Y_i represents respondent i's willingness to pay, $B_i(\cdot)$ represents the nonstochastic portion of the bid function, x_i are explanatory variables related to individual i, and ω_i is the stochastic portion of the bid function. Willingness to pay is unobservable, but is manifested through the respondent's answer to the question above. The probability that the respondent answers "yes" to the question takes the form

$$\Pr(yes) = \Pr(Y_i > t_i) = \Pr(\omega_i > t_i - B_i(x_i)) \tag{5}$$

where t is indexed by i to indicate that the bid is varied over respondents. By dividing by the standard deviation of ω_i, σ_ω, the following statement is derived

$$\Pr(yes) = 1 - F_\omega\left(\frac{t_i - B_i(x_i)}{\sigma_\omega}\right). \tag{6}$$

As with equation (3), this equation can be used, with the assumed function form of $B_i(\cdot)$ and the assumed distribution of ω, to estimate the parameters of the bid function.[4] The following articles are primary examples of discrete choice pooling analysis.

[4] McConnell (1990) has pointed out that the random utility and bid function models are in fact dual to each other under the assumption of constant marginal utility of income.

Adamowicz, Louviere, and Williams (1994)

The model developed by Adamowicz, Louviere, and Williams (1994) links stated and revealed choices between alternative recreation sites in a multinomial logit model (MNL). A multinomial logit model is a random utility model, as described in equations (1) through (3), with gumbel-distributed errors, ε_{ij}. Both the stated and revealed choices are a function of travel distance and quality attributes of the site. The functional form of the utility function is specified, the difference-in-utility function is derived, and maximum likelihood is used to estimate the parameters of the utility function.

The paper makes a couple of important points. First, the authors note that pooling stated and revealed preference data may help with the problem of collinearity, a common problem with travel cost data. This can be achieved in the stated preference portion of the survey through an orthogonal effects design. Pooling a data set that contains orthogonal columns with one that exhibits collinearity will reduce the collinearity. Second, the use of stated preference data allows for the examination of attribute scenarios that are outside the experienced attribute range.

The model is applied using RP and SP data concerning water recreation. Hypothesis tests result in the failure to reject the null hypothesis of equal parameters between the RP and SP data sets, as well as rejection of the null hypothesis of equal variances. They identify this as evidence that the revealed and stated preference data contain similar preference structures.

Louviere (1996)

In this paper Louviere discusses the history of the use of MNL models to link RP and SP data. He focuses on five papers, each of which use a discrete MNL model to pool the two types of data and test for consistency of the RP and SP parameters (one of the papers discussed is Adamowicz, Louviere, and Williams (1994)). Four of the five papers discussed by Louviere find consistency, with the fifth paper finding mixed results. Louviere concludes that the evidence indicates that stated preference models are consistent with revealed preference models if differences in the variance of the two sources of data are accounted for.

From the point of view of this dissertation, another important conclusion of the Louviere paper is that "yes/no" binary response SP data appears to be among the most reliable SP measures.

The difference between the information content of the discrete and continuous stated preference information will be explored in this dissertation. Discrete stated preference questions are often easier for the respondent to answer than continuous questions, but they may differ from continuous questions both in terms of the amount of information embodied in the answer, as well as the quality of the information embodied in the answer.

Adamowicz, Swait, Boxall, Louviere, and Williams (1996)

Like the previous papers, Adamowicz *et al.* (1996) use a random utility model to link stated and revealed trip data. The innovation of this paper is that they also obtain data concerning the respondent's perceptions of certain quality measures. These subjective perceptions of quality are then used as explanatory variables in an $RP_{perceptions}$ model. An $RP_{objective}$ model is also developed using objective perceptions of quality as explanatory variables. Both models are pooled with an SP model and tested for consistency. For each of these pooling models they fail to reject the null hypothesis of parameter equality. Interestingly, they also fail to reject the null hypothesis of equal variance between the data sets. However, the consistency test for a third model, which pools all three data sets, results in the rejection of parameter equality.

Loomis (1997)

Loomis develops a random utility model, but unlike the papers discussed before, he assumes a normal distribution for the error term. He also chooses to add an unobservable characteristic specific to each individual. In this case, equation (1) takes the form

$$u_{ij} = v_{ij}(x) + \pi_i + \varepsilon_{ij} \tag{7}$$

where π_i represents the unobservable characteristic specific to individual i. This model, called the random effects probit model, is appropriate if data is available for more than two dichotomous choice responses.

Loomis tests consistency between the RP and SP data by applying the model using data gathered through a river recreation survey. The survey collected four dichotomous responses from each respondent. The tests of consistency between the revealed and stated data indicate that there are no significant differences between the two data sets.

McConnell, Weninger, and Strand (1999)

McConnell, Weninger, and Strand develop a random utility model that pools revealed travel cost data with stated contingent behavior and valuation data. A bivariate normal model is used to reflect both the revealed decision to take a trip and the stated decision to take a trip if travel costs were higher. This model allows them to explore the relationship between the revealed and stated preference errors.

McConnell *et al.* find mixed results concerning the consistency of revealed and stated preference data. While they do not specifically test the joint hypothesis of equality of RP and SP parameters, they do conclude that revealed and stated preference data are not consistent with each other. This conclusion is based on the presence of large standard errors for the parameter estimates in the model that combines the RP and SP data.

Brownstone, Bunch, and Train (1998)

Brownstone *et al.* (1998) develops a model that is very similar to the Loomis (1997) model. In this case, equation (1) takes the form

$$u_{ij} = v_{ij}(x) + \eta_{ij} + \varepsilon_{ij} \tag{8}$$

where η_{ij} is a random term with zero mean, whose distribution relies on the underlying parameters and observed data relating to alternative j for individual i. The distribution of ε is assumed to be i.i.d, gumbel. Brownstone *et al.* refers to this as a mixed logit model.

The mixed logit model is used to pool RP and SP data to model consumer demand for alternative-fuel vehicles. Models are estimated using the RP data alone, the SP data alone, and both data sets linked. Though no hypothesis tests of consistency between the two data sets are considered, the difference in the parameter estimates between the three models is cited as a justification for linking the RP and SP data. They state that by linking the two data sets the strengths of each type of data are captured, while the weaknesses of each are avoided.

Pooling Models: Continuous

The next group of pooling models uses a different approach than the discrete pooling models. Whereas the previously discussed papers modeled a single choice occasion, these papers estimate continuous models of recreation demand.

Dickie, Fisher, Gerking (1987)

Dickie, Fisher, and Gerking (1987) pools actual market transaction data with hypothetical market transaction data using a continuous demand model. Respondents from whom revealed market transaction data was to be collected were presented with an opportunity to purchase pints of strawberries at a given price. Respondents from whom stated market transaction data were to be collected were first told that they were engaging in a hypothetical experiment. They were then presented with a price and asked how many pints they would purchase.

This data was then used to estimate a linear demand relationship in a Tobit framework. Actual (RP), hypothetical (SP), and pooled models were estimated. The conclusion of the paper is that they cannot reject the hypothesis of structurally identical demand equations obtained by the two data collection methods.

Layman, Boyce, and Criddle (1996)

Layman, Boyce, and Criddle (1996) employs a model they call the hypothetical travel cost model. It links revealed travel cost data and hypothetical travel cost data. Respondents are presented with a hypothetical scenario and asked how many trips they would take under each of three site management scenarios. This information is then pooled with actual travel cost data and used to estimate a demand function whose independent variables are travel costs, sociodemographic variables, and shift variables for the hypothetical scenarios. Both forms of data are used to estimate the parameters.

The conclusion of the Layman, Boyce, and Criddle paper is that the actual and hypothetical travel cost information are compatible in the estimation of a trip demand function in that the travel cost variables affect actual and hypothetical trips in the same manner.

Combining Models

The common theme of the papers discussed above is the pooling of RP and SP data sets, where each contains the same information. Again, the reliability of the information contributed by each data set may not be the same. I will now discuss papers that link RP and SP data sets where each contains different information, and the same caveat about reliability of the two data sets is important. The first paper that will be examined is the pioneering work of Cameron (1992).

Cameron (1992)

One of the first papers to formally combine RP and SP data sets was Cameron (1992). The intuition behind the Cameron paper is that the choices made by economic agents, whether real or hypothetical, should reflect the same underlying preference structure. Cameron, therefore, constructs her model in such a way as to allow consistency between the parameter estimates across the models.

The Cameron model has two parts. The first is a basic contingent valuation model. Using the Hanemann (1984) approach, she defines a direct utility function

$$U(z,q) = U(Y - Mq, q) \tag{9}$$

where q is the current number of trips per year to the recreation site, z is a composite of all other goods and services, M is the respondent's typical travel costs, and Y is the respondent's income. This direct utility function is used to create a utility difference function

$$\Delta U(Y, M, T, q) = \max_{q} U(Y - Mq - T, q) - U(Y, 0) \tag{10}$$

where T is the access fee or tax.

The function ΔU is positive when the maximum utility achieved by taking trips, q, and paying tax, T, is greater than the utility achieved by taking no trips and avoiding the tax.

The assumption is made that the utility difference equation is composed of a systematic portion and a random, unobservable portion as follows

$$\Delta U_i = \Delta U_i^* + \varepsilon_i \tag{11}$$

where ΔU_i^* indicates the systematic portion of the utility difference and $\varepsilon_i \sim N(0, \sigma^2)$. Cameron uses a generic function, $f(x_i, \beta)$, to represent this systematic portion of the utility difference function, where x_i includes various terms in Y, T, M, and q, and β represents the parameters of the function. The utility function is assumed to be deterministic from the point of view of the respondent and stochastic from the perspective of the analyst.

If the utility difference is positive then the respondent gets more utility from taking trips q and paying the tax T than from taking no trips at all. In this case the indicator function, I_i, takes a value of one. Otherwise, it takes a value of zero. Cameron then derives a probabilistic statement in order to form the likelihood function, with

$$\Pr(yes) = \Pr(I_i = 1) = \Pr(\Delta U_i > 0) = \Pr(\varepsilon_i > -f(x_i, \beta)). \tag{12}$$

By dividing by σ, a standard normal variable with cumulative density function $\Phi(\cdot)$ is created, yeilding

$$\Pr(I_i = 1) = 1 - \Phi\left[\frac{f(x_i, \beta)}{\sigma}\right]. \tag{13}$$

The probability of observing the indicator variable equal to zero (i.e., respondent answers "no") is the compliment of the above probability, which leads to the following log-likelihood function

$$\log L = \sum_i \left\{(1 - I_i)\log\left(\Phi\left[\frac{f(x_i, \beta)}{\sigma}\right]\right) + I_i \log\left(1 - \left(\Phi\left[\frac{f(x_i, \beta)}{\sigma}\right]\right)\right)\right\}. \tag{14}$$

The second part of the Cameron model is the travel cost model. The basic problem faced by the economic agent is the maximization of the utility function (9). From this problem emerges the individual's demand function for recreational trips. Cameron specifies a demand function that is composed of a systematic portion and a random portion

$$q_i = g(x_i, \beta) + \eta_i, \tag{15}$$

where $\eta_i \sim N(0, \nu^2)$. The log-likelihood function associated with this demand function is

$$\log L = -\left(\frac{n}{2}\right)\log(2\pi) - n\log\nu - \left(\frac{1}{2}\right)\sum_i \left\{\frac{(q_i - g(x_i, \beta))}{\nu}\right\}^2 . \tag{16}$$

Estimating the two parts simultaneously imposes consistency between stated and revealed models. With the assumption of independent errors, the two log-likelihood functions are summed and the constraint is imposed that the beta coefficients are identical in the two components. It is straightforward to allow for correlated errors by introducing a correlation coefficient between the two error terms.

This is a straightforward method of combining two forms of data. An important limitation to the Cameron model is that she resorts to the use of an ad-hoc demand function. The error term used in the demand specification is not consistent with the error structure in the utility specification. In other words, maximization of the stochastic utility function (11) would result in a demand function with a stochastic structure different from (15).

Cameron's model goes beyond the pooling models by attacking a tougher problem: linking two types of information that have a different "form", or information content. This is a more general way to look at the problem, and is potentially much more useful than pooling models, as it can apply to data of two different types, not just identical forms of data.

Larson (1990)

Larson develops a model that is very similar to Cameron's, except that he deals directly with the issue of how to consistently incorporate the error terms. Larson develops the model from a unified utility-theoretic structure by beginning with an assumed demand specification and integrating back to get the implied utility function. A semilog demand function is first assumed[5]

$$x_i = \exp(\alpha + \beta p_i + \gamma q_i + \delta m_i + \sigma \varepsilon_i), \tag{17}$$

where x_i is trips taken by individual i to a fishing site, p_i is the travel cost of a trip, q_i is a vector of quality characteristics, m_i is income, and $\varepsilon_i \sim N(0,1)$ is the error term.

Roy's Identity can be used to derive the indirect utility function

$$v(p_i, q_i, \varepsilon_i, m_i) = -\left(\frac{1}{\delta}\right)\exp(-\delta m_i) - \left(\frac{1}{\beta}\right)\exp(\alpha + \beta p_i + \gamma q_i + \sigma \varepsilon_i), \tag{18}$$

which can then be used to derive explicit statements for willingness to pay

$$wtp_i = \left(\frac{1}{\delta}\right)\ln\left[1 - \left(\frac{\delta}{\beta}\right)x_i\right]. \tag{19}$$

This willingness to pay statement is then used to form a statement for the probability that the respondent will answer "yes" to the contingent valuation question in the same manner as equation (5). The probability of a "yes" response is equal to the probability that the respondent's willingness to pay is higher than the proposed bid. Note that this is an example of the bid function approach discussed earlier.

---

[5] Larson also presents derivations for linear and log-linear functional forms.

It is now possible to use the stated preference data with the probabilistic statements and the revealed data with the demand specification, much the same as Cameron (1992). The important issue is that though the Larson and Cameron methods are very similar, the Larson model treats the error term in a utility consistent manner.

Niklitschek and Leon (1996)

Niklitschek and Leon (1996) combines CVM response data with data on intended recreation trips using a model that is very similar to those developed by Cameron and Larson. However, in this case both forms of data are hypothetical. While this is not a case of linking stated and revealed data, it is a model that combines different forms of data, and is thus a minor variation on the approach used by Cameron and Larson, where one form of data is revealed.

The emphasis of the Niklitschek and Leon paper is the examination of the difference between use and non-use values. Their model illustrates the ability of the CVM and TCM models to estimate different values.

Huang, Haab, and Whitehead (1997)

Huang, Haab, and Whitehead estimate the bid function and recreation demand function using CVM data and revealed trip data respectively. Their tests for consistency between the two types of data lead them to conclude that stated and revealed data can be fruitfully combined, provided that the two decisions imply the same change in behavior induced by a quality change.

Cameron, Poe, Ethier, and Schulze (1999)

This paper develops a model that combines information from six separate samples of respondents, for each of which a different elicitation method is used. The methods were an actual purchase decision, a single-bid hypothetical dichotomous choice format, an open-ended format, a payment card format, a five-level categorical scale for willingness to pay, and a stated choice among a set of alternatives.

They specifically investigate the role of the error variance of each model and conclude that the valuation results from each model are consistent as long as error variances are permitted to differ between models. However, if preferences are allowed to be heterogeneous across sociodemographic groups, the hypothesis of consistency between models is rejected.

These papers characterize the relevant literature on linking revealed and stated preference data. As mentioned before, this will be the main focus of this dissertation. The issue of modeling time in the recreation demand framework will be discussed next, followed by benefit transfer.

Modeling Time

No matter how one chooses to model recreation demand, it is important to account for a respondent's travel time. When dealing with recreation demand models, the opportunity cost of time plays a potentially important role. For goods that are very time-intensive, as is the case with outdoor recreation, the valuation of travel time is likely to be very important. Bishop and Heberlein (1979) found that valuing travel time at half the wage rate, as opposed to not including it, resulted in a fourfold difference in consumer surplus estimates.

A majority of past authors who have incorporated time into the recreation demand model have chosen to model the opportunity cost of time as some fraction of the full wage rate. The reason this has been done is that the static labor-leisure model implies that labor and leisure are traded at a price equal to the wage rate. Cesario (1976), in a survey of empirical evidence concerning urban commuters, concluded that the opportunity cost of travel time was between one fourth and on half of the wage rate. Based on this evidence he concluded that it would reasonable to value travel time at one-third of the wage rate.

There are a couple of issues that are important to consider when modeling time in the recreation demand framework. The first is whether to include on-site time in the time cost calculation. In general, most authors rely on Wilman's (1980) conclusion that on-site time can be ignored if it can be assumed that the length of the visit does not change as the cost of travel changes.

Another issue is whether travel time has value as a commodity or whether it should be treated as any other scarce input. For example, if the trip to the wetland site takes you through a scenic area, then a portion of the time spent traveling could provide you with utility, and should be considered as a commodity. But if the time spent traveling provides you with no utility, then the time spent in travel should be considered as any other scarce resource. If the alternative use of this time is employment, then the opportunity cost of this time is the wage rate.

Time Papers

For the purposes of this dissertation, three papers will be discussed. The first paper, McConnell and Strand (1981), discusses the estimation of the fraction of the wage rate at which travel time should be valued. The second paper, Bockstael, Strand, and Hanemann (1987), generalizes the McConnell and Strand model by taking a careful look at the form of the time budget constraint. The final paper, Feather and Shaw (1999), also examines the time budget constraint, but takes a different approach from the other two papers.

McConnell and Strand (1981)

McConnell and Strand (1981) approaches the problem by first defining a general utility function, $U(r,x)$, where r is recreation trips and x is a composite bundle of all other goods. The budget constraint takes the following form

$$[F(w)+E](1-t)=px+cr \tag{20}$$

where t is the proportional income tax rate, w is time worked, $F(w)$ is income generated from working w hours, E is fixed income, p is the price of the composite bundle, and c is out-of-pocket costs for each recreational trip. Their time constraint takes the form $T=ar+w$, where T is the total time available and a is the amount of travel time per recreational trip. Maximization of the utility function with respect to the budget and time constraints results in a trip demand function of the form

$$r=f[c+a(1-t)F'(w)] \tag{21}$$

where $(1-t)F'(w)$ represents the marginal opportunity cost of time. They then show that the opportunity cost of time is less than average income, which is defined as $v=[F(w)+E]/w$. This implies that the trip demand function takes the form

$$r=f(c+\lambda va) \tag{22}$$

where $0<\lambda<1$ and λv now represents the marginal opportunity cost of time. The parameter λ represents the fraction that was being estimated by Cesario (1976). In practice, most researchers arbitrarily pick the value of λ, often using Cesario's estimates as a guide. McConnell and Strand (1981) adds λ as a parameter to be estimated.

This is a significant improvement over arbitrarily choosing the value of λ, but there are still shortcomings in the approach. The most obvious shortcoming is that a single λ is estimated for all respondents in the sample regardless of their ability to work extra hours, a problem dealt with in the paper by Bockstael, Strand, and Hanemann (1987). Another shortcoming is the simple specification for the time constraint. This shortcoming is also examined closely in the next paper.

Bockstael, Strand, and Hanemann (1987)

Bockstael, Strand, and Hanemann (1987) develop a model that is very similar to the McConnell and Strand (1981) model, except they take a closer look at the structure of the time constraint. They point out that the nature of an individual's labor supply decision determines whether her wage rate yields information about the marginal value of her time. It may not be possible for a respondent to optimally adjust the number of hours worked. If this is the case, they will be found at a "corner solution" where they choose either to not work, or to work a job with a fixed number of hours. The respondent may choose to work a part-time job with a flexible number of hours in addition to their job with fixed hours, or they may choose not to work at all. It is important to account for these decisions when modeling recreation demand.

Their model begins by specifying a general utility function

$$U(x_R, x_N) \tag{23}$$

where x_R represents the recreational good and x_N is a vector of other commodities. They assume that each recreational visit has a constant marginal cost, p_R, and fixed travel and on-site time requirements, t_R, with all other commodities subject to unit money or time costs. Utility function (23) is maximized subject to the following time and income constraints

$$\begin{aligned} &\overline{T} - t_F - t_D - t_R' x_R - t_N' x_N = 0 \\ &\overline{Y} + w_D t_D - p_R' x_R - p_N' x_N = 0 \end{aligned} \tag{24}$$

where $\overline{T}$ is total time available for discretionary activities, t_F is the number of hours spent working at a job with fixed work week, t_D is hours of discretionary work, t_N is a vector of the time requirements for commodities other than recreation, $\overline{Y}$ is a combination of nonwage income and income from nondiscretionary employment, w_D is the wage associated with t_D, and p_i is the price associated with x_i for $i = R, N$.

If $t_D > 0$ then the respondent is at an interior solution in the labor market, while if $t_D = 0$ the respondent is at a corner solution where they either work a fixed number of hours or choose not to work at all. For a given respondent at a corner solution, the recreation demand function takes the form

$$x_i = h^C\left(p_i, t_i, p^0, t^0, \overline{Y}, \overline{T}\right) \tag{25}$$

where p^0 and t^0 are vectors of money and time costs for goods other than i. If the respondent is at an interior solution, the demand function takes the form

$$x_i = h^I\left(p_i + w_D t_i, p^0 + w_D t^0, \overline{Y} + w_D \overline{T}\right). \tag{26}$$

It is clear from the form of these demand functions that some extra data is required in order to incorporate the respondent's labor market situation. However, the extra data requirements are not prohibitive. In gathering the trip data typically used in travel cost models, respondents must be asked their total work time, whether or not they have discretion to work instead of recreating, and if so, what their discretionary wage would be.[6]

Feather and Shaw (1999)

The model developed in this paper is a generalization of a model developed by Heckman (1974). Heckman's model treated all respondents as having the freedom to choose the number of hours worked. Feather and Shaw (1999) generalize that model by also considering the situation where respondents are faced with a take-it or leave-it decision of working a job with fixed hours. Feather and Shaw classify respondents into four groups: (1) respondents who can freely alter their work hours, (2) respondents who work a fixed number of hours, but desire to work more (underemployed), (3) respondents who work a fixed number of hours, but desire to work less (overemployed), and (4) respondents who choose not to work.

They then use the utility maximization framework to define an expression that describes the monetary value of leisure time for the respondent. This value, called the shadow wage, takes the form

$$W^{*} = k\left(h, wh + A, P_1, \ldots, P_n\right) \tag{27}$$

where W^{*} represents the shadow wage, $k(\cdot)$ is the shadow wage function with continuous first partial derivatives, h is the number of hours worked at wage rate w, A is nonlabor income, and $P_1, \ldots, P_n$ are the prices of market goods i through n. It is also assumed that W^{*} and h are positively related, i.e., the marginal value of leisure time increases as more hours are worked.

---

[6] These extra data requirements do not rely on the functional specification of the demand function.

The relationship between the respondent's shadow wage and the market wage determines the respondent's employment choice. If they can freely choose working hours, then the market wage equals the shadow wage. If the respondent is underemployed, then the respondent's shadow wage is lower than the market wage, which is less than or equal to the shadow wage that the respondent would achieve if they spent no hours in leisure activities. If the respondent is overemployed, then the respondent's shadow wage is higher than the market wage, which is higher than the shadow wage the respondent would achieve if they worked zero hours. If the respondent chooses not to work, then their shadow wage is higher than the market wage.

