ESSAYS ON DEMAND FOR WATER-BASED RECREATION IN OKLAHOMA

By

PHUMSITH MAHASUWEERACHAI

Bachelor of Economics

Khon Kaen University

Khon Kaen, Thailand

1998

Master of Economics

National Institute of Development Administration

Bangkok, Thailand

2003

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2010

ESSAYS ON DEMAND FOR WATER-BASED

RECREATION IN OKLAHOMA

Dissertation Approved:

Dr. Tracy A. Boyer

Dissertation Adviser

Dr. Larry D. Sanders

Dr. Jayson L. Lusk

Dr. Brian E. Whitacre

Dr. Lowell M. Caneday

Dr. A. Gordon Emslie

Dean of the Graduate College

ACKNOWLEDGMENTS

Finally, my long journey of finishing dissertation becomes true. Many people have supported me to complete this dissertation. I would like to express my gratitude to these individuals for their support and assistance. First, there is my mentor who has taught and guided me how to stay on track for this long journey: I would especially like to gratefully and sincerely acknowledge my advisor, Dr. Tracy A. Boyer. She has been a strong and supportive advisor to me throughout my graduate school career. During my journey, the road is not always smooth. However, her mentorship has guided me how to deal and pass the challenge times and provided me a well rounded experience consistent my long-term career goals.

I would like also to thank all of my committee members. Dr. Jason L. Lusk, who has provided helpful suggestions for the survey design and econometrics techniques used in this dissertation. Dr. Brian E. Whitacre, who has always stood beside me whenever I need his help and support. Finally, I would also like to thank Dr. Larry D. Sanders and Dr. Lowell M. Caneday for practical advice as well as inspiration.

Additionally, I would like to thank the Department of Agricultural Economics, Oklahoma State University, for giving me a chance to prove myself with this long and challenging journey. Both the academic and friendship experiences during the time I have been here were priceless.

I would also gratefully acknowledge the funding support that I have received while working on this dissertation. In particular, I thank the Oklahoma Water Resource Board that had provided funding support for this project. In addition, I also thank the Graduate College, Oklahoma State University that had provided me the fellowship when I work on my dissertation.

Additionally, I would like to thank my father and mother back in Thailand for their love. They are my inspiration and their hearts are always with me any time. Finally, I am thankful for my beloved wife for supporting me every direction I have turned. Without her support and love, my long journey would have not become true.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Introduction and Problem Statement

The state of Oklahoma has over 300 lakes, more man-made lakes than any other state, with over one million surface acres of water (Oklahoma Tourism and Recreation Department, 2007). Many of lakes serve several uses such as hydroelectric power, flood control, agriculture, and recreation. Since the mid 1950s, demand for lake recreation in Oklahoma has increased continuously due to the convenience of transportation, communication, and other new technology such as types of vehicles, and types of new watercrafts available to public (Caneday, 2000). The outdoor recreation business was reported as one of the fastest growing businesses in Oklahoma (Caneday et al., 2007). Even though the demand for lake recreation in Oklahoma is increasing, few recent studies have analyzed the demand for lake recreation as well as welfare effects from lake use in term of recreation. Lenard and Badger (1979) studied about lake recreation in Oklahoma in 1977. However, this study focused on the business impact of lake recreation only. Caneday and Jordan (2003) studied the behavior of Oklahomans traveling to state parks, but they did not estimate demand and economic value for water based amenities such as quality and quantity or estimate total visitation across all water-oriented recreational activities. Therefore, currently, there is no comprehensive explanation for lake recreational demand in Oklahoma.

Research objectives

This study proposes to determine the relative value of lake recreation in Oklahoma. The research performed in this study will focus on answering the following questions, "What factors influence demand for lake recreation?", "How do we forecast the number of lake recreational trips"?, "How much does willingness to pay for recreation change according to lake quality improvements?", and "How the potential management changes factors of Close-to-Home Fishing Program (CTHFP) influence anglers preferences?" Answers to these questions will help many interested groups to clearly understand what factors influence lake recreational demand and how this impacts visits to them. In addition, these findings will help policy makers make more informed decisions regarding the Oklahoma state water plan and current and future management scenarios for lakes in Oklahoma.

Because of the questions mentioned earlier, this dissertation will be separated into three papers.[1]

1. The first paper empirically compares the out-of-sample predictive ability of joint revealed preference (RP) and stated preference (SP) model to individual RP and SP models in case of prediction actual and hypothetical trip numbers taken by lake recreationists.
2. The second paper estimates the two-step model with combined RP and SP data to estimate the link between site choice selection and the number of trip

[1] The first two papers used the same data set, while the third paper used another data set.

taken, which would allow us to determine welfare changes from lakes quality improvement in term of changes in numbers of trips and per choice occasion.