These relationships, along with parametric specifications for equation (27) and the wage

$$w = g(Z) \tag{28}$$

where Z is a matrix of exogenous variables, allow for estimation of the respondent's shadow wage rate. These shadow wages are then used as the opportunity cost of travel time in a standard random utility model. This approach is different from that used in the first two papers. Instead of estimating the value of λ within the random utility model, the value of λ is estimated for each respondent in a separate routine. These estimates are then used as data in the random utility model.

These papers illustrate the type of work that has been done towards correctly incorporating time into recreation demand models. Though there are other papers that consider the valuation of time in recreation demand models, these papers characterize the type of research that is relevant to this dissertation. Next, attention is turned to the issue of benefit transfer.

Benefit Transfer

Policy makers are often confronted with the problem of how to value a recreational resource under stringent research and time constraints. Though numerous valuation papers have been written in the past two decades, only a small percent of the recreational resources have been studied. If no past research has been conducted on the resource in question, then the policy maker is left with the problem of deciding on a method of calculating a benefit measure.

The practice of adjusting data from existing studies (study sites) to make them applicable to an unstudied recreational resource (policy site) has been termed benefit transfer. Because the resulting benefit estimates can only be as accurate as the initial benefit estimates, the benefit transfer process is less than ideal. However, the relative low cost and smaller time requirements of benefit transfer methods often make them the preferred choice of policy makers. A number of authors have contributed important research in this area.

Benefit Transfer Papers

Smith and Kaoru (1990)

Smith and Kaoru (1990) uses a meta analysis model to examine estimates of benefits users derive from environmental resources. Meta analysis is a method used in psychology, education, and the health sciences to synthesize the results of numerous controlled experiments. Though controlled experiments are rare in the field of economics, the methodology is still useful.

In practice, a meta model uses the benefit estimate as the dependent variable, explaining that estimate using variables relating to the characteristics of the site being valued as well as the theoretical issues involved in the modeling process. These theoretical issues often involve the judgement of the analyst. The authors point out that ideally, the judgment of the analyst would not be important in explaining the benefit estimates. However, their evidence indicates that often it is.

Smith and Kaoru define consumer surplus per unit of use as the consumer surplus measure divided by the quantity of trips taken, CS/v, where v represents the number of trips. Their estimating model takes the following form

$$\left(\frac{CS}{v}\right)_{Ei} = \beta\alpha_0 + \beta\alpha_S X_{Si} + \beta\alpha_A X_{Ai} + \gamma Z_i + \varepsilon_i \qquad (29)$$

where $\left(\frac{CS}{v}\right)_{Ei}$ is the i'th estimated consumer surplus per trip, X_{Si} contains variables pertaining to the features of the site, X_{Ai} contains variables pertaining to the recreational activities undertaken at the site, Z_i contains variables pertaining to the modeling decisions of the researcher, including behavioral model assumptions, assumptions about the specification of the demand function, and assumptions about the econometric estimators, and ε_i is a stochastic error.

Smith and Kaoru reviewed approximately 200 published and unpublished articles, as well as masters' and Ph.D. theses from 1970 to 1986. Of those, 77 studies both utilized a travel cost model, and contained enough information with which to calculate consumer surplus per unit of use. The authors conclude that meta analysis shows promise in evaluating the effects of modeling assumptions on estimates of values of nonmarket recreation resources.

Loomis (1992)

Loomis develops a model that allows him to test benefit transfer using a study site in one state and a policy site in another. He points out that rather than simply transferring an average net willingness to pay measure from one site to another, a more unbiased approach is to transfer the entire demand equation. This allows the researcher to incorporate characteristics of the new site into the benefit calculation.

Loomis also discusses the issue of convergent validity, or the equality of demand parameters between the study site and the policy site. Identically specified demand models are estimated for two separate regions. If the null hypothesis of identical demand parameters between the two regions cannot be rejected, this would be evidence in favor of convergent validity of benefit transfer.

A zonal travel cost model is specified as

$$\frac{T_{ij}}{POP_i} = B_0 - B_1 TC_{ij} + B_2 TIME_{ij} + B_3 SUBS_{ik} + B_4 INC_i + B_5 QUAL_j \qquad (30)$$

where T_{ij} is trips from origin i to site j, POP_i is the population of origin i, TC_{ij} is the travel cost of origin i to visit site j, $TIME_{ij}$ is the travel time of origin i to visit site j, $SUBS_{ik}$ is a measure of the cost and quality of substitute site k to origin i, INC_i is the average income of origin i, and $QUAL_j$ is the recreation quality at site j. Loomis' chooses a zonal travel cost model, but other benefit transfer research has been conducted using individual observation models.

Convergent validity requires that the parameters at the study site, the B's, be equal to the parameters at the policy site, call them A's. The specific form of the test is as follows

H_0: $B_0 = A_0$ and $B_1 = A_1$ and $B_2 = A_2$ and $B_3 = A_3$ and ····· and $B_n = A_n$

H_A: At least one equality does not hold.

A Chow test is conducted using sport-fishing data first from Oregon and Washington, and then using data from Oregon and Idaho. In both cases the null hypothesis of parameter equality between regions is rejected. This result may be due to a couple of characteristics of the data sets. First, the data sets were collected at different time periods. Angler use levels changed dramatically between the collection of the data sets. Second, the data sets were collected for purposes other than demand estimation. Both of these factors may help explain the results of the hypothesis tests. Despite these factors, these results are taken as evidence against trans-state benefit transfer.

Loomis then considers benefit transfer for different sites within the State of Oregon. A simulation experiment was conducted to explore the effects of using data for $n-1$ rivers to estimate the demand parameters for the n'th river. The percent difference between total recreation benefits estimated with the full model versus the total recreation benefits estimated with the $n-1$ model were calculated. The percent differences ranged from 0.93% to 17.58%. The percent difference was also calculated for another benefit measure, average benefits per trip, and ranged from 3.51% to 39.07%.

Loomis concludes that the results of these tests are not favorable for trans-state benefit transfer, and that in-state benefit transfer works better for total recreation benefits than for average benefits per trip. He suggests further testing of trans-state benefit transfer.

Conclusion

In this chapter I have taken a look at the representative papers in the areas of linking revealed and stated preference information, incorporating time into a recreation demand model, and benefit transfer.

The papers written on the subject of linking RP and SP data have several strengths. Probably the most important is that the papers exhibit a move from using RP and SP as competing methodologies toward using the methods to compliment each other. This allows for an examination of the consistency between revealed and stated data, and facilitates the exploration of possible biases in each methodology. Along this line, one weakness of the papers is that they do not consider specific hypothesis concerning the sources of bias, something that I will explore in this dissertation.

The papers written on the subject of incorporating time into recreation demand models have made strides toward a more realistic framework. In this dissertation I will utilize the McConnell and Strand (1981) model to examine the effects on RP and SP consistency of accurately incorporating time. I will also examine an extension to the their approach.

The papers on benefit transfer deal with the practical side of the benefits estimation process. Policy makers often face constraints that limit their ability to conduct a new study for the site in question. In this case, benefit transfer is often an attractive alternative. The hypothesis tests considered by Loomis (1992) seem to be a very appealing method of investigation convergent validity. This research will consider similar tests to those proposed by Loomis.

CHAPTER 3. THE DATA SET

The data that will be used in this dissertation was gathered as part of a large Iowa Wetlands Survey conducted in 1997 and funded by the U.S. Environmental Protection Agency.[7] Iowa contains two major types of wetlands: prairie pothole and riverine. Prairie pothole wetlands are found in the north central region of the state. Potholes are the result of glacial activity and are characterized by depressions in the land, most of which are less than two feet deep, that are filled with water for at least part of the year. Riverine wetlands are found across the state and are characterized by areas of marshy land near rivers and streams.

A survey instrument was designed to elicit travel cost information, contingent behavior information, contingent valuation information, and socioeconomic information from Iowa residents concerning their use of Iowa wetlands. Construction of the data set proceeded in five steps: survey design, sampling, pre-test, survey refinement, and final survey mailing.

Survey Design

Two main versions of the survey were written. The first was the Iowa River Corridor Project (IRCP) and the second was the Prairie Pothole (PP). The surveys were identical except for the contingent valuation scenario that was presented, which is what gives each type of survey its name. General information on wetlands was gathered from many sources. In order to balance the objectives of presenting the respondent with enough information to make informed decisions, while making the survey simple enough that they would not

[7] This research was supported in part by the U.S. Environmental Protection Agency and by the Western Regional Research Project W-133. Although the research described in this dissertation has been funded in part by the United States Environmental Protection Agency through R82-3362-010, it has not been subject to the Agency's required peer review policy and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

immediately throw it away, the project was discussed with numerous groups and individuals in the wetland field. They gave valuable comments and suggestions that were used to develop an initial version of the survey. A copy of each version of the final survey is included in Appendix 1. Although this is the final survey, the differences between the initial, or pre-test, survey and the final survey are minor. Summary statistics for the data used in this dissertation are shown in Table A1.1 of Appendix one.

The initial version contained a general discussion of wetlands, which included a definition and examples of wetlands for the respondent to keep in mind when they answered the survey questions. The basic definition that survey respondents were provided with was that wetlands could be described 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.

Respondents were told to consider both prairie pothole and riverine wetlands when answering the survey questions. In order to give the respondent some examples to think about, they were told to think of wetlands as including the following types of areas: floodplains, streams and creeks, lowlands, ponds and marshes. They were told not 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 to include the uplands in the vicinity of lakes and rivers.

When answering the survey questions, respondents were told to refer to a map of Iowa shown on page five of the survey. The state was divided into 15 zones based, roughly, on the Iowa crop reporting districts. These 15 zones were further grouped into five megazones, each containing three zones. Zones 1, 2, and 3 make up the 1,2,3 megazone, zones 4, 5, and 8 make up the 4,5,8 megazone, zones 6, 7, and 12 make up the 6,7,12 megazone, zones 9, 10, and 11 make up the 9,10,11 megazone, and zones 13, 14, and 15 make up the 13,14,15 megazone.

Survey question number one was designed to elicit visitation behavior for the past year. They were asked to indicate the number of trips they had taken to each zone over the past year, as well as the activities they engaged in during these trips. It is important to note that only single-day trips taken to wetland areas are of interest. Multi-day trips were intentionally excluded so that the value of the trip to the wetland site could be isolated. When multi-day trips are included, the value of the wetland visit is confounded with the value of the activities engaged in on the other days of the trip.

The respondent was then asked questions concerning the trips they made to the megazone containing their zone of residence. For example, respondents residing in zone 1 were asked, "Consider all of the recreation trips you made to wetlands areas #1, 2, and 3 in Iowa in 1997. Suppose that the **total cost per trip of each of your trips** to these areas had been $B more (for example, suppose that 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 1, 2, or 3 in 1997?" The bid B was varied across respondents according to a bid design described later.

They were asked to choose either "yes" or "no". They were then asked to elaborate on how their behavior would have changed with the increased cost. The next two questions (6a and 6b) elicited information on how many fewer trips they would have taken to their region of residence, as well as how many more trips they would have taken to the other regions.

It is important, for the purposes of this dissertation, to note the difference between the information gathered in the first few questions of the survey. Question number one gathered information about the respondent's travel behavior over the past year. This information takes the form of continuous revealed preference data. Question number five asks for a yes/no response to the question "would you take any trips at the higher price?" This information takes the form of discrete stated preference data. Finally, question number six elaborates on question number five by asking the respondent to state how many fewer trips they would have taken at the higher price. This information takes the form of continuous stated preference data.

All three types of data will be utilized in the linking models. In particular, two types of models will be considered: a model that pools the continuous revealed preference data with the continuous stated preference data (RP/SPc), and a model that combines the continuous revealed preference data with the discrete stated preference data (RP/SPd).

The next section of the survey, questions eight through fifteen, asked respondents about their attitudes toward wetlands and wetland preservation. To this point all surveys were alike, with the exception of variations in the bid, B. The Prairie Pothole and Iowa River surveys differed in the contingent valuation scenario presented to the respondent. In the contingent valuation section the respondent was presented with a scenario and asked about the choices they would make under this scenario.[8] For the IRCP surveys, the scenario

---

[8] The data gathered from the contingent valuation portion of the survey will not be used in this dissertation.

described a plan to purchase 7000 acres of land in the Iowa River corridor. Respondents were then asked, "Would you be willing to contribute $C on a one time basis (payable in annual installments of $C/5 per year over five years) to an Iowa River Corridor Wetlands Management Trust fund that would cover the cost of acquiring this acreage?"

For the PP surveys, the scenario described a program called the Prairie Pothole Joint Venture. The respondent was asked, "Would you be willing to contribute an additional $D on a one time basis (payable in annual installments of $D/5 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." Prior to each of these questions the respondent was reminded to keep in mind any limits their budget might place on their contribution.

Finally, the respondent was asked a series of socioeconomic questions concerning characteristics such as gender, age, income, free time, and money spent on recreation activities. Included in this group of questions were three questions designed to give us information concerning the respondent's ability to work extra hours instead of engaging in leisure activities, as well as information concerning the respondent's marginal wage. This data was designed to allow us to incorporate time into the models developed in the next chapter.

An initial version of the survey was presented to four focus groups. One focus group contained members of Ducks Unlimited and Pheasants Forever, one contained parents of Crawford Elementary School children, one contained Iowa State University students, and the final one contained members of the Bethesda Lutheran Church. Each of these groups gave us useful comments that were used to create the pre-test draft of the survey. The most useful information received from the focus groups concerned the definition of the wetland, as well as the definition of a "trip" to a wetland. Members of the focus groups were paid $10 for their participation.

Sampling

Two samples of Iowa residents were drawn: a sample from the general population of the state and a sample of Iowa hunting/fishing license holders. Our intention was to gather information from both users and non-users of wetland. In order to make sure that users were adequately represented, the portion of the sample was gathered from license holders. Table 3.1 shows the breakdown of the total sample of 6800 names and addresses.

Table 3.1: Distribution of Iowa Wetlands Sample

Total Sample	6800
General Population Sample	4300
99 County	3300
4 County	1000
Hunters/Fishers Sample	2500
Counties	1661
Directs	839

Survey Sarpling Inc., a professional sampling firm, drew the general population sample. This sample was drawn from phone records, and was composed of two subsamples. The first was a random sample of 3300 names and addresses drawn from the 99 county population of all Iowans. The second was a random sample of 1000 names and addresses drawn from Tama, Benton, Poweshiek, and Iowa counties, the four counties surrounding the Iowa River Corridor. This four county portion was drawn in order to allow us to focus on the region around the Iowa River. This data concerned the contingent valuation question, which will not be discussed in this dissertation, but a discussion can be found in Crooker (1998).

In order to obtain the sample of hunting/fishing license holders, permission was obtained from the Iowa Department of Natural Resources to examine the 1996 hunting/fishing license records. These records were divided into two groups: direct sales, which are licenses sold by retail stores such as Wal-Mart, and county sales, which are licenses sold by the counties themselves. Of all licenses sold by the State of Iowa in 1996, 1/3 of the licenses were sold as directs, with 2/3 being sold by counties. Therefore, of the 2500 licenses sampled, 1/3 were sampled from the directs and 2/3 were sampled from the counties.

In order to draw the 839 permits from the directs, 12 boxes were randomly selected from the total number of boxes of direct licenses. Each box contained bundles of permits. In order to sample from a box, the bundles were arranged end to end in a line. The length of the line was measured and five points along the line randomly chosen. From each point, every third permit was pulled. The name and address of the license holder was recorded.

To draw the 1661 permits from the county portion, there were 36 boxes of licenses. The percent of county permits sold in each crop-reporting district was calculated, and the 36 boxes were allocated according to those percentages. The percentage of total sales in each district accounted for by each county was then calculated, and the county's number based on that percentage.[9]

[9] It was not possible to sample according to the percentage sold at each retail outlet (direct) because the permits in the retail boxes were randomly grouped.

Bid Design

The bid design for the contingent behavior question on the final survey was identical to the bid design for the pre-test survey. The data obtained from the pre-test indicated that the bid distribution was adequate. Table 3.2 shows the distribution of bids in both the pre-test and final surveys.

Table 3.2: Bid Distribution

Bid	Prairie Pothole	Iowa River
$5	800	400
$10	800	400
$15	800	400
$20	400	200
$30	400	200
$40	400	200
$50	400	200

Pre-test

In order to test the survey on a sample of Iowa residents, a pre-test was conducted consisting of 600 surveys. The first survey mailing took place in October of 1997.[10] Survey recipients were sent a package that included a cover letter, survey, payment claim form, and return envelope. Respondents were assured that their responses would be completely confidential.[11]

Two weeks after the initial mailing, a reminder postcard was sent to survey recipients who had not yet returned the survey. Two weeks after the postcard, another survey was sent

[10] It may be the case that mailing the survey so close to Christmas resulted in a lower than expected response rate. This timing was unavoidable for the pre-test, but corrected in the final survey mailing.
[11] Respondents were paid four dollars for a completed survey. Payment was made by Iowa State University check.

to recipients who still had not returned the survey. The distribution of surveys sent is shown in Table 3.3.

Of the 400 PP surveys sent, 39 were returned by the post office as undeliverable. Respondents returned 174 surveys, for a 48.2% response rate among deliverable surveys. Of the 200 IRCP surveys sent, 24 were returned as undeliverable. Respondents returned 99 surveys, which is a 56.3% response rate among deliverable surveys.

Table 3.3: Pre-test Survey Distribution

	Prairie Pothole	Iowa River
General Population	240	60
Four County Area	0	100
Hunter/Anglers	160	40
Total	400	200

Final Survey Mailing

As the surveys were returned, it became obvious that there were several minor problems in the survey design that needed to be addressed. One of the main problems was that a significant number of respondents did not provide a complete answer to question number one, the question that elicited information concerning their past travel behavior. Many respondents apparently were not aware that they were supposed to indicate the number of trips they had taken to each area, instead marking only activities. Because this revealed preference data is a very important part of the research, this problem needed to be corrected. To address the problem, the column provided to enter the number of trips was highlighted in order to attract attention.

The other changes that were made concerned the 'free time' and 'leisure expenditures' questions at the end of the survey. The question that asked about the amount of 'free time' respondents typically have in a week was changed from an open-ended response to a categorical response. The question that asked how much money their household typically spends per month on all leisure activities was also changed from an open-ended response to a categorical response question. The most disturbing problem was the low response rate. In order to boost the response rate it was decided that the incentive would be increased from four dollars to ten dollars.

The final survey was mailed in February of 1998. This meant that the survey would not be competing with the holiday season, which may have been a factor in the low response rate of the pre-test. The same mailing procedure was followed for the final survey as for the pre-test, viz. the initial survey was sent, those who had not returned the survey after two weeks were sent a reminder card, and those who still had not returned the survey after four weeks were sent another survey. The distribution of surveys sent is shown in Table 3.4.

Of the 4000 PP surveys sent, 443 were returned by the post office as undeliverable. 2094 surveys were returned, which is a 58.9% response rate among deliverable surveys. Of

Table 3.4: Final Survey Distribution

	Prairie Pothole	Iowa River
General Population	2400	600
Four County Area	0	1000
Hunter/Anglers	1600	400
Total	4000	2000

the 2000 IRCP surveys sent, 151 were returned as undeliverable. 1045 surveys were returned by respondents, which is a 56.5% response rate among deliverable surveys.

The response rate for the PP survey, 58.9%, was 10.7 points higher than the pre-test PP response rate. For the IRCP, the final response rate of 56.5% was 0.2 points higher than the pre-test response rate of 56.3%.

Two individuals coded the data into two separate Access databases. Each person had their own database into which they coded the entire data set. This allowed comparisons between the data sets to ensure accurate coding. Comparisons between the data sets revealed that both were accurately entered, with one being slightly better than the other. The most accurate of the two was used.

The revealed and stated preference data gathered in this survey will be applied using a model developed in the next chapter. Both an RP/SPc and an RP/SPd model will be developed. These models will be used to test hypotheses concerning the consistency of the revealed and stated preference data, as well as various sources of bias within each type of data.

CHAPTER 4. LINKING REVEALED AND STATED PREFERENCE DATA: THE MODEL

In this chapter, I develop a model designed to analyze household responses to questions one through six of the survey discussed in Chapter 3. A model will be developed to link the revealed preference data with both the continuous and discrete stated preference data. This model will be applied using data from the Iowa Wetlands survey, and used to explore the difference in the information content of the continuous SP data and the discrete SP data. Finally, the results and interpretation of various hypothesis tests will be discussed.

Development of the Model: Revealed and Stated Preference

Suppose there is a single recreational site to visit, site i, and respondent $n's$ demand function for that site takes the form

$$q_{in} = \alpha + \beta P_{in} + \gamma M_n + \sigma \varepsilon_{in} \qquad (31)$$

where q_{in} is the quantity of recreation trips demanded per year to site i by individual n, P_{in} is the travel cost to individual n of visiting site i, M_n is the income of individual n, and $\varepsilon_{in} \sim N(0,1)$. The error term can arise from a variety of sources, including variations in preferences across individuals in the population as well as omitted variables. Also, suppose that individual n's travel cost takes the following form

$$P_{in} = C_{in} + \lambda T_{in} w_n \qquad (32)$$

where C_{in} is individual n's out-of-pocket travel cost, λ is the proportion at which travel time is valued, T_{in} is individual n's round-trip travel time to site i, and w_n is individual n's wage rate.

To get wage data, respondents were asked whether they had the ability to work extra hours, and if so, what their wage rate would be. If respondents provided a wage, then this data was used. If respondents did not provide a wage, then their wage was calculated from their household income by calculating a per/week income and using their typical number of hours worked per week to calculate an hourly wage. If no information was given about the number of hours they work in a typical week, then a 40 hour work week was assumed.

This specification of the price structure follows the McConnell and Strand (1981) approach discussed in Chapter 2. Again, the respondent's ability to choose between work and leisure determines whether their wage reveals any information about their opportunity cost of time. The simplifying assumption being made here is that all respondents can choose to work at the margin.

Modeling the Revealed Preference Data

The survey provides revealed preference data on the number of trips taken to each site.[12] The demand for site i is assumed to take the form

$$q_{in}^{RP} = f_{in}^{RP} + \sigma_{RP}\varepsilon_{in}^{RP} \tag{33}$$

where $\varepsilon_{in}^{RP} \sim N(0,1)$ and $f_{in}^{RP} = \alpha_{RP} + \beta_{RP}(C_{in} + \lambda_{RP}T_{in}w_n) + \gamma_{RP}M_n$. Note that σ_{RP} need not be the same as σ. In particular, additional error components may arise due to recall error on the part of the survey respondent, optimization error, random preferences, or omitted variables. Equation (33) can be solve for ε_{in}^{RP} to yield

$$\varepsilon_{in}^{RP} = \frac{q_{in}^{RP} - f_{in}^{RP}}{\sigma_{RP}}. \tag{34}$$

[12] The following section entitled "Data Construction" will illustrate how the sites were defined.

In many circumstances, equation (34) can be estimated by minimizing the sum of squared errors. However, in this case OLS would yield biased results since the data in this application are censored. Censored data arise when values of the dependent variable in a certain range are reported as a single value. In our case, the quantity of trips taken by the respondent must be greater than or equal to zero, even though the respondent may desire to take a negative number of trips at the price proposed.

What is observed is a censored dependent variable.

$$\tilde{q}_{in}^{RP} = \begin{cases} 0 & if \quad q_{in}^{RP} \leq 0 \\ q_{in}^{RP} & if \quad q_{in}^{RP} > 0 \end{cases}$$

$$= \begin{cases} 0 & if \quad \varepsilon_{in}^{RP} \leq \dfrac{-[\alpha_{RP} + \beta_{RP}(C_{in} + \lambda_{RP} T_{in} w_n) + \gamma_{RP} M_n]}{\sigma_{RP}} \\ \alpha_{RP} + \beta_{RP}(C_{in} + \lambda_{RP} T_{in} w_n) + \gamma_{RP} M_n + \sigma_{RP}\varepsilon_{in}^{RP} & if \quad \varepsilon_{in}^{RP} > \dfrac{-[\alpha_{RP} + \beta_{RP}(C_{in} + \lambda_{RP} T_{in} w_n) + \gamma_{RP} M_n]}{\sigma_{RP}} \end{cases}$$

In this case the relevant distribution is a combination of discrete and continuous distributions. The log likelihood function is

$$\log L = \sum_{n=1}^{N} \left\{ I_n^{RP} \left[\log\left(\frac{1}{\sigma_{RP}} \right) + \log \phi\left(\frac{\tilde{q}_{in}^{RP} - f_{in}^{RP}}{\sigma_{RP}} \right) \right] + \left(1 - I_n^{RP}\right) \log \Phi\left(\frac{-f_{in}^{RP}}{\sigma_{RP}} \right) \right\} \tag{35}$$

where I_n^{RP} is an indicator variable that equals one if $\tilde{q}_{in}^{RP} > 0$ and zero otherwise. Equation (35) is the standard Tobit model.

Maximum likelihood estimation of equation (35) will yield parameter estimates that can be used to compute welfare estimates. These parameter values are based solely on the revealed preference data. Next, I consider how the two types of stated preference data can be used to obtain parameter estimates.

Modeling the Discrete Stated Preference Data

Just as in the revealed preference case, the first step is specification of the underlying demand function for stated preference responses. In particular it is assumed that demand for site i takes the form

$$q_{in}^{SP} = \alpha_{SP} + \beta_{SP}(C_{in} + \lambda_{SP} T_{in} w_n) + \gamma_{SP} M_n + \sigma_{SP} \varepsilon_{in}^{SP} . \tag{36}$$

Again, it need not be the case that $\sigma_{SP} = \sigma$ or $\sigma_{SP} = \sigma_{RP}$. Alternative sources of error can enter the consumer's stated preference revelation. In general, it is reasonable to believe that the stated preference error derives from a different source than the revealed preference error. In addition to having been formed at a different time, the stated preference error is likely to result more from the survey instrument than it would from the sources of the revealed preference error previously mentioned.

At this point, the relevant question is what information is revealed about the above equation by the survey responses. The first stated preference question was, "Suppose that the total cost per trip of each of your trips had been \$B more. Would you have taken **any** recreation trips?" In Figure 4.1, suppose P^0 is the respondent's initial price. The respondent will answer "yes" to the question if the new price, $P^1 = P^0 + B$, is less than $P_{Q=0}$, the price at which they demand zero trips. Equivalently, the respondent will answer "yes" if the quantity, evaluated at the new price of P^1, is greater than or equal to zero.

Thus, the probability of observing a "yes" response to this question can be written

$$\Pr(yes) = \Pr\left[q_{in}^{SP}(P_{in}', M_n) > 0\right] \tag{37}$$

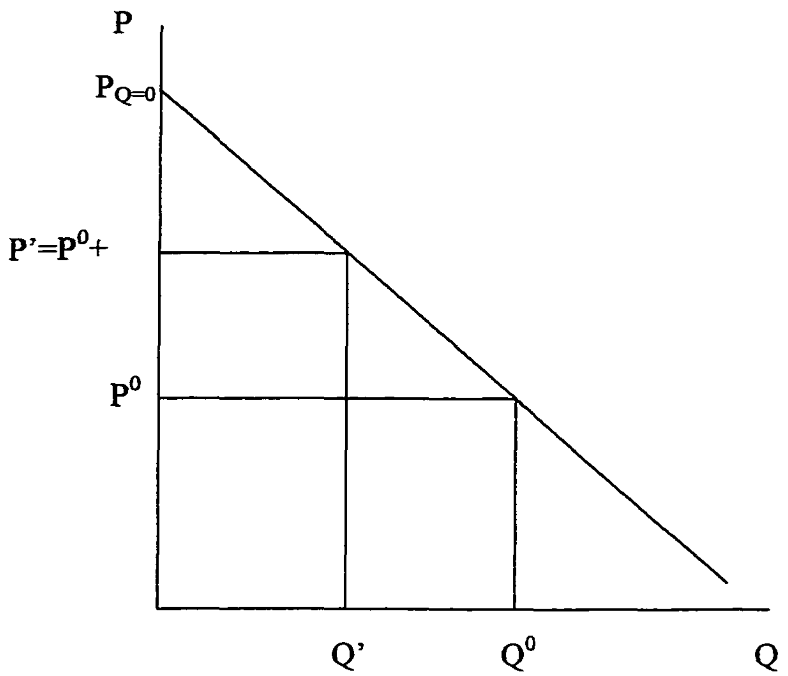


Figure 4.1: Trip demand function

where $P'_{in} = C_{in} + B + \lambda_{SP} T_{in} w_n$. This yields the following probability statement

$$\Pr(yes) = \Pr\left[\varepsilon_{in}^{SP} \geq \frac{-\alpha_{SP} - \beta_{SP}(C_{in} + B + \lambda_{SP} T_{in} w_n) - \gamma_{SP} M_n}{\sigma_{SP}}\right]. \tag{38}$$

Inspection of equation (38) indicates that it is not possible to separately identify the five parameters using this stated preference data alone. Instead, the parameter ratios α/σ, β/σ, λ/σ, and γ/σ are recovered. The associated likelihood function is

$$\log L = \sum_{n=1}^{N} \left\{ I_n^{SP} \log\left[1 - \Phi\left(\frac{-f_{in}^{SP}}{\sigma_{SP}}\right)\right] + \left(1 - I_n^{SP}\right) \log \Phi\left(\frac{-f_{in}^{SP}}{\sigma_{SP}}\right) \right\} \tag{39}$$

where $f_{in}^{SP} = \alpha_{SP} + \beta_{SP}(C_{in} + B + \lambda_{SP} T_{in} w_n) + \gamma_{SP} M_n$, the indicator function, I_n^{SP}, takes on a value of 1 if the respondent answered yes to the stated preference question and 0 otherwise, and $n = 1,\ldots,N$. Maximum likelihood estimation will yield estimates of the parameter ratios based on the stated preference responses. It is important to note that these parameter estimates can not be used to compute welfare estimates. Because the discrete stated preference model is only capable of estimating parameter ratios, the price parameter estimate, which is necessary for computation of consumer surplus, can not be isolated. Next, I consider how the continuous stated preference data can be used to generate parameter estimates.

Modeling the Continuous Stated Preference Data

The model used to generate parameter estimates from the continuous stated preference data is identical to the revealed preference model. In this case the "RP" subscripts and superscripts are changed to "SP." As in equation (33), the demand for site i is assumed to take the form

$$q_{in}^{SP} = f_{in}^{SP} + \sigma_{SP} \varepsilon_{n}^{SP} \tag{40}$$

where $\varepsilon_{in}^{SP} \sim N(0,1)$, and $f_{in}^{SP} = \alpha_{SP} + \beta_{SP}(C_{in} + B + \lambda_{SP} T_{in} w_n) + \gamma_{SP} M_n$. Again, it need not be the case that $\sigma_{SP} = \sigma$ or $\sigma_{SP} = \sigma_{RP}$. Equation (40) can be solve for ε_{in}^{SP} to yield

$$\varepsilon_{in}^{SP} = \frac{q_{in}^{SP} - f_{in}^{SP}}{\sigma_{SP}}. \tag{41}$$

As in the RP case, I am dealing with censored data and the standard Tobit model must be used. The likelihood function takes the following form

$$\log L = \sum_{n=1}^{N} \left\{ I_n^{SP} \left[\log\left(\frac{1}{\sigma_{SP}}\right) + \log\phi\left(\frac{\tilde{q}_{in}^{SP} - f_{in}^{SP}}{\sigma_{SP}}\right) \right] + \left(1 - I_n^{SP}\right) \log\Phi\left(\frac{-f_{in}^{SP}}{\sigma_{SP}}\right) \right\} \tag{42}$$

where I_n^{SP} is an indicator variable that equals one if $\tilde{q}_{in}^{SP} > 0$, and zero otherwise.

Maximum likelihood estimation of equation (42) will yield parameter estimates that can be used to compute welfare estimates. These parameter values are based solely on the continuous stated preference data.

In both the discrete and continuous cases, the stated preference data can be used alone to estimate the parameters of the model. In the discrete stated preference case, it is only possible to estimate parameter ratios because of the limited information contained in the yes/no response, but both the continuous revealed and stated preference models allow for estimation of all parameters of the model. Next, I consider how the two types of stated preference data can be linked with the revealed preference data to obtain parameter estimates.

Linking the Revealed and Stated Preference Data

In developing the model that will be used to link the revealed and stated preference data it will be useful to reparameterize the stated preference model. Setting $k_\alpha^{SP} \equiv \alpha_{SP}/\alpha_{RP}$, $k_\beta^{SP} \equiv \beta_{SP}/\beta_{RP}$, etc., the following demand specification is obtained

$$q_{in}^{SP} = \alpha_{RP} k_\alpha^{SP} + \beta_{RP} k_\beta^{SP} \left(C_{in} + \lambda_{RP} k_\lambda^{SP} T_{in} w_n\right) + \gamma_{RP} k_\gamma^{SP} M_n + \sigma_{RP} k_\sigma^{SP} \varepsilon_{in}^{SP}. \tag{43}$$

This demand function can now be used to develop the likelihood function for both the discrete and continuous stated preference cases, as shown in the previous two sections. This is a very convenient parameterization. In order to use the data to estimate one set of preference parameters in the linked model, all k parameters are restricted to be equal to one. If one wishes to estimate a set of revealed preference parameters and a set of stated preference parameters using the linked model, then both the revealed preference parameters and the stated preference k parameters can be estimated.

Assuming that the revealed and stated error components (ε_{in}^{RP} and ε_{in}^{SP}) are independent, linking the revealed and stated preference data is simply a matter of summing the revealed and stated preference likelihood functions and estimating by maximum likelihood estimation. In the RP/SPd case, the log likelihood function takes the following form

$$\begin{aligned} LogL = \sum_{n=1}^{N} \Bigg\{ & I_n^{RP} \left[\log\left(\frac{1}{\sigma_{RP}}\right) \log\phi\left(\frac{\tilde{q}_{in}^{RP} - f_{in}^{RP}}{\sigma_{RP}}\right) \right] + \left(1 - I_n^{RP}\right) \log\Phi\left(\frac{-f_{in}^{RP}}{\sigma_{RP}}\right) \\ & + I_n^{SP} \log\left[1 - \Phi\left(\frac{-f_{in}^{SP}}{\sigma_{SP}}\right)\right] + \left(1 - I_n^{SP}\right) \log\Phi\left(\frac{-f_{in}^{SP}}{\sigma_{SP}}\right) \Bigg\} \end{aligned} \tag{44}$$

where I_n^{RP} is an indicator variable that equals one if respondent n took positive trips and zero otherwise, and I_n^{SP} is an indicator variable that equals one if the respondent answered "yes" to the stated preference question and zero otherwise. This likelihood function is only applicable if the revealed and stated preference errors are assumed to be independent, an assumption that is clearly not realistic when the data are from the same individuals.

In the more realistic case where the error variances are correlated, the likelihood function for the RP/SPd case takes the form

$$\log L=\sum_{n=1}^{N}\left\{I_n^{RP}\left[\ln\phi\left(\frac{q_{in}^{RP}-f_{in}^{RP}}{\sigma_{RP}}\right)-\ln(\sigma_{RP})\right]+I_n^{RP}I_n^{SP}\left[\ln\Phi\left(\frac{\left(\frac{q_{in}^{SP}-f_{in}^{SP}}{\sigma_{SP}}\right)-\rho\left(\frac{q_{in}^{RP}-f_{in}^{RP}}{\sigma_{RP}}\right)}{\left(1-\rho^2\right)^{1/2}}\right)\right]\right.$$
$$\left.+I_n^{RP}\left(1-I_n^{SP}\right)\left[\ln\Phi\left(\frac{\frac{-f_{in}^{SP}}{\sigma_{SP}}-\rho\left(\frac{q_{in}^{RP}-f_{in}^{RP}}{\sigma_{RP}}\right)}{\left(1-\rho^2\right)^{1/2}}\right)\right]+\left(1-I_n^{RP}\right)\left(1-I_n^{SP}\right)\ln\int_{-\infty}^{\frac{-f_{in}^{RP}}{\sigma_{RP}}}\int_{-\infty}^{\frac{-f_{in}^{SP}}{\sigma_{SP}}}\phi_2(\eta_1,\eta_2;\rho)d\eta_1 d\eta_2\right\} \quad (45)$$

where ρ is the correlation coefficient associated with the revealed and stated preference errors, and $\phi_2(\cdot,\cdot;\rho)$ is the standard normal bivariate pdf.[13] As discussed in Chapter 2, this is an example of a model that combines, as opposed to pools, two types of data with different information content.

For the RP/SPc case, the log likelihood function takes the following form

$$\log L=\sum_{n=1}^{N}\left\{I_n^{RP}\left[\ln\phi\left(\frac{q_{in}^{RP}-f_{in}^{RP}}{\sigma_{RP}}\right)-\ln(\sigma_{RP})\right]+I_n^{RP}I_n^{SP}\left[\ln\Phi\left(\frac{\left(\frac{q_{in}^{SP}-f_{in}^{SP}}{\sigma_{SP}}\right)-\rho\left(\frac{q_{in}^{RP}-f_{in}^{RP}}{\sigma_{RP}}\right)}{\left(1-\rho^2\right)^{1/2}}\right)-\ln\left(\left(1-\rho^2\right)^{1/2}\right)\right]\right.$$
$$\left.+I_n^{RP}\left(1-I_n^{SP}\right)\left[\ln\Phi\left(\frac{\frac{-f_{in}^{SP}}{\sigma_{SP}}-\rho\left(\frac{q_{in}^{RP}-f_{in}^{RP}}{\sigma_{RP}}\right)}{\left(1-\rho^2\right)^{1/2}}\right)\right]+\left(1-I_n^{RP}\right)\left(1-I_n^{SP}\right)\ln\int_{-\infty}^{\frac{-f_{in}^{RP}}{\sigma_{RP}}}\int_{-\infty}^{\frac{-f_{in}^{SP}}{\sigma_{SP}}}\phi_2(\eta_1,\eta_2;\rho)d\eta_1 d\eta_2\right\}. \quad (46)$$

This is a model that pools two types of data, both of which contain the same information, though the two types of information may not be equally reliable.

---

[13] Derivation of the RP/SPd and RP/SPc log-likelihood functions is shown in Appendix 3.

Welfare Measurement

Parameter estimates from any of the models discussed above, with the exception of the SPd model, can be used to calculate consumer surplus, which takes the following form for the linear demand function

$$CS_{in} = \frac{q_{in}^{2}}{-2\hat{\beta}}. \tag{47}$$

This is the consumer surplus associated with complete elimination of the good, site i in this example.[14] Obviously, if the revealed and stated preference data is used to estimate a single set of preference parameters (all k parameters restricted to one) then only one consumer surplus can be calculated. However, if a separate set of revealed and stated parameters are estimated, then both a revealed and stated consumer surplus can be calculated. The revealed preference consumer surplus would be constructed using $\hat{\beta}_{RP}$ and the stated preference consumer surplus would be constructed using $\hat{\beta}_{SP}$.

In the next section I examine the various hypothesis tests that can be considered using the models developed above.

Hypothesis Testing

The construction of this model allows for a number of interesting hypothesis tests. The most basic test would be a general test of consistency between the revealed and stated preference data sets. In other words, can I reject the null hypothesis that the revealed and stated data sets produce identical estimates of the preference parameters? For the pooling

[14] The error imposed by using consumer surplus rather than compensating variation is likely to be small for recreation goods due to the small size of consumer surplus as a percent of income.

model that links revealed and continuous stated preference data, the test takes the following form:

Test 1(a): General Consistency RP/SPc

H_0: $k_\alpha^{SP} = 1$ and $k_\beta^{SP} = 1$ and $k_\gamma^{SP} = 1$ and $k_\lambda^{SP} = 1$ and $k_\sigma^{SP} = 1$

H_A: At least one equality does not hold.

A likelihood ratio statistic is used, with

$$\psi_{1(a)} = 2\log\left(\frac{L_{1(a)}^0}{L_{1(a)}}\right) \tag{48}$$

where $L_{1(a)}^0$ is the restricted model's likelihood function value and $L_{1(a)}$ is the unrestricted model's likelihood value.

For the combining model that links revealed and discrete stated preference data, the test takes a slightly different form due to the fact that a value must be chosen for k_σ^{SP}. The value chosen for k_σ^{SP} in each of the discrete models was the value of the parameter estimate from the corresponding continuous model. If the stated preference model were being estimated alone it would be standard practice to normalize the value of k_σ^{SP}, usually to unity. However, the continuous model provides an estimate of k_σ^{SP}, which was an obvious choice.

This test takes the following form:

Test 1(b): General Consistency RP/SPd

H_0: $k_\alpha^{SP} = 1$ and $k_\beta^{SP} = 1$ and $k_\gamma^{SP} = 1$ and $k_\lambda^{SP} = 1$

H_A: At least one ratio is different.

The test statistic, $\psi_{1(b)}$ similar to equation (48), can be calculated, and is distributed χ_4^2.

In order to conduct these tests, an unrestricted model, as well as a model that restricts the parameters to be equal between the two data sets needs to be estimated. In the linking models developed above, this is simply a matter of allowing the k parameters to be freely estimated in the first case, and setting all k parameters equal to one in the second case.

Cameron (1999) has proposed another test that will be considered. Cameron was interested in the effects on consistency of allowing the variance to differ between the revealed and stated preference data sets. She calls this the heteroscedasticity hypothesis. The specific form of this test is the same for both the continuous and discrete stated preference data sets, and takes the following form:

Test 2: Heteroscedasticity

H_0: $k_{\alpha}^{SP} = 1$ and $k_{\beta}^{SP} = 1$ and $k_{\gamma}^{SP} = 1$ and $k_{\lambda}^{SP} = 1$

H_A: At least one equality does not hold.

As in the above cases, the test statistic ψ_2 is calculated as in equation (48), and is distributed χ_4^2.

In addition to these general consistency hypothesis tests, the model also allows for hypothesis testing concerning various sources of bias in each of the revealed and stated preference methodologies. The idea behind these tests is that one of the sources of data is assumed to be accurate and the other source of data is tested against it. For example, Cameron (1992) has suggested that revealed and stated preference data can be linked to impose the "discipline of market behavior" on the stated preference data.

One objection to the use of stated preference data is that respondents may ignore their budget constraint when answering willingness to pay questions, which would result in an overstatement of willingness to pay. This suggests that the RP data could be used as a basis from which to test the validity (or bias) in the SP data. In essence, the RP data is assumed to be accurate and the validity of the SP data is tested against it. This type of test is very straightforward with the model developed above. To conduct this test, variance and income parameter estimates are allowed to vary, while all other parameters are restricted to be equal between the revealed and stated portions of the model. The test takes the form:

Test 3: Variance and Income Inconsistency

H_0: $k_{\alpha}^{SP} = 1$ and $k_{\beta}^{SP} = 1$ and $k_{\lambda}^{SP} = 1$

H_A: At least one equality does not hold.

The test statistic ψ_3 is distributed χ_3^2.

As an alternative approach, the roles of the revealed and stated data can be reversed. The stated preference data can be used as a basis for a validity test of the revealed preference data. In particular, Randall (1994) has argued that the price term in revealed preference data is measured with a high degree of error, and is likely the cause of significant bias. As in the test considered above, variance and price parameter estimates are allowed to vary, while all other parameters are restricted to be equal between the revealed and stated portions of the model. The test takes the form:

Test 4: Variance and Price Inconsistency

H_0: $k_{\alpha}^{SP} = 1$ and $k_{\lambda}^{SP} = 1$ and $k_{\gamma}^{SP} = 1$

H_A: At least one equality does not hold.

The test statistic ψ_4 is distributed χ_3^2.

The same type of test can be conducted concerning the opportunity cost of time parameter, λ. If it is reasonable to think that respondents might ignore their budget constraint when answering willingness to pay questions, then it seems reasonable to believe that they might ignore their time constraint when answering the same questions. It is common for recreationalists to cite time much more than money as a constraining element in their recreation decisions. For this test, the RP data is used as a basis from which to test the validity of the SP data.

Test 5: Variance and Time Inconsistency

H_0: $k_\alpha^{SP} = 1$ and $k_\beta^{SP} = 1$ and $k_\gamma^{SP} = 1$

H_A: At least one equality does not hold.

The test statistic ψ_5 is distributed χ_3^2.

As mentioned in Chapter 3, the state of Iowa was divided into fifteen zones, which were then grouped into five megazones. The next section presents the results of applying the models and hypothesis tests developed above, using the data from each of the five megazones.

An Application to the Wetland Data Base

In this section I provide results from applying the above models using the wetlands survey data. The data was used to estimate two models for each megazone: the first model linked RP data with discrete SP data, and the second linked RP data with continuous SP data. I will first describe how the wetland "sites" were defined, present the parameter estimates, and discuss the results and interpretation of the various hypothesis tests.

Data Construction

The first step in the estimation of the parameters of the demand function was to construct a set of travel costs for each respondent in the data set. An important step in any travel cost model is definition of the sites.

Survey question number two asked the respondent to place an X on a map of Iowa at the location of their most recent wetland visit. This information was intended to give us an idea about the location of the important wetland sites in each of the 99 counties. This information was coded into the database using geographic information system (GIS) equipment. The longitude/latitude coordinates for the visitation points in each county were averaged to find the mean visitation point in that county.

The next step was to calculate the travel time and travel cost for each respondent from their residence to each of the 99 mean visitation points. The software used was PC Miler, a package designed for use in the transportation and logistics industry. PC Miler calculated the travel time and cost using the most practical route from point to point.[15] See Appendix 2 for a description of how PC Miler calculates mileage and routes.

PC Miler used a per-mile cost of $0.21, and calculated travel time by using the mile-per-hour figures shown in Table 4.1. This gave us a data set with 99 travel times and costs for each respondent in the data set.

The goal was to get a travel time and cost for each respondent from their residence to each of the fifteen zones (see map on page four of the survey). For each respondent, I

[15] The most practical route is not always the shortest in terms of miles, but is often the shortest in terms of travel time.

Table 4.1: Road Types and Mile-Per Hour Figures

Road Type	Miles Per Hour
Multi-Lane Toll Free	65
Toll Roads	65
Divided Highways	50
Non-Divided Highways	40
Local Roads	30
Urban Highways	45
Urban Access Roads	50

calculated a weighted average of the county travel costs within each zone, with each county's weight being determined by the percentage of trips within that county's zone taken to that county. For example, zone three is made up of Pottawattamie, Mills, and Fremont counties. There were 92 trips taken to zone three. 41 trips were taken to Pottawattamie County, 24 trips were taken to Mills County, and 27 trips were taken to Fremont County. Therefore, the weight for Pottawattamie County was 0.45, the weight for Mills County was 0.26, and the weight for Fremont County was 0.29. This method was used to construct a set of fifteen travel times and costs for each respondent in the data set, one for each zone.

The fifteen zones were then grouped into five megazones, each containing three zones. Zones with similar regions of wetlands were grouped together. For example, zones 4, 5, and 8 were grouped into the 4,5,8 megazone, which encompasses the prairie pothole region of Iowa, while the 1,2,3 megazone encompasses mainly riverine wetlands. This reduced the state to five possible sites. Zones 1, 2, and 3 were grouped into one megazone, zones 4, 5, and 8 another, zones 6, 7, and 12 another, zones 9, 10, and 11 another, and zones 13, 14, and 15 made up the final megazone.

The total number of respondents in each megazone who provided useable data was 219 for the 1,2,3 megazone, 276 for the 4,5,8 megazone, 370 for the 6,7,12 megazone, 1433 for the 9,10,11 megazone, and 358 for the 13,14,15 megazone. This is a total of 2656 observations.

The demand function specified in equation (31) includes only one good, single day wetland visits. The price for each site, or megazone, was a weighted average of the prices of the three zones making up the megazone. The weights were determined in the following manner:

- For zone i in a megazone, the total number of trips taken to each of the three zones in the megazone, by residents of zone i was determined. For example, residents of zone four took 1454 trips to zone four, 32 trips to zone five, and 26 trips to zone eight.
- For members of zone i, the weight used for zone j was the percent of trips taken to zone j by residents of zone i. For example residents of zone four took 96% of their trips to zone four, 2% of their trips to zone five, and 2% of their trips to zone eight. Therefore, their weight for zone four was 0.96, for zone five 0.02, and for zone eight 0.02.[16]

The quantity of wetland recreation trips taken by each respondent to megazone i was simply the sum of their trips taken to the zones that make up megazone i. This provided us

[16] This gave us a separate set of weights for respondents from each zone.

with a data set that contained a single price (travel cost) and quantity (number of trips) for each respondent. This data set was used with the model developed above to generate parameter estimates for each of the five megazones.

Parameter Estimates

The following tables show the coefficient estimates for the 4,5,8 megazone. Coefficient estimates for the other megazones are shown in Tables A4.1 through A4.20 of Appendix 4. Parameter estimates for the 4,5,8 megazone are shown here, and are grouped into two tables.

Table 4.2 shows parameter estimates for the 4,5,8 megazone using the model that assumes the revealed and stated preference errors are correlated, and pools revealed preference data with continuous stated preference data. Parameters are presented for the unrestricted model, the fully restricted model (all k's equal to one), and the heteroscedasticity model (all k's equal one except k_{σ}^{SP}). For the revealed preference parameters (first five in Table 4.2), the t-statistics are shown in parenthesis below the parameter estimate. For the stated preference k parameters (the next five in Table 4.2), the estimate of k is given, along with the implied stated preference parameter estimate in brackets. Below these two numbers, in parenthesis, is the value of the test statistic for a test of whether that k estimate is significantly different from one. Consumer surplus estimates are included in the last two rows of the table.

Table 4.3 shows the 4,5,8 megazone parameter estimates for the model that assumes that the revealed and stated preference errors are correlated, and combines revealed

Table 4.2: 4,5,8 Megazone: RP and SP correlated, continuous SP data

	Unrestricted	All k's equal one	Heteroscedasticity
Constant RP	23.38 (8.03)**	14.52 (7.48)**	14.22 (6.84)**
Price RP	-0.94 (-7.19)**	-0.47 (-11.34)**	-0.46 (-8.02)**
Income RP	0.08 (1.40)	0.21 (4.29)**	0.22 (4.58)**
Lambda RP	0.06 (0.89)	0.44 (3.93)**	0.47 (3.30)**
Sigma RP	13.84 (19.00)**	14.34 (18.40)**	14.48 (18.10)**
k constant SP	0.48 [11.22] (-3.69)**	1.00	1.00
k price SP	0.38 [-0.36] (-7.90)**	1.00	1.00
k income SP	3.33 [0.27] (1.07)	1.00	1.00
k lambda SP	14.69 [0.88] (0.88)	1.00	1.00
k sigma SP	1.04 [14.39] (0.44)	1.00	0.98 [14.19] (-0.24)
Rho	0.72 (17.43)**	0.72 (16.28)**	0.73 (16.92)**
-log L	1114.37	1127.56	1127.50
CS RP	99.61	197.39	203.50
CS SP	264.65	197.39	203.50
n=276			

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

preference data with discrete stated preference data. Since it is not possible to identify all of the parameters in the discrete model, the value of k_{σ}^{SP} was set to the value of the parameter estimate from the continuous model.

For both the RP/SPc and the RP/SPd models, the parameter estimates have the expected signs. As is common in applied studies, income is not significant, at least in the unrestricted models. All models exhibit a high degree of correlation between the revealed and stated preference errors, with ρ lying between 0.62 and 0.73 and statistically significant at any reasonable confidence level.

For both the RP/SPc and RP/SPd models, the k parameters indicate that there are significant differences between the revealed preference parameters and the stated preference parameters. This is particularly important for the price parameter, β, which is used for consumer surplus calculations. Both the RP/SPc and RP/SPd models indicate significant difference between the price parameters. This is reflected in the difference between the revealed preference consumer surplus and the stated preference consumer surplus.

In general, the consumer surplus estimates from both the fully restricted models (all k's equal one) and the heteroscedasticity model fall between the RP and SP consumer surplus estimates from the unrestricted model. This is an important point in relation to the hypothesis tests considered in the next section. Though certain models may be statistically different from each other, if the consumer surplus estimates do not vary considerably between the models it might be academic as to which model is chosen. However, as these estimates show, there is considerable variation in the consumer surplus estimates. Therefore,

Table 4.3: 4,5,8 Megazone: RP and SP correlated, discrete SP data

	Unrestricted	All k's equal one	Heteroscedasticity
Constant RP	25.09 (9.28)**	16.25 (8.11)**	21.46 (6.01)**
Price RP	-1.04 (-9.87)**	-0.57 (-11.67)**	-0.85 (-5.00)**
Income RP	0.04 (1.10)	0.17 (3.41)**	0.10 (1.62)
Lambda RP	0.005 (3.42)**	0.26 (3.00)**	0.09 (0.98)
Sigma RP	13.65 (19.51)**	14.07 (17.45)**	13.68 (16.44)**
k constant SP	0.49 [11.80] (-3.48)**	1.00	1.00
k price SP	0.42 [-0.44] (-8.01)**	1.00	1.00
k income SP	5.88 [0.24] (0.95)	1.00	1.00
k lambda SP	110.73 [0.55] (2938.60)**	1.00	1.00
k sigma SP	1.04 [14.20] not estimated	1.00	1.63 [22.30] (1.59)
Rho	0.67 (11.88)**	0.66 (10.76)**	0.62 (9.11)**
-log L	875.25	886.99	885.22
CS RP	90.36	163.66	110.68
CS SP	217.58	163.66	110.68
n=276			

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

the results of the hypothesis tests will be important in the choice of the consumer surplus estimate.

Judging from the fact that several of the k parameters are significantly different from one, as well as the significant t-statistics for several of the revealed preference parameter estimates, it seems unlikely that consistency will be observed in either the RP/SPc or RP/SPd cases.

Though they are difficult to compare, the SPc and SPd parameter estimates are very similar. Despite the smaller amount of information contained in the SPd answer, this model exhibits the ability to produce parameter estimates comparable to those estimated by the SPc data, which contains more information.

Parameter estimates shown in Tables A4.1 through A4.20 of Appendix 4 are consistent with those reported in Tables 4.2 and 4.3 in terms of the magnitudes of the parameters, the signs of the parameters, and the conclusions that can be drawn from the parameter estimates.

Table 4.4 shows the value of the test statistic for the five hypothesis tests outlined above. Results are shown for both the pooling model that links revealed and continuous stated preference data and the combining model that links revealed and discrete stated preference data.

These results indicate that consistency between the revealed and stated preference parameters is rejected for all consistency and bias hypotheses. This was not an unexpected result given the parameter estimates reported in Tables 4.2 and 4.3. Consistency and bias hypothesis tests conducted for the other megazones exhibit the same result: in all cases the

Table 4.4: 4,5,8 Megazone: Hypothesis tests

	RP/SPc	RP/SPd
Test 1(a) and 1(b): General Consistency	26.37 reject	23.48 reject
Test 2: Heteroscedasticity	26.26 reject	19.94 reject
Test 3: Variance and Income	25.96 reject	19.88 reject
Test 4: Variance and Price	25.89 reject	19.94 reject
Test 5: Variance and Time	14.08 reject	15.69 reject

null hypothesis of parameter equality between the revealed and stated preference parameters was rejected. Hypothesis test results for the other megazones are shown in Tables A5.1 through A5.5 of Appendix 5.

Interpreting the Results of the Hypothesis Tests

The issue of how to interpret the results of the hypothesis tests has received little attention in the literature. The basic question I am trying to answer with these hypothesis tests is whether or not the revealed and stated preference data sets are consistent with each other i.e., do the two data sets generate identical parameter estimates? If it were observed that the two models did generate parameter estimates for which the hypothesis of parameter equality would not be rejected, then I would simply conclude that the two sets of data appear to reflect the same set of underlying preferences. However, in this case a rejection of all hypotheses testing consistency is observed. Previous analyses have simply ended at this point. Here, I wish to argue that there are alternative interpretations of rejection, that depend upon one's previously held opinion concerning the relative reliability of each data source.

The different interpretations have important implications for which parameter estimates would be used for prediction and welfare analysis. Along these lines, I will consider three groups of people, which I will call RP Lovers, SP Lovers, and the Agnostics.

RP Lovers

This group is made up of people who strongly believe in the relative reliability of revealed data over stated data. They believe that it is only possible to learn something meaningful by observing actual behavior, and since stated preference methods are based on hypothetical scenarios, the data obtained through such methods is practically useless. Diamond and Hausman (1994) write

> Most economic analyses aim at explaining market transactions. Data on transactions, or potentially collectible data on transactions, are the touchstone for recognizing interesting economic analyses. However loose the connection between a theoretical or empirical analysis and transactions, this connection is the basis of the methodology of judging the credibility and reliability of economic analysis.
>
> [Our skepticism of stated preference methods] comes from the belief that the internal consistency problems come from an absence of preferences, not a flaw in survey methodology. That is, we do not think that people generally hold views about individual environmental sites (many of which they have never heard of); or that, within the confines of the time available for survey instruments, people will focus successfully on the identification of preferences, to the exclusion of other bases for answering survey questions. This absence of preferences shows up as inconsistency in responses across surveys and implies that the survey responses are not satisfactory bases for policy. (p. 63)

The statement that respondents do not even have preferences over individual environmental sites is very strong. However, this group also includes people who may believe that respondents do have preferences over environmental sites, but that stated preference techniques are "fatally flawed" in terms of their ability to flesh out those preferences.

This group might logically view a rejection of consistency as validation of their opinion that the revealed preference data is "correct" and that the "incorrect" stated preference data led to the rejection of consistency. It should also be pointed out that the most hardcore members of this group would tend to view the finding of consistency as a mere coincidence. In either case, this group would tend to view the test of consistency as a test of the validity of the stated preference data against the valid revealed preference data.

This group would naturally use the revealed preference data for parameter estimation as well as consumer surplus measurement.

SP Lovers

This group is made up of people who strongly believe in the relative reliability of stated data over revealed data. This opinion could be held for a number of reasons. For example, the nature of the "revealed" price in the travel cost model is different than the price in a typical market. Again, as noted by Randall (1994), the travel cost price must be calculated by the researcher, and likely contains a large amount of measurement error. This may not be a problem with many forms of stated preference questions.[17] In essence, the researcher can arguably design a stated preference survey instrument in such a way as to make clear the payment mechanism and the cost structure, avoiding the measurement error problem.[18] This could lead some researchers to believe that the large amount of price measurement error accompanying the travel cost model makes is unreliable compared to stated preference methods.

[17] It should be pointed out that this could be a potential source of error in the stated preference questions analyzed in this dissertation. The travel cost used in the stated preference analysis was constructed in the same manner as the travel cost is used in the revealed analysis.

[18] Other sources of error could be present.

This group would tend to view a rejection of consistency as validation of their opinion that the stated preference data is "correct" and that the "incorrect" revealed preference data led to the rejection of consistency. As in the case of RP Lovers, some SP Lovers might view a finding of consistency as a mere coincidence. Others in this group might believe that while it could be possible to extract some useable information from a revealed preference data set, the payoff does not justify the extra work when compared to stated preference techniques. The important point is that SP lovers will view the test of consistency as an implicit test of the RP data against the valid stated preference data.

SP lovers would naturally use the stated preference data for parameter estimation, as well as consumer surplus measurement.

Agnostics

This group is made up of people who have not yet made up their mind concerning the relative reliability of revealed and stated preference methods. Members of this group might not worry about conducting tests of consistency between the revealed and stated preference data, instead preferring to use the most convenient data set for the problem to be analyzed. They might point out that tests would probably indicate inconsistency, but that this should not come as a surprise, especially if you subscribe to the opinion that each form of data can be used to fill in holes in the other.

Agnostics might not have a preference for either revealed or stated preference data for parameter estimation and consumer surplus measurement. If both forms of data are available, they might choose to use both to estimate one set of parameter estimates for consumer surplus measurement.

It becomes clear that the results of the hypothesis tests are not quite as easy to interpret as one would like. The same result can mean different things to different people. The inherent problem is that neither source of data can be validated externally. The solution to the problem seems to lie in the search for a way to validate one or the other methods. For example, consider a scenario where the researcher has information concerning road construction along a major highway planned for next year. The researcher then asks respondents the number of trips they would take if their travel time were increased by M minutes because of road construction, with all other costs unchanged. This data would be ex ante stated preference data. Now suppose that the researcher was able to observe the actual number of trips taken by the same respondent during the year of the road construction. This data would be ex post revealed data concerning the same situation as the ex ante data, and could therefore be used to test consistency.

In some sense, this experiment still suffers from the criticism that a potentially biased methodology is being used to validate another potentially biased methodology. On the other hand, the exercise could serve to strengthen the link between revealed and stated methods. In either case, it is an example of the types of research that can be used to further explore the validity of the methods. Collection of both ex ante and ex post data from the same respondent seems especially promising.

Conclusion

In this chapter I have examined the use of revealed, stated, and linked models in the estimation of the parameters of a recreation demand function for respondents of each megazone. In particular, both a pooling model and a combining model were developed that allow for the estimation of a single set of parameter estimates that use both the revealed and stated data, as well as separate revealed and stated parameter estimates. These models also allow for examination of various hypothesis tests concerning revealed and stated preference consistency, as well as hypotheses concerning specific biases in the two sets of data. In the next chapter, I will examine the performance of the model across the megazones in the benefits transfer context.

CHAPTER 5. COMPARING THE RESULTS ACROSS ZONES: BENEFIT TRANSFER

The data collected in the Iowa Wetlands Survey provides an excellent opportunity to explore the issue of benefits transfer. The approach discussed in Loomis (1992) will be used to investigate convergent validity with the Iowa Wetlands data set.

As discussed in Chapter 2, Loomis (1992) uses a zonal travel cost model to test convergent validity between sport fishing data gathered from different states. In particular, a 1977 survey of Oregon freshwater steelhead sport fishing and ocean sport salmon anglers was used in conjunction with two other surveys: a 1983 Idaho freshwater steelhead sport fishing survey and a 1983 survey of Washington salmon punch card holders. He points out that these data sets suffer from three problems: (1) they were gathered at different dates, (2) they were gathered for purposes other than demand estimation, and (3) they were only gathered from anglers (users of the resource), implying that the model should account for this truncation. Loomis found evidence against convergent validity in all models examined, but states that the presence of these problems makes the hypothesis testing of convergent validity more comparable to an actual benefit transfer scenario.

Though these data sets are characteristic of those encountered in a benefit transfer situation, it would be instructive to explore the issue of convergent validity by abstracting from these problems. The Iowa Wetlands data set allows for just such an investigation. The Iowa Wetlands data set was gathered over a period of weeks, during which time it is unlikely that wetland recreational activity changed significantly. The data set was collected primarily for recreation demand modeling, and was gathered from both users and non-users of wetland recreational resources. In addition, both revealed and stated data was gathered, which allows for an investigation of the convergent validity properties of the RP model, the SP model, and the linked models.

The zonal travel cost model is convenient for truncated data like that used by Loomis. However, the Iowa Wetlands data is not truncated. It is therefore appropriate to use the individual observation travel cost model developed in Chapter 4. It is generally agreed that sociodemographic characteristics are important when constructing a benefit transfer model. In order to account for the sociodemographic characteristics of the respondents, equation (31) will be generalized. Assume that the trip demand function for site i, respondent n, takes the form

$$q_{in} = \alpha + \beta P_{in} + \gamma M_n + \delta_1 G_n + \delta_2 A_n + \delta_3 L_n + \sigma \varepsilon_{in} \tag{49}$$

where G_n is a dummy variable that takes a value of one if respondent n is a male, A_n is respondent n's age, and L_n is a dummy variable that takes a value of one if the respondent purchases either a hunting or fishing license. As before, respondent n's round trip travel cost to site i takes the following form

$$P_{in} = C_{in} + \lambda T_{in} w_n . \tag{50}$$

This model will be applied using the data from the 1,2,3 megazone and the 13,14,15 megazone. These regions, though not in separate states, are geographically separate. Loomis suggests that convergent validity studies be conducted for sites in separate states. However, there seems to be nothing critical about inter-state comparisons. If convergent validity is found not to hold for intra-state comparisons, then it is probably less likely that it will hold for inter-state comparisons. If evidence in favor of convergent validity is found first at the intra-state level, then attention can be turned to inter-state comparisons.

The 1,2,3 and 13,14,15 megazones are good candidates because they are located on the west and east sides of the state respectively. They encompass similar wetland regions, riverine wetlands, and they both border major river ways, the Missouri and Mississippi Rivers. The main difference between the two regions is population density. The 13,14,15 megazone is relatively more densely populated than the 1,2,3 megazone.

In order to test convergent validity, a likelihood ratio test will be used. Both the 1,2,3 megazone and the 13,14,15 megazone models will be independently estimated. Then a model will be estimated that restricts the parameters to be equal between the two models. The sum of the unrestricted, independent likelihood values will be compared to the restricted likelihood value using a likelihood ratio test. Table 5.1 shows parameter estimates for estimations of equation (49) using the RP data alone (top set of estimates) and the SPc data alone (bottom set of estimates). Parameter estimates are shown for the 1,2,3 megazone, the 13,14,15 megazone, and the model that restricts the parameters to be equal between megazones.

The structure of the hypothesis test for convergent validity between megazone k and megazone l takes the form:

$H_0: \Theta^k = \Theta^l$

$H_A: \Theta^k \neq \Theta^l$

Table 5.1: Revealed preference and continuous stated preference models

	1,2,3 Megazone	13,14,15 Megazone	Both Megazones
Constant RP	19.53 (4.16)**	20.45 (5.76)**	20.04 (7.09)**
Price RP	-0.66 (-6.81)**	-0.85 (-9.11)**	-0.77 (-11.49)**
Income RP	-0.03 (-0.71)	-0.04 (-1.14)	-0.03 (-1.25)
Lambda RP	-0.05 (-1.31)	-0.07 (-2.43)*	-0.06 (-2.69)**
Gender RP	2.87 (1.34)	3.98 (2.19)*	3.79 (2.67)**
Age RP	-0.17 (-2.69)**	-0.10 (-2.03)*	-0.12 (-3.21)**
License RP	6.00 (2.58)**	5.55 (3.41)**	5.46 (4.08)**
Sigma RP	11.82 (14.51)**	11.71 (21.71)**	11.80 (25.79)**
CS RP	105.80	103.83	103.86
-log L	510.59	966.06	1479.46
Constant SP	8.65 (1.62)	18.60 (2.94)**	14.25 (3.40)**
Price SP	-0.29 (-4.21)**	-0.54 (-6.30)**	-0.46 (-7.71)**
Income SP	0.19 (3.35)**	0.14 (2.86)**	0.15 (4.14)**
Lambda SP	0.39 (2.62)**	0.27 (2.73)**	0.27 (3.79)**
Gender SP	1.85 (0.71)	-0.34 (-0.12)	1.08 (0.54)
Age SP	-0.18 (-2.36)*	-0.15 (-1.76)	-0.16 (-2.81)**
License SP	4.07 (1.43)	7.10 (2.64)**	6.22 (2.90)**
Sigma SP	10.32 (8.96)**	15.51 (12.03)**	13.85 (15.32)**
CS SP	259.89	166.45	177.90
-log L	229.51	396.02	632.14

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

where Θ represents the vector of parameter estimates. The value of the test statistic for a test of convergent validity of the revealed preference model between the 1,2,3 megazone and the 13,14,15 megazone is 5.62. This is less than the critical value of $\chi^2_{8,95} = 15.51$, which leads to a failure to reject the null hypothesis.

The value of the test statistic for a test of convergent validity of the continuous stated preference model between the 1,2,3 megazone and the 13,14,15 megazone is 13.21, compared to the critical value of 15.51. Again, a failure to reject the null hypothesis is observed.

Table 5.2 shows parameter estimates for the pooling model that links revealed preference data with continuous stated preference data. Parameter estimates are shown for the 1,2,3 megazone, the 13,14,15 megazone, and the model that restricts the parameters to be equal between megazones.

Using the pooling model, the value of the test statistic for a hypothesis test of convergent validity between the 1,2,3 megazone and the 13,14,15 megazone is 20.22. Again, when compared to the critical value of $\chi^2_{16,95} = 26.30$, evidence in favor of the null hypothesis is observed.

Next, attention is turned to the discrete stated preference data. Table 5.3 shows parameter estimates of equation (49) using the revealed preference data alone (these estimates were also shown in Table 5.1), the discrete stated preference data alone, and the model that restricts the parameters to be equal between the 1,2,3 megazone and the 13,14,15 megazone.

Table 5.2: Revealed and continuous stated preference data, pooling model

	1,2,3 Megazone	13,14,15 Megazone	Both Megazones
Constant RP	18.52 (32.16)**	17.14 (4.83)**	17.76 (6.90)**
Price RP	-0.62 (-9.14)**	-0.67 (-7.55)**	-0.65 (-11.51)**
Income RP	0.01 (3.27)**	0.03 (0.90)	0.02 (0.84)
Lambda RP	0.001 (2.99)**	0.03 (0.58)	0.01 (0.48)
Gender RP	2.67 (2.84)**	3.91 (2.25)*	3.66 (2.70)**
Age RP	-0.19 (-3.87)**	-0.11 (-2.25)*	-0.14 (-3.75)**
License RP	6.14 (6.19)**	5.75 (3.51)**	5.68 (4.49)**
Sigma RP	12.24 (16.52)**	12.03 (18.90)**	12.09 (26.17)**
k constant SP	0.69 [12.79] (-1.37)**	0.91 [15.60] (-0.34)	0.84 [14.92] (-0.95)
k price SP	0.51 [-0.32] (-5.01)**	0.64 [-0.43] (-3.46)**	0.63 [-0.41] (-4.84)**
k income SP	11.95 [0.01] (523.12)**	4.50 [0.14] (0.77)	5.81 [0.12] (0.73)
k lambda SP	395.16 [0.40] (26138.94)**	12.03 [0.36] (0.56)	18.63 [0.19] (0.47)
k gender SP	0.84 [2.24] (-0.21)	-0.19 [-0.74] (-1.69)	0.27 [0.99] (-1.74)
k age SP	1.12 [-0.21] (0.33)	1.05 [-.12] (0.09)	1.15 [-0.16] (0.42)
k license SP	0.64 [3.93] (-0.97)*	1.15 [6.61] (0.37)	1.02 [5.79] (0.06)
k sigma SP	0.80 [9.79] (-2.20)*	1.14 [13.71] (1.54)	1.03 [12.45] (0.53)
Rho	0.63 (10.83)**	0.66 (12.47)**	0.64 (17.03)**
CS RP	109.21	129.04	121.23
CS SP	214.04	203.18	191.95
-log L	717.90	1319.47	2047.48

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table 5.3: Revealed preference and discrete stated preference models

	1,2,3 Megazone	13,14,15 Megazone	Both Megazones
Constant RP	19.53 (4.16)**	20.45 (5.76)**	20.04 (7.09)**
Price RP	-0.66 (-6.81)**	-0.85 (-9.11)**	-0.77 (-11.49)**
Income RP	-0.03 (-0.71)	-0.04 (-1.14)	-0.03 (-1.25)
Lambda RP	-0.05 (-1.31)	-0.07 (-2.43)*	-0.06 (-2.69)**
Gender RP	2.87 (1.34)	3.98 (2.19)*	3.79 (2.67)**
Age RP	-0.17 (-2.69)**	-0.10 (-2.03)*	-0.12 (-3.21)**
License RP	6.00 (2.58)**	5.55 (3.41)**	5.46 (4.08)**
Sigma RP	11.82 (14.51)**	11.71 (21.71)**	11.80 (25.79)**
CS RP	103.65	103.06	103.09
-log L	510.59	966.06	1479.46
Constant SP	23.13 (3.02)**	37.05 (4.32)**	32.47 (5.33)**
Price SP	-0.47 (-4.77)**	-0.66 (-6.02)**	-0.61 (-8.00)**
Income SP	0.24 (3.44)**	0.28 (4.05)**	0.27 (5.28)**
Lambda SP	0.34 (3.08)**	0.35 (3.30)**	0.33 (4.59)**
Gender SP	1.23 (0.37)	-3.53 (-0.97)	-1.30 (-0.50)
Age SP	-0.34 (-3.35)**	-0.34 (-3.28)**	-0.36 (-4.75)**
License SP	5.19 (1.32)	3.81 (1.07)	4.20 (1.58)
Sigma SP	10.32 not estimated	15.51 not estimated	13.85 not estimated
-log L	57.24	103.38	162.86

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

For the discrete stated preference model, the value of the test statistic for a test of convergent validity between the 1,2,3 megazone and the 13,14,15 megazone is 4.50. This is smaller than the critical value of $\chi^2_{7,.95} = 14.07$, which leads to a failure to reject the null hypothesis.

Table 5.4 shows parameter estimates for the combining model that links revealed preference data with discrete stated preference data. Parameter estimates are shown for the 1,2,3 megazone, the 13,14,15 megazone, and the model that restricts the parameters to be equal between megazones. It was necessary to fix the value of λ in order to achieve convergence in these models. In each case, the value of λ was fixed at the value of the parameter estimate from the appropriate pooling model.

For the combining model, the value of the test statistic is 12.81. When compared to the critical value of $\chi^2_{14,.95} = 23.69$, evidence in favor of convergent validity is observed. It should be noted that this hypothesis test was also conducted for the pooling model that restricts all RP and SPc parameters to be equal, the pooling model that restricts all RP and SPc parameters except the variance to be equal, the combining model the restricts all RP and SPd parameters to be equal, and the combining model that restricts all RP and SPd parameters except the variance to be equal. In all cases, evidence in favor of convergent validity was observed. Table 5.5 summarizes these results.

The parameter estimates shown in Tables 5.1 through 5.4 are very similar to those reported in Chapter 4. Travel cost, income, and travel time affects the quantity of trips taken in the expected manner, given the results reported in Chapter 4. Gender has a mixed effect.

Table 5.4: Revealed and discrete stated preference data, combining model

	1,2,3 Megazone	13,14,15 Megazone	Both Megazones
Constant RP	18.14 (4.43)**	17.48 (5.11)**	17.75 (6.78)**
Price RP	-0.60 (-8.43)**	-0.69 (-10.96)**	-0.66 (-13.85)**
Income RP	0.01 (0.35)	0.03 (1.35)	0.02 (1.18)
Lambda RP	0.0006 not estimated	0.03 not estimated	0.01 not estimated
Gender RP	2.65 (1.23)	3.82 (2.22)*	3.64 (2.66)**
Age RP	-0.18 (-3.12)**	-0.10 (-2.10)*	-0.13 (-3.41)**
License RP	6.11 (2.81)**	5.77 (3.63)**	5.67 (4.50)**
Sigma RP	12.13 (16.69)**	12.05 (23.10)**	12.09 (27.26)**
k constant SP	1.38 [25.03] (0.83)	1.74 [30.42] (1.75)	1.66 [29.47] (1.91)
k price SP	0.82 [-0.49] (-1.08)	0.78 [-0.54] (-1.57)	0.83 [-0.55] (-1.58)
k income SP	18.36 [0.18] (0.34)	7.79 [0.23] (1 21)	10.54 [0.21] (1.09)
k lambda SP	362.44 [0.22] (3.33)**	13.38 [0.40] (3.50)**	21.05 [0.21] (4.94)**
k gender SP	0.68 [1.80] (-0.32)	-0.76 [-2.90] (-1.85)	-0.24 [-0.09] (-1.99*)
k age SP	1.92 [-0.35] (1.29)	2.84 [-0.28] (1.30)	2.54 [-0.33] (1.87)
k license SP	0.70 [4.28] (-0.53)	0.70 [4.04] (-0.59)	0.71 [4.03] (-0.76)
k sigma SP	0.82 [9.95] not estimated	1.14 [13.74] not estimated	1.03 [12.45] not estimated
Rho	0.50 (5.54)**	0.50 (7.62)**	0.48 (8.42)**
CS RP	113.56	124.67	120.21
CS SP	139.04	159.23	144.28
-log L	561.82	1057.01	1625.24

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table 5.5: Convergent Validity, 1,2,3 megazone and 13,14,15 megazone

Model	Test Statistic	Critical Value	
RP alone	5.62	15.51	fail to reject
SPc alone	13.21	15.51	fail to reject
Pooling: All k=1	10.28	16.92	fail to reject
Pooling: Heteroscedasticity	12.82	18.31	fail to reject
Pooling: No Restrictions	20.22	27.59	fail to reject
SPd alone	4.50	14.07	fail to reject
Combining: All k=1	7.34	16.92	fail to reject
Combining: Heteroscedasticity	10.38	18.31	fail to reject
Combining: No Restrictions	12.81	25.00	fail to reject

Gender positively affects the quantity of trips taken using the revealed data. However, using the stated data, gender has a positive effect for the 1,2,3 megazone and a negative effect for the 13,14,15 megazone.

The respondent's age is inversely related to the quantity of trips in all models except the stated portion of both the pooling and combining models. In all cases, license ownership is positively related to the quantity of trips takes, as expected.

All tests of convergent validity between the 1,2,3 megazone and the 13,14,15 megazone indicate that the null hypothesis of parameter equality cannot be rejected, as shown in Table 5.5. This evidence indicates that benefit transfer between these megazones would be acceptable using this demand specification.

To further investigate the issue of convergent validity, the same set of hypothesis tests was considered between the 6,7,12 megazone and the 9,10,11 megazone. These two megazones contain mainly riverine wetlands, but differ from the 1,2,3 and 13,14,15 megazones in that they do not border either the Missouri or the Mississippi Rivers. Table 5.6 shows the value of the test statistic and the critical value for each of the models. Unlike the

Table 5.6: Convergent Validity, 6,7,12 megazone and 9,10,11 megazone

Model	Test Statistic	Critical Value	
RP alone	14.64	15.51	accept
SPc alone	23.30	15.51	reject
Pooling: All $k=1$	23.05	16.92	reject
Pooling: Heteroscedasticity	23.13	18.31	reject
Pooling: No Restrictions	42.43	27.59	reject
SPd alone	19.16	14.07	reject
Combining: All $k=1$	20.32	16.92	reject
Combining: Heteroscedasticity	21.18	18.31	reject
Combining: No Restrictions	30.35	26.30	reject

unrestricted combining model shown in Table 5.4, no restrictions had to placed on λ to achieve convergence

. These results indicate that convergent validity is rejected in all models except the model that uses only the revealed preference data. This evidence against convergent validity between two geographically close megazones is rather surprising, given that evidence in favor of convergent validity was found between the 1,2,3 megazone and the 13,14,15 megazone, which are located on opposite sides of the state. This result may be due in part to the presence of the major rivers (Missouri and Mississippi) in the 1,2,3 and 13,14,15 megazones. Perhaps the wetland recreation taking place in these megazones is more homogenous due to the presence of these two rivers.

Another explanation is that the 6,7,12 and 9,10,11 megazones differ in their rural/urban makeup. The 9,10,11 megazone contains Des Moines, whereas the 6,7,12 megazone contains few urban areas. Though some socioeconomic differences are controlledfor, the rural/urban difference is not modeled, and may contribute to the rejection of convergent validity.

This result highlights the difference between the proximity of the policy and study sites, and the homogeneity of wetland areas contained in the policy and study sites. In this application, proximity of the policy and study sites was not necessarily an indication of the ability to conduct benefit transfer between the sites.

CHAPTER 6. A CLOSER LOOK AT THE OPPORTUNITY COST OF TIME

The model developed in Chapter 4 estimated the opportunity cost of time using the McConnell and Strand (1981) approach. This chapter will compare the estimates of λ across megazones, generalize the model to account for the employment status of the respondent, and examine the effect on consistency of choosing a fixed value for λ, as is often done in the literature.

Comparing λ Across Megazones

The model developed in Chapter 4 has been used to estimate λ for all megazones. While the parameter estimates for λ appear in the tables of parameter estimates for each megazone, it is instructive to compare these estimates across megazones.

Table 6.1 shows the value of λ for the unrestricted model that pools RP and continuous SP data with correlated error and the unrestricted model that combines RP and discrete SP data with correlated error. The t-statistic for a test of difference from zero is shown in parenthesis. For both the pooled and combined models, the number in bold brackets is the t-statistic for a hypothesis of significant difference between the revealed preference λ and the stated preference λ i.e., $k_{\lambda}^{SP} = 1$. Parameter estimates are shown for all megazones, as well as a stacked model that stacked the individual megazone data sets into one large data set.

The parameter results for the pooling and the combining model are very similar. As can be seen from Table 6.1, the revealed preference λ is consistently lower than the stated preference λ. In each case the revealed preference parameter is very near zero and the stated

Table 6.1: The Value of λ Across Models for RP and SPc Data

	Pooled Model		Combined Model	
	RP	**SPc**	**RP**	**SPd**
1,2,3 megazone	-0.01 (-0.19)	0.26 {-0.19}	-0.02 (-0.66)	0.25 {-283.31}**
4,5,8 megazone	0.06 (0.89)	0.82 {0.88}	0.005 (3.42)**	0.59 {2938.60}**
6,7,12 megazone	0.06 (1.46)	0.51 {1.34}	0.04 (1.10)	0.33 {0.95}
9,10,11 megazone	0.07 (2.91)**	0.34 {2.42}*	0.04 (2.04)*	0.37 {1.85}
13,14,15 megazone	0.05 (1.97)*	0.42 {1.98}*	-0.02 (-0.77)	0.40 (-599.90)**
Stacked model	0.06 (3.52)**	0.36 {3.05}**	0.03 (1.67)	0.35 {1.59}

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

preference parameter is larger. The revealed preference λ is negative in only three cases, and in both cases it is not statistically different from zero. In many cases it is between 0.03 and 0.07, and is statistically different from zero. The pooling and combining models confirm the general result that the revealed preference λ is very near zero, while the stated preference λ is in the range of 0.30 to 0.40.

The hypothesis test for significant difference between the revealed and stated preference λ indicates that the parameters are significantly different in three of the six cases for the continuous model and four of the six cases for the discrete model.

These results indicate that the use of one-third of the wage rate as the opportunity cost of travel time would not be a good idea for the revealed preference data, although the stated preference data seem to indicate that the opportunity cost of travel time lies in this general region.

Allowing λ to Vary According to Employment Situation

The model developed in Chapter 4 treats all respondents identically when estimating the fraction of the wage at which to value travel time. Bockstael, Strand, and Hanemann (1987) have suggested that it is important to consider the employment status of the respondent when estimating the opportunity cost of time. They argue that the respondent must be able to freely adjust their work hours for their wage to reveal anything about their opportunity cost of time.

In order to explore the effects of accounting for the employment status of the respondent (i.e., whether the respondent can optimally adjust their work hours at the margin), a slight change was made to the demand function. The price specification for site i, respondent n, was changed to the following form

$$P_{in} = C_{in} + I_n \lambda_a T_{in} w_n + (1 - I_n) \lambda_f T_{in} w_n, \tag{51}$$

where I_n is an indicator variable that equals one if the respondent has the option of working additional hours to increase their total income, and zero if they cannot adjust their work hours, λ_a is the fraction of the wage rate at which respondents with adjustable work hours value travel time, and λ_f is the fraction of the wage rate at which respondents with fixed work hours value travel time. In order to allow for convenient comparison between λ_a, and λ_f, the model was reparameterized as shown in equation (43). Setting $k_\lambda^f = \lambda_f / \lambda_a$, the following expression is obtained

$$P_{in} = C_{in} + I_n \lambda_a T_{in} w_n + (1 - I_n) k_\lambda^f \lambda_a T_{in} w_n . \qquad (52)$$

It should be noted that this is a slightly different use of the k parameter than in Chapter 4. Here the parameter is used to relate λ_a, and λ_f in the same model (either RP or SP), whereas in Chapter 4, the k parameter was used to relate parameters between the RP and the SP model.

This model was applied using data from the 4,5,8 megazone. The parameter estimates for the revealed preference, continuous stated preference, and the pooling model are shown in Table 6.2. The t-statistics are shown below the parameter estimates. For the fixed lambda parameter, the value of k_λ^f is reported, along with the implied value of λ_f in brackets. The number in bold brackets below the fixed lambda parameter estimate is the t-statistic for a test of significant difference from one.

The results of this estimation are very similar to the results obtained when a single λ is estimated for all respondents. The revealed preference model estimates both λ's to be zero. The t-statistic indicates that the value of k_λ^f is significantly different from one. However, both λ_a and λ_f are estimated to be close to zero. For the stated preference model, λ_a and λ_f are estimated to be 0.41 and 0.48 respectively, with no significant difference between the parameter estimates. For the pooling model, λ_a and λ_f are

Table 6.2: Estimating λ_a and λ_f, 4,5,8 Megazone

	Revealed	Stated	Pooling, all k=1
Constant	25.04 (8.90)**	9.77 (2.41)*	14.33 (7.03)**
Price	-1.03 (-8.61)**	-0.41 (-4.70)**	-0.47 (-11.69)**
Income	0.04 (1.10)	0.20 (3.10)**	0.22 (3.93)**
Lambda (adjustable)	0.00005 (0.08)	0.41 (2.36)*	0.42 (3.57)**
k Lambda (fixed)	41.85 [0.002] {8.34}**	1.16 [0.48] {0.91}	1.15 [0.48] {0.89}
Sigma	13.33 (18.30)**	15.27 (10.46)**	14.31 (17.74)**
Rho	--	--	0.72 (15.25)**
-log L	782.77	379.17	1122.51

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

estimated to be 0.42 and 0.48 respectively, again with no significant difference between the parameter estimates.

In order to investigate the effect of generalizing the model to account for the employment status of respondents on the consistency between the revealed and stated preference data, the model was used to estimate a pooling model using the 4,5,8 megazone data. Table 6.3 shows the parameter estimates for the unrestricted model, the fully restricted model (all k parameters equal one), and the model that allows for difference in the variance of the RP and SP data (all k parameters except k_σ equal one).

For the second two models, the lambda (fixed) parameter was estimated as a proportion of the lambda (adjustable) parameter. The t-statistics in bold brackets beneath the parameter estimate are for a test of significant difference from one. For the unrestricted model, both the revealed lambda (fixed) and the revealed lambda (adjustable) were freely estimated. The t-statistic appearing in parenthesis below the parameter estimate is for significant difference from zero. Both the stated lambda (fixed) and the stated lambda (adjustable) were estimated as a proportion of the respective revealed preference parameters. Therefore, the t-statistic appearing below these parameter estimates, in bold brackets, is for significant difference from one.

The last row of Table 6.3 shows the results of the consistency tests. Both the general consistency hypothesis (equivalent of Test 1(a)) and the heteroscedasticity hypothesis (equivalent of Test 2) are rejected. These results are very similar to those shown in Chapter 4. In general, estimates for revealed preference λ are small compared to estimates for the stated preference λ. The hypothesis tests for consistency between the RP and SP data are rejected. It is also interesting to note that the adjustment made for respondents with fixed work hours does not seem to make much of a difference. For both the model that restricts all k's to one and the heteroscedasticity model, there is no statistical difference between λ_a and λ_f.

Table 6.3: Accounting for Employment Status, Pooling Model, 4,5,8 Megazone

	Unrestricted	All k=1	Heteroscedasticity
Constant RP	23.19 (7.79)**	14.33 (7.03)**	14.00 (6.63)**
Price RP	-0.93 (-7.12)**	-0.47 (-11.69)**	-0.45 (-8.12)**
Income RP	0.08 (1.43)	0.22 (3.93)**	0.22 (4.43)**
Lambda RP adjustable	0.04 (0.74)	0.42 (3.57)**	0.44 (3.16)**
Lambda RP fixed	0.07 (1.16)	1.15 [0.48] {0.89}	1.14 [0.50] {0.84}
Sigma RP	13.89 (18.41)**	14.37 (17.47)**	14.52 (17.67)**
k constant SP	0.48 [11.13] {-4.02}**	1	1
k price SP	0.37 [-0.34] {-7.88}**	1	1
k income SP	3.23 [0.26] {1.06}	1	1
k lambda SP adjustable	18.07 [0.72] {0.74}	1	1
k lambda SP fixed	13.28 [0.93] {1.00}	1	1
k sigma SP	1.03 [14.31] {0.43}	1	0.98 [14.23] {-0.30}
-log L	1109.31	1122.51	1122.43
CS RP	100.21	199.66	206.57
CS SP	272.04	199.66	206.57
Consistency Test		Reject	Reject

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Setting $\lambda = 1/3$: The Effect on Consistency

The practice of setting λ equal to one-third is very common. This is often done as a means of simplifying the model. As discussed in Chapter 1, the consumer surplus estimates can be dramatically affected by the researcher's choice of a value for λ. This problem has been discussed in the literature, and was noted by McConnell and Strand (1981) as a reason to investigate the modeling of time.

However, this may not be the only effect of an arbitrarily fixed λ. As I will show, this practice can have a dramatic impact on the hypothesis tests of consistency between revealed and stated preference data. Each of the models developed in Chapter 4 (i.e., with a single λ for all respondents) were estimated using a fixed λ of one-third. Table 6.4 shows the value of the test statistic for the continuous pooling model with correlated error. Results for both the general consistency hypothesis and the heteroscedasticity hypothesis are shown. Both hypothesis of consistency are rejected for the 9,10,11 megazone data. For all other megazones the evidence implies consistency between the revealed and stated preference data when the fraction at which travel time is valued is set to one-third.

Table 6.5 shows the value of the test statistic for the discrete combining model with correlated error. Again, results for both the general consistency hypothesis and the heteroscedasticity hypothesis are shown. As in the continuous case, the null hypothesis of consistency between the revealed and stated preference data is rejected in only a few cases. It is rejected for both hypothesis with the 9,10,11 megazone data, and the general consistency hypothesis for the 1,2,3 megazone data.

These results are very interesting when compared to Table 4.4, which show the values of the test statistics for both hypotheses of consistency when λ is estimated freely. When λ

Table 6.4: Testing consistency for the RP/SPc model with correlated error, $\lambda = 1/3$

	General Consistency	Heteroscedasticity
1,2,3 megazone	8.24 fail to reject	0.31 fail to reject
4,5,8 megazone	1.58 fail to reject	1.54 fail to reject
6,7,12 megazone	4.92 fail to reject	4.71 fail to reject
9,10,11 megazone	23.39** reject	8.13* reject
13,14,15 megazone	1.54 fail to reject	1.14 fail to reject
Critical value	$\chi^2_{4,95}$=9.49	$\chi^2_{3,95}$=7.82

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

is freely estimated, all hypothesis concerning consistency between the revealed and stated preference data sets are rejected.

To further investigate the effect of choosing a fixed λ, a search procedure was conducted that tested consistency between revealed and stated preference data using a different fixed value of λ for each test. Figure 6.1 shows the result of this search procedure for each megazone as well as for the overall data set. The test conducted is Test 1(a), which tests general consistency between the revealed continuous stated preference data.

For each data set there exists a range of values of λ that will result in a failure to reject the null hypothesis of consistency between the revealed and stated preference data.

Table 6.5: Testing consistency for the RP/SPd model with correlated error, $\lambda = 1/3$

	General Consistency	Heteroscedasticity
1,2,3 megazone	15.41** reject	1.14 fail to reject
4,5,8 megazone	0.90 fail to reject	0.89 fail to reject
6,7,12 megazone	4.80 fail to reject	1.01 fail to reject
9,10,11 megazone	41.70** reject	13.12** reject
13,14,15 megazone	6.83 fail to reject	0.99 fail to reject
Critical value	$\chi^2_{3,95}=7.82$	$\chi^2_{2,95}=5.99$

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

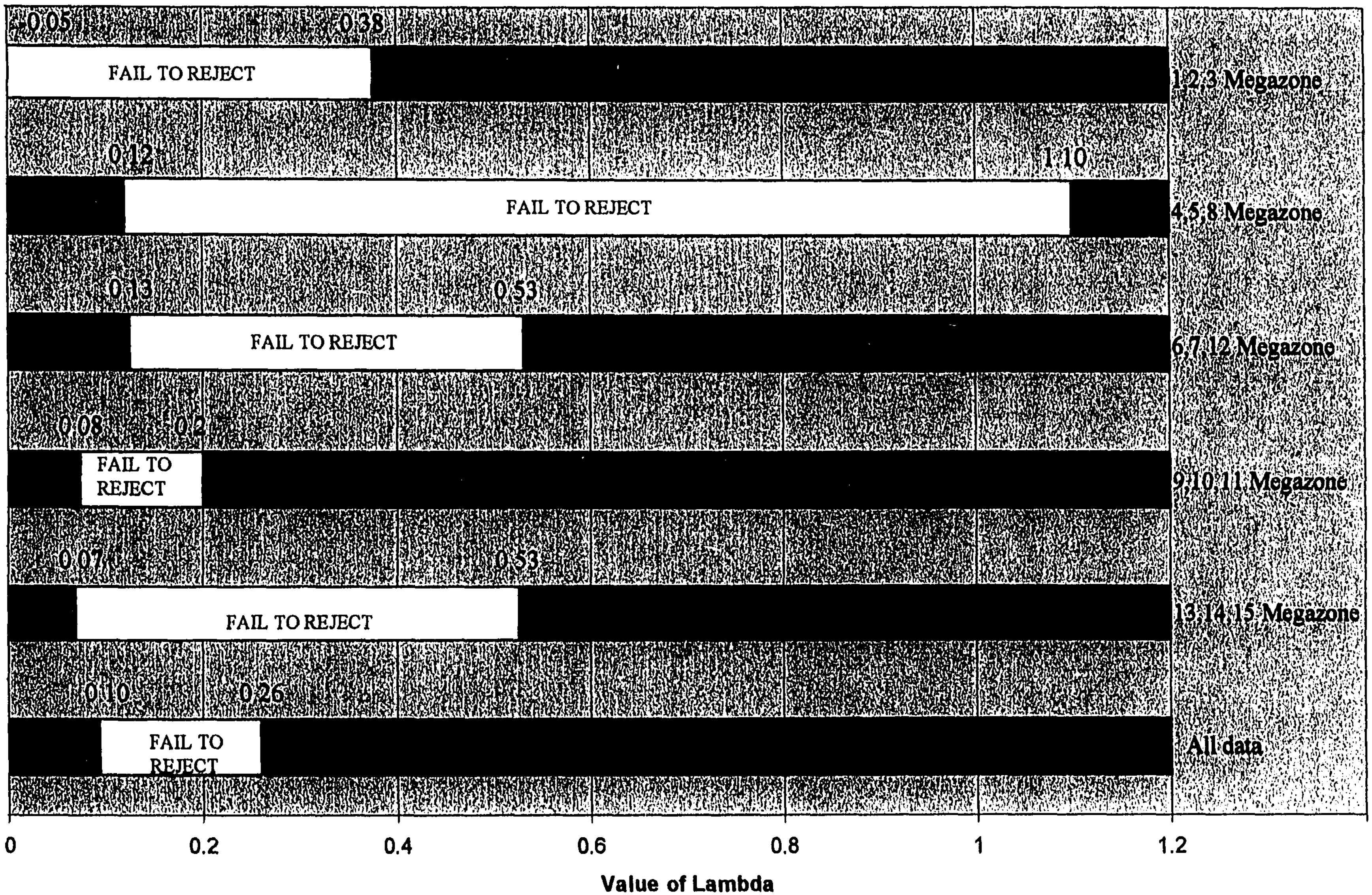


Figure 6.1: Testing general consistency with fixed lambda

The only megazone data set for which a value of λ equal to 1/3 results in a rejection of consistency is the 9,10,11 Megazone, as shown in Table 6.4.

This is a very important result. Testing for consistency between revealed and stated preference data is often a primary goal of papers that link both forms of data. As these results show, when estimating the model with a fixed λ, the consistency results depend on whether the value of λ chosen falls into the range of "consistent λ's" for that data set. However, if the opportunity cost of travel time is modeled in a more realistic manner, all tests result in a rejection of the null hypothesis of consistency.

CHAPTER 7. CONCLUSIONS

In this chapter I will summarize the main conclusions of the research. I will follow the order the results were presented, beginning with the linking models and hypothesis tests. I will then discuss the benefit transfer results, and conclude with a discussion of the results from the opportunity cost of time models.

The results of applying the model developed in Chapter 4 indicate that the data gathered in the Iowa Wetlands Survey do not exhibit consistency between the revealed preference data and the stated preference data. This result is somewhat unexpected, given the results obtained in the literature. Though there are exceptions, with some papers finding mixed results of consistency, the general result found in the literature is that consistency is often observed.

However, this result is not surprising. As discussed in Chapter 2, it has been standard practice for researchers to choose a fixed fraction for the percent of the full wage at which the respondent's travel time will be valued. When I chose a fixed value of one-third to value the opportunity cost of time, I too observed consistency between the revealed and stated preference data in most cases.

The discrete stated preference data performs comparable to the continuous stated preference data in terms of its ability to estimate the parameters of the demand model. It does not appear that the smaller amount of information contained in the respondent's discrete answer, compared to their continuous answer, hinders its performance materially. As discussed in Chapter 4, the discrete model is only capable of identifying parameter ratios when used alone, but if used in conjunction with continuous data, the discrete variance can be normalized to the value of the parameter estimated in the continuous model. The general result of RP/SP inconsistency does not depend on which type of stated preference data (discrete versus continuous) is used.

All hypothesis tests considered indicated that the null hypothesis of parameter equality between the revealed and stated preference data should be rejected. As discussed, the interpretation of this result depends on your feeling about the relative validity of revealed versus stated preference data. The hypothesis tests concerning specific biases of each form of data do not allow a persuasive argument to be made in favor of either form of data. Despite this result, these types of hypothesis tests appear promising.

In Chapter 5, I present mixed results concerning the issue of benefits transfer. The Loomis (1992) approach is used to examine whether or not a demand function estimated using wetland recreation data from one geographic region could be transferred to another geographic region. Hypothesis tests of convergent validity indicate that benefit transfer would be acceptable between the 1,2,3 megazone and the 13,14,15 megazone. However, in all models (with the exception of the model that uses only the revealed preference data) convergent validity is rejected between the 6,7,12 megazone and the 9,10,11 megazone, which are geographically closer than the 1,2,3 and 13,14,15 megazones. This may indicate the expected result that geographic proximity between the regions under consideration plays a relatively small role compared to the homogeneity of the wetland sites contained in the regions.

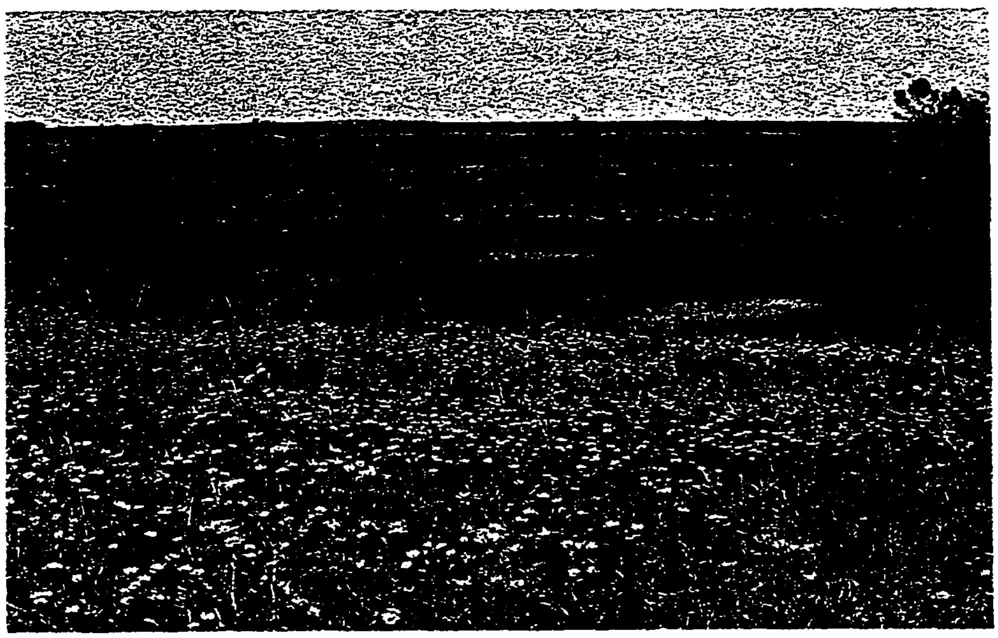
In Chapter 6, the role of the opportunity cost of travel time is examined. Cross-megazone comparisons indicate the general result that the value of λ estimated from the revealed preference data is significantly smaller than the value estimated from the stated preference data. Using the revealed preference data results in λ estimates very near zero. Often the RP parameter estimates of λ are not significantly different from zero. Using either form of stated preference data results in λ estimates that usually fall in the 0.25 to 0.50 range. Using both the revealed and stated preference data to estimate a single λ results in parameter estimates that vary from 0.07 to 0.47.

These results highlights the importance of examining the way that the opportunity cost of travel time is modeled. It also highlights the possibility of hypothesis tests concerning the respondent's time budget constraint.

APPENDIX 1. IOWA WETLANDS SURVEY

Prairie Pothole 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?

Ducks Unlimited

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.

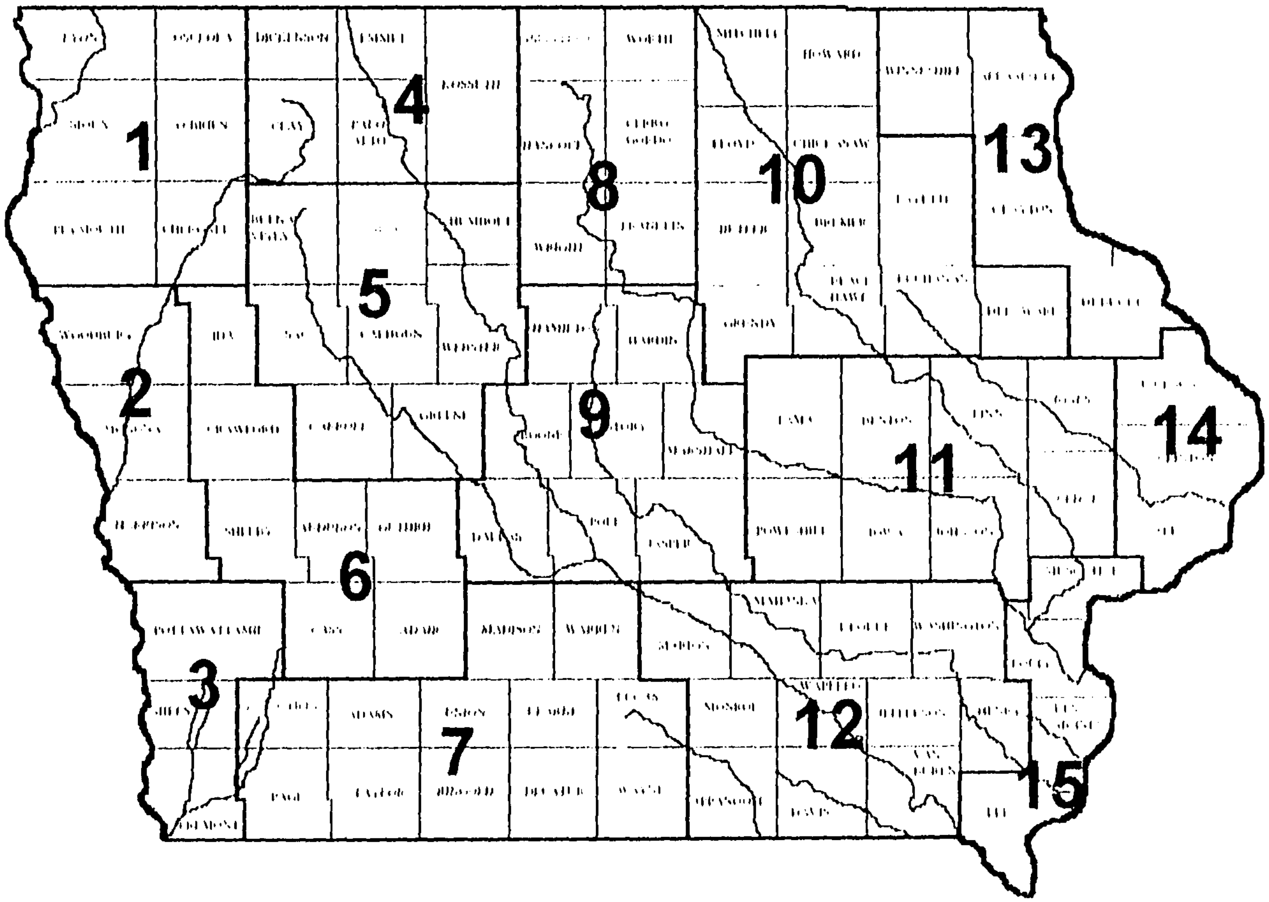
[1] Bishop, R. A., and A. van der Valk. 1982. *Wetlands*. In Cooper, T., *Iowa's Natural Heritage*. Iowa Natural Heritage Foundation and Iowa Academy of Science, Des Moines, pp. 208-29.

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.



*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 # 11), 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 $50 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 the **additional** cost of $20 per trip of visiting areas 9, 10, and 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 $20 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 mcy 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:

Scott Nelson/Ducks Unlimited

- **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.

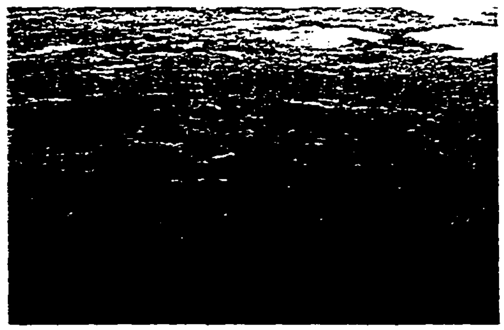
- **Water Quality.** Prairie potholes help to improve the water quality of local lakes and rivers by absorbing excess nutrients and chemicals that runoff from both farms and communities in the region.

- **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 $25 on a one time basis (payable in annual installments of $5 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
- ❑ $10,000-$14,999
- ❑ $15,000-$19,999
- ❑ $20,000-$24,999
- ❑ $25,000-$29,999
- ❑ $30,000-$34,999
- ❑ $35,000-$39,999
- ❑ $40,000-$49,999
- ❑ $50,000-$59,999
- ❑ $60,000-$74,999
- ❑ $75,000-$99,999
- ❑ $100,000-$124,999
- ❑ $125,000-$149,999
- ❑ 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.)

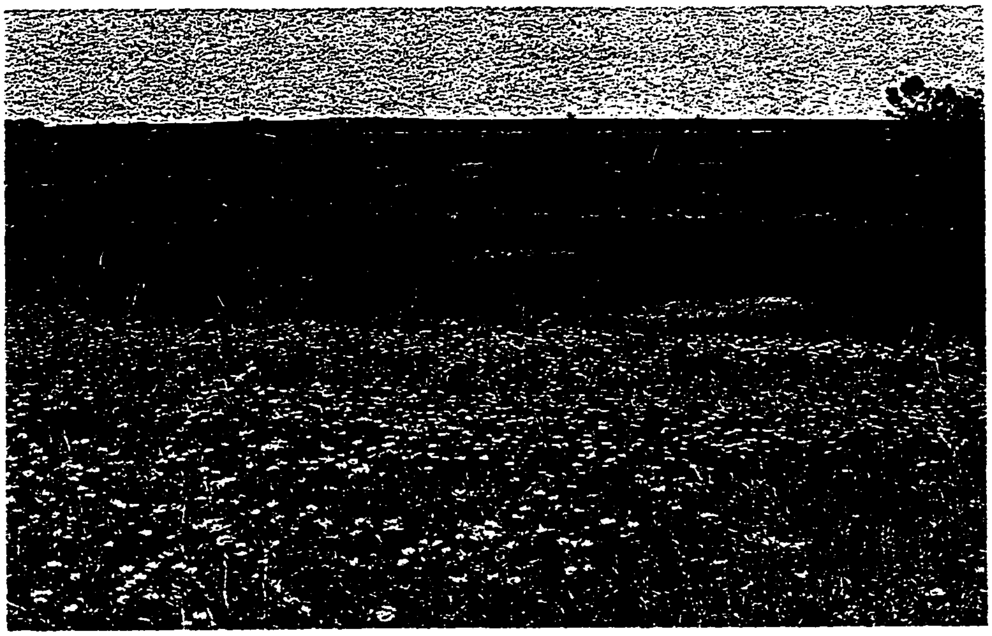
- ❑ 0 to 5%
- ❑ 5 to 10%
- ❑ 10 to 15%
- ❑ 15 to 20%
- ❑ over 20%

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 A256
IOWA STATE UNIVERSITY
AMES, IA 50011-1070

Iowa River Corridor 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?

Ducks Unlimited

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.

---

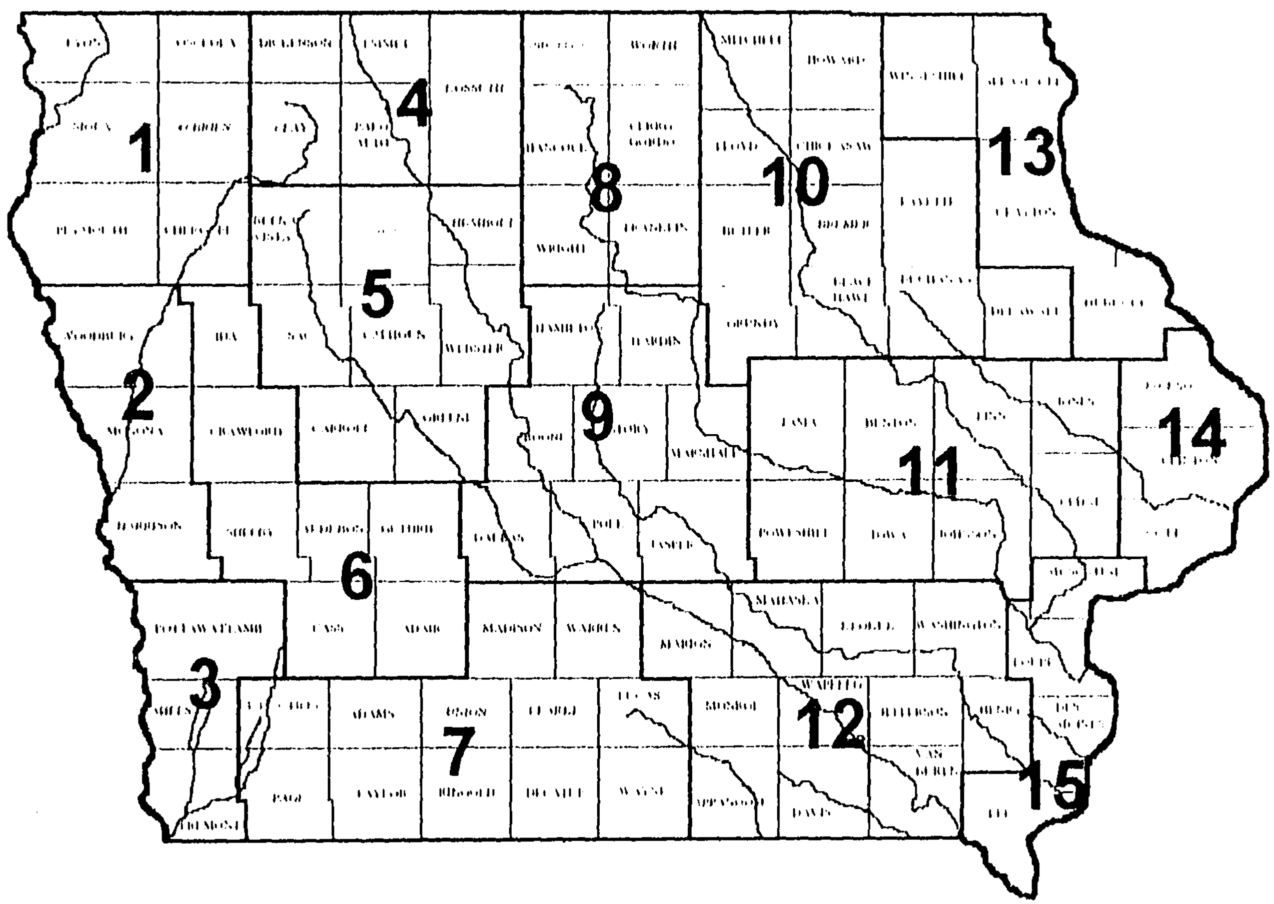
[1] Bishop, R. A., and A. van der Valk. 1982. *Wetlands.* In Cooper, T., *Iowa's Natural Heritage*. Iowa Natural Heritage Foundation and Iowa Academy of Science, Des Moines, pp. 208-29.

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.



*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 # 11), 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 $50 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 the **additional** cost of $20 per trip of visiting areas 9, 10, and 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 $20 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	☐	☐	☐	☐	☐

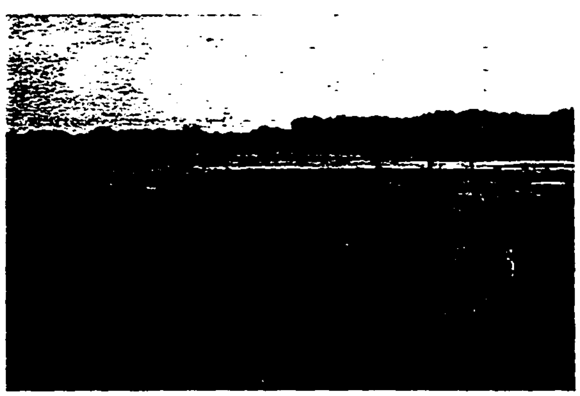
14. To protect and/or restore wetlands often costs money. How do you think wetland conservation efforts should be paid for? (Please check all that you think apply)

- ❑ 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

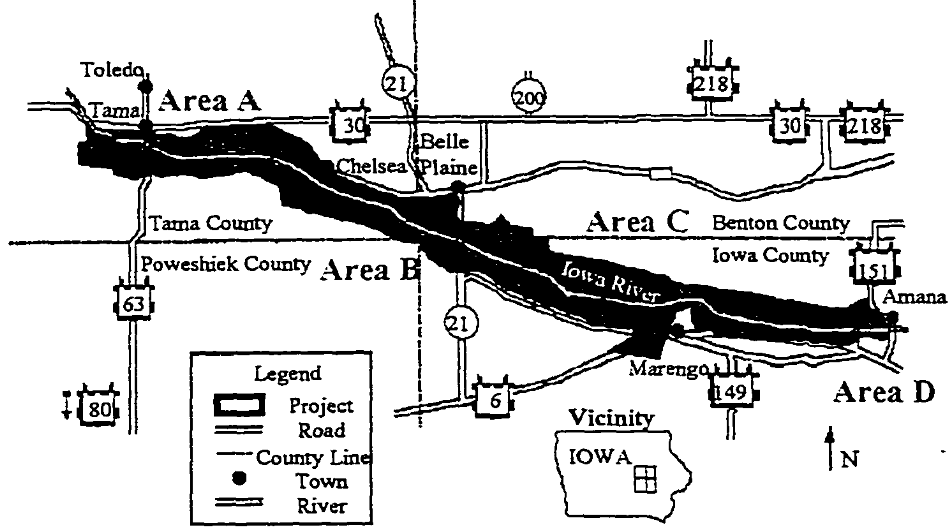
In this section, we want to focus your attention on a particular stretch of wetlands located along the Iowa River and possible changes to the extent of these wetlands. Please **do not** *go back to change your responses to earlier questions once you have read this section.*



It is estimated that prior to the 1750's, Iowa had around 2.3 million acres of wetlands. Today, Iowa has about 35,000 acres, with over 98% of the original wetlands having been converted to other uses. The majority of these wetland conversions have been to put the land in agricultural production. In the last several decades, scientists, policy makers, and landowners have begun to realize wetlands provide numerous environmental benefits that were lost by these conversions and that there may be reason to restore some of the lost areas to their wetland state.

As a consequence of the 1993 floods, many landowners in the region referred to as the Iowa River Corridor became interested in alternatives to traditional farming practices. The Iowa River Corridor is an area of saturated soils that experiences frequent flooding and which encompasses roughly a 50 mile stretch along the Iowa River between Tama and the Amana Colonies (see the map on the following page). In response to landowners concerns in the region, the Natural Resource Conservation Service initiated the Iowa River Corridor Project. Under this program, interested landowners could enroll their land in the Emergency Wetlands Reserve Program, providing them with a one time payment equal roughly to the value of their farm crops ($837 to $905/acre) in exchange for a permanent easement on the land.

Iowa River Corridor Project



The easement implies that the land is retired from agricultural production and is restored by the Natural Resource Conservation Service to wetlands. Out of the 50,000 acres in this region, about 11,600 acres have been entered into permanent easements. Under these arrangements, landowners retain their property rights to the land so that, although the land is restored to its previous wetland and upland habitat, there is not necessarily open public access to the area.

In addition, some landowners indicated a preference to sell their land outright. The US Fish and Wildlife Service responded to this interest by purchasing over 8000 acres outright in this region and converting this land to wetlands. These lands have been made part of the National Wildlife Refuge System and are open to the public.

Although it is too early to measure all of the environmental services provided by these newly reclaimed wetlands, there are several types of likely benefits:

- **Wildlife.** Wetlands provide critical areas for breeding, raising young, and gathering food for many mammals, waterfowl, birds, reptiles and amphibians. Many game species such as pheasants, ducks, and deer have greatly increased in the area. Also, sandhill cranes and bald eagles have been seen in the area in the past four years and trumpeter swans have been reintroduced into the Amana wetlands.

- **Water quality.** Wetlands help to improve the water quality of local lakes and rivers by absorbing excess nutrients and chemicals that run off from both farms and towns. Although it is too early to have scientific data on the water quality effects, experts in the area believe that the benefits will be valuable since the Iowa River currently contributes some of the largest concentrations of sediment and nutrients to the Mississippi River in the upper Midwest and atrazine and nitrate levels in Iowa City regularly exceed EPA standards.

- **Flood attenuation and storage.** Wetlands store large quantities of water temporarily thus helping to reduce the severity of floods when they occur. Because of the Iowa River Corridor project, several private levees will not be rebuilt. This should reduce the height of flood peaks and thereby reduce flood damages to lands downstream.

- **Recreation.** 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. The increased availability of wetland acres should provide increased access to recreational activities in the region.

- **Taxpayer savings/flood damage reduction.** Due to the project, crops and roads in the area will no longer suffer flood damage. The Natural Resource Conservation Service estimates that taxpayers paid upwards of $761/acre over the last ten years in disaster payments and crop subsidies which will no longer be necessary.

One objective of this survey is to determine how valuable a project like the Iowa River Corridor 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 to the project.

16. During the past year, have you visited the Iowa River Corridor project area?

- ❑ No
- ❑ Yes ⇨ If yes, how many visits did you make to each of the areas marked on the map on the previous page?

Area A ________ Area B ________ Area C ________ Area D ________

17. Although over 8000 acres of land in the Iowa River Corridor will be purchased and added to the National Wildlife Refuge System, there are additional landowners interested in selling their land which could be added to this system. This land is similar to that which has already been purchased. Suppose that 7000 acres located in Area D (the brown area) were available for acquisition over the next two years. Would you be willing to contribute $50 on a one time basis (payable in annual installments of $10 per year over five years) to an Iowa River Corridor Wetlands Management Trust fund that would cover the cost of acquiring this acreage?

- ❑ Yes
- ❑ No

18. If this acreage were acquired, how many visits do you believe you would take each year to the four wetland areas?

Area A ________ Area B ________ Area C ________ Area D ________

19. Suppose this 7000 acres were located in Area A (the red area) instead of Area D, how many visits do you believe you would take each year to the four wetlands areas?

Area A ________ Area B ________ Area C ________ Area D ________

20. To help us better understand your answers, please indicate what motivated your response to question 17:

- ❑ In general, the plan is not 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.

21. Are you

❑ male? ❑ female?

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

23. What is your age?

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

24. 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

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

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

27. 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

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

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

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

30. 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

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

32. 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

33. 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

34. 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 L175
IOWA STATE UNIVERSITY
AMES, IA 50011-1070

Table A1.1: Summary Statistics for data used in this dissertation. Standard deviation is shown in brackets

	Megazone *i,j,k*					
	1,2,3	4,5,8	6,7,12	9,10,11	13,14,15	All data
Number of respondents	214	269	363	1419	354	2619
Quantity of trips to zone *i*	619	604	437	2780	749	5189
Quantity of trips to zone *j*	317	1101	810	2674	915	5817
Quantity of trips to zone *k*	393	507	1599	4811	1070	8380
Average out-of-pocket cost[a]	26.48 [13.66]	22.62 [10.26]	30.56 [13.55]	25.37 [9.75]	25.20 [14.58]	25.88 [11.65]
Average travel time[b]	1.38 [0.68]	1.32 [0.55]	1.70 [0.70]	1.38 [0.49]	1.41 [0.79]	1.42 [0.60]
Average out-of-pocket cost with *B* increase[a]	54.38 [19.44]	47.91 [16.84]	59.02 [20.49]	50.40 [17.88]	55.49 [20.37]	52.35 [18.95]
Percent answering "yes" to q#5	20.56	23.42	21.76	20.58	20.62	21.04
Average income[c]	44.52 [31.79]	41.24 [25.56]	37.80 [24.81]	44.78 [29.93]	45.24 [30.09]	43.49 [29.12]
Percent who could adjust work hours	41.59	47.58	43.53	47.36	48.31	45.51
Wage[a]	20.41 [23.72]	17.34 [11.20]	15.62 [11.04]	19.35 [15.39]	19.33 [13.14]	18.71 [15.13]
Percent male	74.30	75.46	75.21	75.62	77.68	75.71
Average age[d]	49.67 [16.16]	48.13 [15.63]	46.86 [15.06]	48.36 15.97]	48.14 [15.41]	48.20 [15.76]
Percent who own a hunting or fishing license	72.90	77.32	78.79	68.36	68.93	71.17

[a] Measured in dollars
[b] Measured in hours
[c] Measured in thousands of dollars
[d] Measured in years

APPENDIX 2. PC MILER COST CALCULATIONS

The PC Miler documentation provides the following description for how the software calculates mileage and routes:

> All PC Miler products are based on ALK Associates' proprietary computer representation of the North American Highway system. The mileages contained in PC Miler are derived from official State highway maps, State DOT maps, county maps, local maps, and information received from thousands of industry contacts.
>
> Distances and routing directions are calculated by first determining which roads a vehicle will travel over to get between two points and then adding up the distances over each section of road to arrive at a total mileage, similar to the manual task that one might follow if using an atlas.
>
> Exact distances and driving instructions are generated if both origin and destination points are Key Cities on the ALK Highway Network. If one of the points is not a Key City, then the nearest Key City, in the direction of the other point, is selected to calculate the exact portion of the trip using actual roads. A distance estimation for the local portion of the route is then calculated. This distance is generated from the chosen Key City to the geographic center of the Non-Key City point and added to the total distance. Due to the large number of Key Cities (at least one per county), the local portion of any mileage calculation is usually a very small percentage of the total distance. (p. 6)

APPENDIX 3. DERIVATIONS

The quantity of trips taken for data set j has been specified as

$$Q^j = X^j \beta^j + \sigma_j \varepsilon^j \tag{53}$$

where $j = RP$ for revealed preference and SP for stated preference, and

$$\varepsilon \equiv \begin{pmatrix} \varepsilon^{RP} \\ \varepsilon^{SP} \end{pmatrix} \sim N\left(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right). \tag{54}$$

Equation (53) can be written equivalently as

$$\varepsilon^j = \frac{Q^j - X^j \beta^j}{\sigma_j}. \tag{55}$$

Consider two error transformations for this model. First, let

$$\begin{aligned} \tilde{\varepsilon}^{RP} &\equiv \varepsilon^{RP} - \rho \varepsilon^{SP} \\ &= \left(\frac{Q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}}\right) - \rho\left(\frac{Q^{SP} - X^{SP}\beta^{SP}}{\sigma_{SP}}\right). \end{aligned} \tag{56}$$

Since $E\left[\tilde{\varepsilon}^{RP}\varepsilon^{SP}\right] = E\left[\varepsilon^{RP}\varepsilon^{SP} - \rho\left(\varepsilon^{SP}\right)^2\right] = 0$, and

$$\begin{aligned} E\left[\left(\tilde{\varepsilon}^{RP}\right)^2\right] &= E\left[\left(\left(\varepsilon^{RP}\right)^2 - 2\rho\varepsilon^{SP} - \rho^2\left(\varepsilon^{SP}\right)^2\right)\right] \\ &= 1 - 2\rho E\left[\varepsilon^{SP}\right] - \rho^2 \\ &= 1 - \rho^2 \end{aligned} \tag{57}$$

then

$$\begin{pmatrix} \tilde{\varepsilon}^{RP} \\ \varepsilon^{SP} \end{pmatrix} \sim N\left(0, \begin{bmatrix} 1-\rho^2 & 0 \\ 0 & 1 \end{bmatrix}\right). \tag{58}$$

Second, let

$$\tilde{\varepsilon}^{SP} \equiv \varepsilon^{SP} - \rho\varepsilon^{RP}$$
$$= \left(\frac{Q^{SP} - X^{SP}\beta^{SP}}{\sigma_{SP}}\right) - \rho\left(\frac{Q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}}\right). \quad (59)$$

Then

$$\begin{pmatrix} \varepsilon^{RP} \\ \tilde{\varepsilon}^{SP} \end{pmatrix} \sim N\left(0, \begin{bmatrix} 1 & 0 \\ 0 & 1-\rho^2 \end{bmatrix}\right). \quad (60)$$

These transformations essentially purge one of the error terms of its correlation with the other error term.

Linking model: RP/SPc

When modeling the continuous revealed and stated preference data, we have two trip variables, each of which can be censored. This means there will be four components to the likelihood function.

Case 1: $Q^{RP} = q^{RP} > 0$, $Q^{SP} = q^{SP} > 0$

This information can be written equivalently as

$$\varepsilon^{RP} = \frac{q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}} \quad (61)$$

and

$$\tilde{\varepsilon}^{SP} = \left(\frac{q^{SP} - X^{SP}\beta^{SP}}{\sigma_{SP}}\right) - \rho\left(\frac{q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}}\right). \quad (62)$$

Since ε^{RP} and $\tilde{\varepsilon}^{SP}$ are independent, the corresponding contribution to the log-likelihood function is given by

$$LL_1 = \ln\phi\left(\frac{q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}}\right) - \ln(\sigma_{RP})$$
$$+\ln\phi\left(\frac{\left(\frac{q^{SP} - X^{SP}\beta^{SP}}{\sigma_{SP}}\right) - \rho\left(\frac{q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}}\right)}{\left(1-\rho^2\right)^{1/2}}\right) - \ln\left(\left(1-\rho^2\right)^{1/2}\right) \tag{63}$$

where ϕ denotes the standard normal pdf.

Case 2: $Q^{RP} \leq 0$, $Q^{SP} = q^{SP} > 0$

This information can be written equivalently as

$$\varepsilon^{SP} = \frac{q^{SP} - X^{SP}\beta^{SP}}{\sigma_{SP}} \tag{64}$$

and

$$\tilde{\varepsilon}^{RP} \leq \left(\frac{0 - X^{RP}\beta^{RP}}{\sigma_{RP}}\right) - \rho\left(\frac{q^{SP} - X^{SP}\beta^{SP}}{\sigma_{SP}}\right). \tag{65}$$

The corresponding contribution to the log-likelihood function is given by

$$LL_2 = \ln\phi\left(\frac{q^{SP} - X^{SP}\beta^{SP}}{\sigma_{SP}}\right) - \ln(\sigma_{SP}) + \ln\Phi\left(\frac{\left(\frac{-X^{RP}\beta^{RP}}{\sigma_{RP}}\right) - \rho\left(\frac{q^{SP} - X^{SP}\beta^{SP}}{\sigma_{SP}}\right)}{\left(1-\rho^2\right)^{1/2}}\right) \tag{66}$$

where Φ denotes the standard normal cdf.

Case 3: $Q^{RP} = q^{RP} > 0$, $Q^{SP} \leq 0$

This information can be written equivalently as

$$\varepsilon^{RP} = \frac{q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}} \tag{67}$$

and

$$\tilde{\varepsilon}^{SP} \leq \left(\frac{0 - X^{SP}\beta^{SP}}{\sigma_{SP}}\right) - \rho\left(\frac{q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}}\right). \tag{68}$$

The corresponding contribution to the log-likelihood function is given by

$$LL_3 = \ln\phi\left(\frac{q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}}\right) - \ln(\sigma_{RP}) + \ln\Phi\left(\frac{\left(\frac{-X^{SP}\beta^{SP}}{\sigma_{SP}}\right) - \rho\left(\frac{q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}}\right)}{(1-\rho^2)^{1/2}}\right). \tag{69}$$

Case 4: $Q^{RP} \leq 0$, $Q^{SP} \leq 0$

This information can be written equivalently as

$$\varepsilon^{RP} \leq \frac{-X^{RP}\beta^{RP}}{\sigma_{RP}} \tag{70}$$

and

$$\varepsilon^{SP} \leq \frac{-X^{SP}\beta^{SP}}{\sigma_{SP}}. \tag{71}$$

The corresponding contribution to the log-likelihood function is

$$LL_4 = \ln \int_{-\infty}^{\frac{-X^{RP}\beta^{RP}}{\sigma_{RP}}} \int_{-\infty}^{\frac{-X^{SP}\beta^{SP}}{\sigma_{SP}}} \phi_2(\eta_1, \eta_2; \rho) d\eta_1 d\eta_2 \tag{72}$$

where $\phi_2(\eta_1, \eta_2; \rho)$ denotes the standard bivariate normal cdf. Notice that the boundaries placed on the untransformed error terms do not provide boundaries on the transformed error terms (since they are error differences).

The resulting log-likelihood function is simply the combination of the appropriate components. In my data set, Case 2 would not be observed. If

$$I^j = \begin{cases} 1 & Q^j > 0 \\ 0 & otherwise \end{cases} \tag{73}$$

then the log-likelihood function can be written in the form

$$\log L = \sum_{n=1}^{N} \left\{ I_n^{RP} \left[\ln \phi\left(\frac{q_{in}^{RP} - f_{in}^{RP}}{\sigma_{RP}} \right) - \ln(\sigma_{RP}) \right] + I_n^{RP} I_n^{SP} \left[\ln \Phi \left(\frac{\left(\frac{q_{in}^{SP} - f_{in}^{SP}}{\sigma_{SP}} \right) - \rho \left(\frac{q_{in}^{RP} - f_{in}^{RP}}{\sigma_{RP}} \right)}{(1-\rho^2)^{1/2}} \right) - \ln\left((1-\rho^2)^{1/2} \right) \right] \right. \tag{74}$$

$$\left. + I_n^{RP}(1-I_n^{SP}) \left[\ln \Phi \left(\frac{\frac{-f_{in}^{SP}}{\sigma_{SP}} - \rho \left(\frac{q_{in}^{RP} - f_{in}^{RP}}{\sigma_{RP}} \right)}{(1-\rho^2)^{1/2}} \right) \right] + (1-I_n^{RP})(1-I_n^{SP}) \ln \int_{-\infty}^{\frac{-f_{in}^{RP}}{\sigma_{RP}}} \int_{-\infty}^{\frac{-f_{in}^{SP}}{\sigma_{SP}}} \phi_2(\eta_1, \eta_2; \rho) d\eta_1 d\eta_2 \right\}$$

Combining model: RP/SPd

With discrete stated preference data, the log-likelihood components are slightly different.

Case 1: $Q^{RP} = q^{RP} > 0,\ Q^{SP} > 0$

This information can be written equivalently as

$$\varepsilon^{RP} = \frac{q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}} \tag{75}$$

and

$$\tilde{\varepsilon}^{SP} \le \left(\frac{0 - X^{SP}\beta^{SP}}{\sigma_{SP}} \right) - \rho \left(\frac{q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}} \right). \tag{76}$$

The corresponding contribution to the log-likelihood function is

$$LL_3 = \ln \phi \left(\frac{q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}} \right) - \ln(\sigma_{RP}) + \ln \Phi \left(\frac{\left(\frac{-X^{SP}\beta^{SP}}{\sigma_{SP}} \right) - \rho \left(\frac{q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}} \right)}{(1-\rho^2)^{1/2}} \right). \tag{77}$$

Case 2: $Q^{RP} \le 0,\ Q^{SP} > 0$

This information can be written equivalently as

$$\varepsilon^{SP} > \frac{-X^{SP}\beta^{SP}}{\sigma_{SP}} \tag{78}$$

and

$$\varepsilon^{RP} \leq \frac{-X^{RP}\beta^{RP}}{\sigma_{RP}}. \tag{79}$$

The corresponding contribution to the log-likelihood function is

$$LL_4 = \ln \int_{-\infty}^{\frac{-X^{RP}\beta^{RP}}{\sigma_{RP}}} \int_{-\infty}^{\frac{X^{SP}\beta^{SP}}{\sigma_{SP}}} \phi_2(\eta_1, \eta_2; \rho) d\eta_1 d\eta_2 \tag{80}$$

where Φ denotes the standard normal cdf.

Case 3: $Q^{RP} = q^{RP} > 0$, $Q^{SP} \leq 0$

This is identical to the continuous choice Case 3, so the corresponding contribution to the log-likelihood function is

$$LL_3 = \ln \phi\left(\frac{q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}}\right) - \ln(\sigma_{RP}) + \ln \Phi\left(\frac{\left(\frac{-X^{SP}\beta^{SP}}{\sigma_{SP}}\right) - \rho\left(\frac{q^{RP} - X^{RP}\beta^{RP}}{\sigma_{RP}}\right)}{(1-\rho^2)^{1/2}}\right). \tag{81}$$

Case 4: $Q^{RP} \leq 0$, $Q^{SP} \leq 0$

This is identical to the continuous choice Case 4, so the corresponding contribution to the log-likelihood function is given by

$$LL_4 = \ln \int_{-\infty}^{\frac{-X^{RP}\beta^{RP}}{\sigma_{RP}}} \int_{-\infty}^{\frac{-X^{SP}\beta^{SP}}{\sigma_{SP}}} \phi_2(\eta_1, \eta_2; \rho) d\eta_1 d\eta_2 . \tag{82}$$

The overall log-likelihood function is the combination of the appropriate components. In my data set, Case 2 would not be observed. The corresponding log-likelihood function is given by

$$\log L = \sum_{n=1}^{N} \left\{ I_n^{RP}\left[\ln\phi\left(\frac{q_{in}^{RP}-f_{in}^{RP}}{\sigma_{RP}}\right)-\ln(\sigma_{RP})\right] + I_n^{RP} I_n^{SP}\left[\ln\Phi\left(\frac{\left(\frac{q_{in}^{SP}-f_{in}^{SP}}{\sigma_{SP}}\right)-\rho\left(\frac{q_{in}^{RP}-f_{in}^{RP}}{\sigma_{RP}}\right)}{\left(1-\rho^2\right)^{1/2}}\right)\right] \right.$$
$$\left. + I_n^{RP}\left(1-I_n^{SP}\right)\left[\ln\Phi\left(\frac{\frac{-f_{in}^{SP}}{\sigma_{SP}}-\rho\left(\frac{q_{in}^{RP}-f_{in}^{RP}}{\sigma_{RP}}\right)}{\left(1-\rho^2\right)^{1/2}}\right)\right] + \left(1-I_n^{RP}\right)\left(1-I_n^{SP}\right)\ln\int_{-\infty}^{\frac{-f_{in}^{RP}}{\sigma_{RP}}}\int_{-\infty}^{\frac{-f_{in}^{SP}}{\sigma_{SP}}}\phi_2(\eta_1,\eta_2;\rho)\,d\eta_1 d\eta_2 \right\} \tag{83}$$

APPENDIX 4. PARAMETER ESTIMATES

Table A4.1: 1,2,3 Megazone: RP and SP uncorrelated, continuous SP data

	Independent
Constant RP	19.57 (6.88)**
Price RP	-0.79 (-7.79)**
Income RP	-0.02 (-0.44)
Lambda RP	-0.06 (-2.00)*
Sigma RP	12.45 (15.30)**
Constant SP[a]	4.60 (1.21)
Price SP	-0.30 (-4.40)**
Income SP	0.22 (3.97)**
Lambda SP	0.42 (2.70)**
Sigma SP	10.94 (9.10)**
-log L	764.76
CS RP	84.65
CS SP	228.50

[a] These stated preference estimates are not estimates of the k^{SP} parameters. Rather, they are estimates of the parameters themselves, with the standard t-statistic in parenthesis.

** Denotes significance at the 99% confidence level

* Denotes significance at the 95% confidence level

Table A4.2: 1,2,3 Megazone: RP and SP uncorrelated, discrete SP data

	Independent
Constant RP	19.57 (6.88)**
Price RP	-0.79 (-7.79)**
Income RP	-0.02 (-0.44)
Lambda RP	-0.06 (-2.00)*
Sigma RP	12.45 (15.30)**
k constant SP	0.52 [10.21] (-1.89)
k price SP	0.59 [-0.47] (-3.26)**
k income SP	-15.26 [0.38] (-71.18)**
k lambda SP	-6.33 [0.38] (-2.65)**
k sigma SP	0.88 [10.94] not estimated
-log L	67.47
CS RP	84.85
CS SP	cannot be estimated

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.3: 1,2,3 Megazone: RP and SP correlated, continuous SP data

	Unrestricted	All k's equal one	Heteroscedasticity
Constant RP	16.54 (6.34)**	11.07 (5.51)**	9.82 (4.16)**
Price RP	-0.69 (-7.89)**	-0.46 (-11.02)**	-0.41 (-7.10)**
Income RP	0.03 (0.78)	0.11 (2.77)**	0.12 (3.15)**
Lambda RP	-0.01 (-0.19)	0.12 (2.42)*	0.17 (2.27)*
Sigma RP	12.79 (14.15)**	12.63 (15.96)**	12.98 (15.03)**
k constant SP	0.48 [7.94] (-2.93)**	1	1
k price SP	0.50 [-0.35] (-5.21)**	1	1
k income SP	4.96 [0.15] (0.67)	1	1
k lambda SP	-36.40 [0.36] (-0.19)	1	1
k sigma SP	0.82 [10.49] (-2.00)*	1	0.90 [11.68] (-1.01)
Rho	0.66 (10.46)**	0.66 (11.44)**	0.67 (11.59)**
-log L	739.62	749.65	749.16
CS RP	95.46	144.03	160.44
CS SP	192.32	144.03	160.44

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.4: 1,2,3 Megazone: RP and SP correlated, discrete SP data

	Unrestricted	All k's equal one	Heteroscedasticity
Constant RP	17.05 (413.20)**	13.16 (5.68)**	10.99 (3.89)**
Price RP	-0.71 (-17.11)**	-0.54 (-10.58)**	-0.45 (-4.62)**
Income RP	0.02 (0.78)	0.09 (2.19)*	0.11 (2.66)**
Lambda RP	-0.02 (-0.66)	0.07 (1.74)	0.13 (1.65)
Sigma RP	12.67 (23.60)**	12.62 (15.95)**	12.77 (15.01)**
k constant SP	0.71 [12.11] (-6.91)**	1	1
k price SP	0.70 [0.50] (-7.25)**	1	1
k income SP	12.85 [0.26] (287.15)**	1	1
k lambda SP	-10.69 [0.21] (-283.31)**	1	1
k sigma SP	0.82 [10.39] not estimated	1	0.77 [9.83] (-1.16)
Rho	0.54 (36.20)**	0.53 (6.12)**	0.57 (6.34)**
-log L	587.07	596.61	596.07
CS RP	93.96	122.34	148.69
CS SP	134.00	122.34	148.69

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.5: 4,5,8 Megazone: RP and SP uncorrelated, continuous SP data

	Independent
Constant RP	27.11 (8.77)**
Price RP	-1.15 (-8.37)**
Income RP	-0.02 (-0.32)
Lambda RP	-0.06 (-1.59)
Sigma RP	13.26 (18.30)**
Constant SP[a]	9.77 (2.31)*
Price SP	-0.42 (-4.64)**
Income SP	0.20 (2.92)**
Lambda SP	0.43 (2.39)*
Sigma SP	15.31 (11.25)**
-log L	1165.56
CS RP	82.80
CS SP	239.85

[a] These stated preference estimates are not estimates of the k^{SP} parameters. Rather, they are estimates of the parameters themselves, with the standard t-statistic in parenthesis.
** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.6: 4,5,8 Megazone: RP and SP uncorrelated, discrete SP data

	Independent
Constant RP	27.11 (8.77)**
Price RP	-1.15 (-8.37)**
Income RP	-0.02 (-0.32)
Lambda RP	-0.06 (-1.59)
Sigma RP	13.26 (18.30)**
Constant SP[a]	10.85 (2.33)*
Price SP	-0.47 (-4.99)**
Income SP	0.20 (2.87)**
Lambda SP	0.38 (2.30)*
Sigma SP	15.31 not estimated
-log L	913.19
CS RP	82.80
CS SP	cannot be estimated

[a] These stated preference estimates are not estimates of the k^{SP} parameters. Rather, they are estimates of the parameters themselves, with the standard t-statistic in parenthesis.
** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.7: 4,5,8 Megazone: RP and SP correlated, continuous SP data

	Unrestricted	All k's equal one	Heteroscedasticity
Constant RP	23.38 (8.03)**	14.52 (7.48)**	14.22 (6.84)**
Price RP	-0.94 (-7.19)**	-0.47 (-11.34)**	-0.46 (-8.02)**
Income RP	0.08 (1.40)	0.21 (4.29)**	0.22 (4.58)**
Lambda RP	0.06 (0.89)	0.44 (3.93)**	0.47 (3.30)**
Sigma RP	13.84 (19.00)**	14.34 (18.40)**	14.48 (18.10)**
k constant SP	0.48 [11.22] (-3.69)**	1.00	1.00
k price SP	0.38 [-0.36] (-7.90)**	1.00	1.00
k income SP	3.33 [0.27] (1.07)	1.00	1.00
k lambda SP	14.69 [0.88] (0.88)	1.00	1.00
k sigma SP	1.04 [14.39] (0.44)	1.00	0.98 [14.19] (-0.24)
Rho	0.72 (17.43)**	0.72 (16.28)**	0.73 (16.92)**
-log L	1114.37	1127.56	1127.50
CS RP	99.61	197.39	203.50
CS SP	264.65	197.39	203.50

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.8: 4,5,8 Megazone: RP and SP correlated, discrete SP data

	Unrestricted	All k's equal one	Heteroscedasticity
Constant RP	25.09 (9.28)**	16.25 (8.11)**	21.46 (6.01)**
Price RP	-1.04 (-9.87)**	-0.57 (-11.67)**	-0.85 (-5.00)**
Income RP	0.04 (1.10)	0.17 (3.41)**	0.10 (1.62)
Lambda RP	0.005 (3.42)**	0.26 (3.00)**	0.09 (0.98)
Sigma RP	13.65 (19.51)**	14.07 (17.45)**	13.68 (16.44)**
k constant SP	0.49 [11.80] (-3.48)**	1.00	1.00
k price SP	0.42 [0.44] (-8.01)**	1.00	1.00
k income SP	5.88 [0.24] (0.95)	1.00	1.00
k lambda SP	110.73 [0.55] (2938.60)**	1.00	1.00
k sigma SP	1.04 [14.20] not estimated	1.00	1.63 [22.30] (1.59)
Rho	0.67 (11.88)**	0.66 (10.76)**	0.62 (9.11)**
-log L	875.25	886.99	885.22
CS RP	90.36	163.66	110.68
CS SP	217.58	163.66	110.68

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.9: 6,7,12 Megazone: RP and SP uncorrelated, continuous SP data

	Independent
Constant RP	26.11 (10.33)**
Price RP	-0.80 (-9.50)**
Income RP	0.04 (0.96)
Lambda RP	0.01 (0.37)
Sigma RP	13.13 (22.26)**
k constant SP	0.98 [25.54] (-0.13)
k price SP	0.73 [-0.58] (-2.21)*
k income SP	4.80 [0.19] (0.77)
k lambda SP	32.98 [0.44] (0.38)
k sigma SP	1.16 [15.25] (1.58)
-log L	1451.03
CS RP	114.87
CS SP	160.44

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.10: 6,7,12 Megazone: RP and SP uncorrelated, discrete SP data

	Independent
Constant RP	26.11 (10.33)**
Price RP	-0.80 (-9.50)**
Income RP	0.04 (0.96)
Lambda RP	0.01 (0.37)
Sigma RP	13.13 (22.26)**
k constant SP	1.23 [32.19] (7.92)**
k price SP	0.93 [-0.74] (-2.54)*
k income SP	5.23 [0.21] (143.78)**
k lambda SP	24.12 [0.32] (785.26)**
k sigma SP	1.16 [15.25] not estimated
-log L	1153.34
CS RP	114.87
CS SP	cannot be estimated

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.11: 6,7,12 Megazone: RP and SP correlated, continuous SP data

	Unrestricted	All k's equal one	Heteroscedasticity
Constant RP	25.95 (11.43)**	21.15 (12.23)**	21.43 (11.71)**
Price RP	-0.79 (10.11)**	-0.55 (-14.85)**	-0.56 (-11.23)**
Income RP	0.08 (1.98)*	0.13 (3.45)**	0.13 (3.33)**
Lambda RP	0.06 (1.46)	0.23 (4.09)**	0.22 (3.30)**
Sigma RP	13.50 (20.93)**	13.54 (21.99)**	13.45 (21.01)**
k constant SP	0.90 [23.36] (-0.87)	1	1
k price SP	0.63 [0.50] (-4.42)**	1	1
k income SP	2.69 [0.22] (1.37)	1	1
k lambda SP	8.33 [0.50] (1.34)	1	1
k sigma SP	0.98 [13.23] (-0.35)	1	1.03 [13.85] (0.49)
Rho	0.74 (21.07)**	0.72 (18.22)**	0.71 (19.61)**
-log L	1374.89	1401.50	1401.42
CS RP	114.34	165.71	161.58
CS SP	180.53	165.71	161.58

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.12: 6,7,12 Megazone: RP and SP correlated, discrete SP data

	Unrestricted	All k's equal one	Heteroscedasticity
Constant RP	25.83 (10.62)**	22.60 (12.21)**	22.41 (10.49)**
Price RP	-0.79 (-9.52)**	-0.62 (-13.94)**	-0.61 (-8.15)**
Income RP	0.06 (1.70)	0.10 (2.71)**	0.10 (2.69)**
Lambda RP	0.04 (1.10)	0.14 (2.96)**	0.14 (2.53)*
Sigma RP	13.34 (21.84)**	13.18 (21.22)**	13.18 (22.85)**
k constant SP	1.03 [26.60] (0.17)	1	1
k price SP	0.77 [-0.61] (-1.94)	1	1
k income SP	2.94 [0.18] (1.04)	1	1
k lambda SP	8.37 [0.33] (0.95)	1	1
k sigma SP	0.98 [13.07] not estimated	1	0.98 [12.92] (-0.13)
Rho	0.54 (7.65)**	0.52 (7.23)**	0.52 (7.46)**
-log L	1132.56	1140.67	1140.66
CS RP	114.30	146.15	148.10
CS SP	148.22	146.15	148.10

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.13: 9,10,11 Megazone: RP and SP uncorrelated, continuous SP data

	Independent
Constant RP	20.95 (15.18)**
Price RP	-0.78 (-15.15)**
Income RP	0.04 (2.25)*
Lambda RP	0.01 (0.48)
Sigma RP	13.47 (40.81)**
Constant SP[a]	10.31 (5.70)**
Price SP	-0.49 (-11.59)**
Income SP	0.22 (8.23)**
Lambda SP	0.35 (6.09)**
Sigma SP	14.58 (22.78)**
-log L	5548.33
CS RP	101.56
CS SP	163.89

[a] These stated preference estimates are not estimates of the k^{SP} parameters. Rather, they are estimates of the parameters themselves, with the standard t-statistic in parenthesis.
** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.14: 9,10,11 Megazone: RP and SP uncorrelated, discrete SP data

	Independent
Constant RP	20.95 (15.18)**
Price RP	-0.78 (-15.15)**
Income RP	0.04 (2.25)*
Lambda RP	0.01 (0.48)
Sigma RP	13.47 (40.81)**
Constant SP[a]	10.89 (4.98)**
Price SP	-0.53 (-12.14)**
Income SP	0.29 (9.21)**
Lambda SP	0.39 (6.77)**
Sigma SP	14.51 not estimated
-log L	4467.32
CS RP	101.56
CS SP	cannot be estimated

[a] These stated preference estimates are not estimates of the k^{SP} parameters. Rather, they are estimates of the parameters themselves, with the standard t-statistic in parenthesis.
** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.15: 9,10,11 Megazone: RP and SP correlated, continuous SP data

	Unrestricted	All k's equal one	Heteroscedasticity
Constant RP	18.62 (13.97)**	15.75 (17.62)**	15.04 (14.75)**
Price RP	-0.69 (-14.11)**	-0.58 (-27.09)**	-0.55 (-20.21)**
Income RP	0.09 (4.37)**	0.13 (7.56)**	0.14 (8.03)**
Lambda RP	0.07 (2.91)**	0.17 (6.87)**	0.19 (6.51)**
Sigma RP	13.75 (44.75)**	13.82 (44.38)**	14.00 (42.42)**
k constant SP	0.58 [10.80] (-5.04)**	1	1
k price SP	0.66 [-0.46] (-5.77)**	1	1
k income SP	2.31 [0.21] (2.69)**	1	1
k lambda SP	4.66 [0.33] (2.42)*	1	1
k sigma SP	0.98 [13.48] (-0.59)	1	0.95 [13.30] (-1.63)
Rho	0.65 (34.07)**	0.65 (30.18)**	0.65 (31.13)**
-log L	5356.34	5381.35	5380.40
CS RP	113.97	136.92	143.75
CS SP	172.17	136.92	143.75

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.16: 9,10,11 Megazone: RP and SP correlated, discrete SP data

	Unrestricted	All k's equal one	Heteroscedasticity
Constant RP	19.78 (16.32)**	16.71 (17.67)**	15.33 (11.86)**
Price RP	-0.74 (-16.19)**	-0.63 (-26.48)**	-0.57 (-13.28)**
Income RP	0.07 (3.51)**	0.13 (7.00)**	0.14 (7.41)**
Lambda RP	0.04 (2.04)*	0.13 (5.74)**	0.16 (4.99)**
Sigma RP	13.74 (42.28)**	13.88 (38.90)**	13.95 (38.14)**
k constant SP	0.56 [11.08] (-4.50)**	1	1
k price SP	0.69 [-0.51] (-5.08)**	1	1
k income SP	3.91 [0.27] (2.73)**	1	1
k lambda SP	8.28 [0.33] (1.85)	1	1
k sigma SP	0.98 [13.47] not estimated	1	0.88 [12.28] (1.56)
Rho	0.53 (18.30)**	0.52 (16.12)**	0.53 (15.29)**
-log L	4379.85	4410.14	4409.19
CS RP	106.17	125.39	137.59
CS SP	153.74	125.39	137.59

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.17: 13,14,15 Megazone: RP and SP uncorrelated, continuous SP data

	Independent
Constant RP	24.36 (10.37)**
Price RP	-0.94 (-9.76)**
Income RP	-0.03 (-0.76)
Lambda RP	-0.06 (-2.21)*
Sigma RP	12.54 (21.33)**
Constant SP[a]	19.08 (4.15)**
Price SP	-0.61 (-6.42)**
Income SP	0.16 (2.77)**
Lambda SP	0.25 (2.73)**
Sigma SP	16.69 (11.34)**
-log L	1405.96
CS RP	95.46
CS SP	149.35

[a] These stated preference estimates are not estimates of the k^{SP} parameters. Rather, they are estimates of the parameters themselves, with the standard t-statistic in parenthesis.
** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.18: 13,14,15 Megazone: RP and SP uncorrelated, discrete SP data

	Independent
Constant RP	24.36 (10.37)**
Price RP	-0.94 (-9.76)**
Income RP	-0.03 (-0.76)
Lambda RP	-0.06 (-2.21)*
Sigma RP	12.54 (21.33)**
k constant SP	0.93 [22.72] (809.95)**
k price SP	0.76 [-0.72] (-8.01)*
k income SP	-10.92 [0.30] (-396.37)**
k lambda SP	-5.26 [0.34] (-208.03)**
k sigma SP	1.33 [16.69] not estimated
-log L	1105.13
CS RP	94.68
CS SP	cannot be estimated

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.19: 13,14,15 Megazone: RP and SP correlated, continuous SP data

	Unrestricted	All k's equal one	Heteroscedasticity
Constant RP	19.61 (8.83)**	14.39 (8.68)**	14.92 (8.32)**
Price RP	-0.72 (-9.66)**	-0.44 (-12.14)**	-0.46 (-9.11)**
Income RP	0.06 (1.83)	0.15 (4.20)**	0.14 (3.97)**
Lambda RP	0.05 (1.97)*	0.31 (3.98)**	0.28 (2.99)**
Sigma RP	12.88 (21.70)**	13.54 (20.42)**	13.30 (19.86)**
k constant SP	0.78 [15.30] (-1.42)	1	1
k price SP	0.62 [-0.54] (-3.95)**	1	1
k income SP	2.77 [0.17] (1.31)	1	1
k lambda SP	8.05 [0.40] (1.98)*	1	1
k sigma SP	1.12 [14.43] (1.55)	1	1.06 [14.10] (0.92)
Rho	0.69 (16.41)**	0.72 (16.83)**	0.71 (15.45)**
-log L	1356.63	1366.07	1365.70
CS RP	123.83	202.37	191.22
CS SP	201.03	202.37	191.22

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

Table A4.20: 13,14,15 Megazone: RP and SP correlated, discrete SP data

	Unrestricted	All k's equal one	Heteroscedasticity
Constant RP	22.63 (746.80)**	16.83 (9.87)**	15.46 (7.27)**
Price RP	-0.87 (-28.65)**	-0.56 (-12.62)**	-0.49 (-6.34)**
Income RP	0.01 (0.34)	0.11 (3.36)**	0.12 (3.67)**
Lambda RP	-0.02 (-0.77)	0.15 (3.02)**	0.20 (2.50)*
Sigma RP	12.74 (34.50)**	12.83 (20.98)**	12.94 (20.78)**
k constant SP	0.60 [13.58] (-13.11)**	1	1
k price SP	0.48 [-0.42] (-17.14)**	1	1
k income SP	18.27 [0.18] (569.83)**	1	1
k lambda SP	-17.18 [0.34] (-599.90)**	1	1
k sigma SP	0.82 [10.45] not estimated	1	0.85 [11.00] (-1.04)
Rho	0.50 (45.79)**	0.48 (6.66)**	0.50 (6.21)**
-log L	1089.10	1104.70	1104.32
CS RP	102.20	158.55	179.93
CS SP	212.61	158.55	179.93

** Denotes significance at the 99% confidence level
* Denotes significance at the 95% confidence level

APPENDIX 5. HYPOTHESIS TESTS

Table A5.1: 1,2,3 Megazone: Hypothesis tests

	RP/SPc	RP/SPd
Test 1(a) and 1(b):	20.06	19.08
General Consistency	reject	reject
Test 2:	19.08	18.00
Heteroscedasticity	reject	reject
Test 3:	18.28	15.70
Variance and Income	reject	reject
Test 4:	16.98	14.66
Variance and Price	reject	reject
Test 5:	18.68	17.82
Variance and Time	reject	reject

Table A5.2: 4,5,8 Megazone: Hypothesis tests

	RP/SPc	RP/SPd
Test 1(a) and 1(b):	26.37	23.48
General Consistency	reject	reject
Test 2:	26.26	19.94
Heteroscedasticity	reject	reject
Test 3:	25.96	19.88
Variance and Income	reject	reject
Test 4:	25.89	19.94
Variance and Price	reject	reject
Test 5:	14.08	15.69
Variance and Time	reject	reject

Table A5.3: 6,7,12 Megazone: Hypothesis tests

	RP/SPc	RP/SPd
Test 1(a) and 1(b):	53.22	16.22
General Consistency	reject	reject
Test 2:	53.06	16.20
Heteroscedasticity	reject	reject
Test 3:	33.04	15.82
Variance and Income	reject	reject
Test 4:	31.70	13.28
Variance and Price	reject	reject
Test 5:	49.75	15.68
Variance and Time	reject	reject

Table A5.4: 9,10,11 Megazone: Hypothesis tests

	RP/SPc	RP/SPd
Test 1(a) and 1(b):	50.02	60.58
General Consistency	reject	reject
Test 2:	48.12	58.68
Heteroscedasticity	reject	reject
Test 3:	47.42	58.54
Variance and Income	reject	reject
Test 4:	45.02	57.98
Variance and Price	reject	reject
Test 5:	42.88	54.31
Variance and Time	reject	reject

Table A5.5: 13,14,15 Megazone: Hypothesis tests

	RP/SPc	RP/SPd
Test 1(a) and 1(b): General Consistency	18.88 reject	31.20 reject
Test 2: Heteroscedasticity	18.14 reject	30.44 reject
Test 3: Variance and Income	17.78 reject	27.80 reject
Test 4: Variance and Price	17.78 reject	26.88 reject
Test 5: Variance and Time	16.96 reject	30.25 reject

REFERENCES

Adamowicz, W., Louviere, J., and Williams, M. "Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities." *Journal of Environmental Economics and Management* 10(1994): 271-292.

Adamowicz, W., Swait, J., Boxall, P., Louviere, J., and Williams, M. "Perceptions Versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation" Presented at the Fourth Annual Meeting of the Canadian Resource and Environmental Economics Study Group 1996.

Andreoni, J. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy* 97(1989): 1447-1458.

Bockstael, N., Strand, I., and Hanemann, M. "Time and the Recreational Demand Model." *American Journal of Agricultural Economics* 69(1987): 293-302

Brownstone, D., Bunch, D, and Train, K. "Joint Mixed Logit Models of Stated and Revealed Preferences for Alternative-fuel Vehicles." *Transportation Research*, (1998): 1-35.

Cameron, T. "A New Paradigm for Valuing Non-market Goods Using Referendum Data: Maximum Likelihood Estimation by Censored Logistic Regression." *Journal of Environmental Economics and Management* 15(1988): 355-379.

Cameron, T. "Combining Contingent Valuation and Travel cost Data for the Valuation of Nonmarket Goods." *Land Economics* 68(1992): 302-317.

Cameron, T., Poe, G., Ethier, R., and Schulze, W., "Alternative Nonmarket value-elicitation Methods: Are the Underlying Preferences the Same?" Paper presented at 1999 W133 Conference in Tucson, AZ.

Cesario, F. J., "Value of Time in Recreation Benefit Studies." *Land Economics* 52(1976):32-41.

Crooker, J. "Valuing environmental amenities with nonparametric and semiparametric methods" Doctoral Dissertation, Iowa State University, (1998).

Diamond, P., and Hausman, J. "Contingent Valuation: Is Some Number Better than No Number?" *Journal of Economic Perspectives* 8(1994): 45-64.

Dickie, M., Fisher, A., and Gerking, S. "Market Transactions and Hypothetical Demand Data: A Comparative Study." *Journal of the American Statistical Association* 82(1987): 69-75.

Feather, P. and Shaw, W. D. "Estimating the Cost of Leisure Time for Recreation Demand Models." *Journal of Environmental Economics and Management* 38(1999): 49-65.

Hanemann, W. "Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses." *American Journal of Agricultural Economics* 66(1984): 332-341.

Heckman, J. "Shadow Prices, Market Wages, and Labor Supply." *Econometrica* 42(1974): 679-694.

Huang, J., Haab, T., and Whitehead, J. "Willingness to Pay for Quality Improvements: Should Revealed and Stated Preference Data be Combined?" Working paper 1997.

Kahneman, D. and Ritov, I. "Determinants of Stated Willingness to Pay for Public Goods: A Study in the Headline Method." Mimeo, Department of Psychology, University of California, Berkeley, 1993.

Larson, D. "Testing Consistency of Direct and Indirect Methods for Valuing Nonmarket Goods." Working Paper 1990.

Layman, R., Boyce, R., and Criddle, K. "Economic Valuation of the Chinook salmon Sport Fishery of the Gulkana River, Alaska, under current and Alternate Management Plans." *Land Economics* 72(1996): 113-128.

Loomis, J. "The Evolution of a More Rigorous Approach to Benefit Transfer: Benefit Function Transfer." *Water Resources Research* 28(1992): 701-705.

Loomis, J. "Panel Estimators to Combine Revealed and Stated Preference Dichotomous Choice Data." *Journal of Agricultural and Resource Economics* 22(1997): 233-245.

Louviere, J. "Combining Revealed and Stated Preference Data: The Rescaling Revolution." Prepared for the Association of Environmental and Resource Economists 1996.

McConnell, K., and Strand, I. "Measuring the Cost of Time in Recreation Demand Analysis: An Application to Sportfishing." *American Journal of Agricultural Economics* 63(1981): 153-156.

McConnell, K., Weninger, Q., and Strand, I. "Joint Estimation of Contingent Valuation and Truncated Recreational Demands." *Valuing Recreation and the Environment,* J. Herriges and C. Kling (eds.), Elger (1999).

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, New York, North Holland (1978).

Niklitschek, M. and Leon, J. "Combining Intended Demand and Yes/No Responses in the Estimation of contingent Valuation Models." *Journal of Environmental Economics and Management* 31(1996): 387-402.

PC Miler User's Guide, Princeton, New Jersey, ALK Associates, Inc., 1998.

Randall, A. "A Difficulty with the Travel Cost Method." *Land Economics* 70(1994): 88-96.

Smith, V. Kerry and Kaoru, Y. "Signals or Noise? Explaining the Variation in Recreation Benefit Estimates." *American Journal of Agricultural Economics* (1990): 419-433.

Wilman, E. "The Value of Time in Recreation Benefit Studies." *Journal of Environmental Economics and Management* 7 (1980): 272-286.