3. The third paper estimates management of urban recreation use using a discrete choice experiment.

Valuation Environmental and Natural Resources

Services provided by environmental and natural resources often fall outside of the market's pricing system such that they are called non-market goods. Markets often inefficiently allocate environmental and natural resources because their property rights are not clearly defined. Often actions of private individuals impose external costs upon others' use of non market goods in ways for which they are not compensated, something called a negative externality. In addition, because there is no clear no market system to value most non-market goods, it is difficult to place a value on them and to efficiently manage them. Because private markets often underprovide public goods or do not adequately protect them, government agencies often must justify actions to manage natural resources using costs-benefit analysis. To deal with benefits and costs analysis, the first step is to measure the benefits of the existence or improvement of a non-market resource, an activity called non-market valuation. .

There are a number of methodologies used to value the non-market goods. These methodologies can be classified as revealed preference (RP) approaches and stated preference (SP) approaches. RP approaches use actual behavioral data to value the non-market goods. Researchers observe individual behavior in response to changes in quantity and quality of the non-market assets, and use this behavior to attempt to value them.

Travel cost and hedonic price methods are the common RP methods. Instead of using behavioral data, SP approaches rely on hypothetical data to value non-market goods. Respondents are directly asked to answer hypothetical questions that model tradeoffs between changes in their attributes and some monetary measure paid by respondents. Contingent valuation and discrete choice methods are the example of SP approach. In the next section, the details of travel cost method and discrete choice method, which are applied in this study, will be discussed.

Revealed Preference Approach

Travel Cost Method

The travel cost method is the oldest method used to value the environmental and natural resources (Kjaer, 2005). It has usually been applied for valuing recreational demand such as hunting, fishing, and forest visitation. The travel cost method is a method that uses variations in travel costs to a recreational site to estimate the demand for that site. Specifically, although the experience to visit a recreational site has no market to value its price, the costs incurred by visitors to visit a site can be used as surrogate values for that resource. These costs usually include travel costs, entry fees, and time costs. The rationale behind the travel cost method is that if the price of visit a site (i.e. cost of travel) increases, the visit rate tends to fall (Hanley and Spash, 1993). By using regression analysis to estimate the relationship between these two variables, it is possible to construct the demand curve and hence consumer surplus from visiting particular recreational site.

The travel cost method also assumes that there is the weak complementarity between the environment goods and consumption expenditure. This can imply that when consumption expenditure is zero, the marginal utility of environmental goods is also zero. If fishing trip in Illinois River is too expensive and nobody takes a trip to this river, the marginal social cost of a decrease in the quality of this river is also zero, for example. Hence, from this assumption, the travel cost method can only estimate use-value, but it cannot estimate non-use value. This is one of the problems of using travel cost method for valutation of environmental and natural resources in that cannot estimate the total use value which includes use and non-use values. , Non-use value, refers to the value that people have (WTP) for specific goods (i.e. rivers, and forest) to keep them available even though they have never used or plan to use them (Tietenberg and Lewis, 2009). Time costs are also the vexing problem for the travel cost method. It is difficult to specify what exactly value of time to individuals for recreational activities. The common approach to deal with time costs is to value time at fixed percentage of the wage rate. There is, however, another question of whether just travel time should be included, or whether on-site time should be included as well (Randall, 1994; Kjaer, 2005). Functional form is also another concern of the travel cost method. A variety of functional forms has been used (i.e. linear, log-log, and quadratic) for travel cost model studies. The difference functional form can produce large changes in consumer’s surplus estimates from a given data set (Hanley, 1989).

Discrete Choice: A Stated Preference Approach

Discrete Choice Method

Choice techniques have been introduced to the marketing field since the early 1970s (Kjaer, 2005). It is one of several versions of the method known as conjoint analysis in marketing. Not long after its introduction to marketing field, economists started to apply this technique to fit with economic theory, known today as random utility theory (Ben-Akiwa, and Lerman, 1985). The development of random utility theory became the benchmark for the use of choice technique in economics because it provides the linkage between observed consumer behavior and economic theory (Kjaer, 2005). Because this technique used in economics form is relied on the random utility theory, the new term that separates this technique applied in economic field from other fields is "discrete choice experiment" (DCE) (Ryan and Wordsworth, 2000). This technique was introduced to the environmental economic literature in the early 1990s (Hanley et al., 2003).

Even though, the DCE technique uses surveys to ask respondents hypothetical questions such as in the contingent valuation method (CVM), the DCE is able to compare multiple options with different attributes such that the marginal rate of substitution between goods is able to be calculated. In a discrete choice experiment, respondents have choice sets comprised with two or more alternatives, which vary along several characteristics or attributes of interest, and they are asked to choose one alternative. This allows the researcher to break down the preferences of respondents by each attribute instead of just entire products or situations. Another advantage of DCE is that it could provide more information than CVM by inducing more choices for each respondent. This would also imply more information about the respondent's preferences, and hence the better precision on preferences parameter estimates (Habb and McConnell, 2002).

Combined Revealed and Stated Preference Data

As shown in the previous section, travel cost and discrete choice methods represent a subset of the RP and SP approaches respectively that have been employed to value non market goods and amenities. In the past, the data from these two approaches are separately used to estimate the value of non market goods and amenities. However, since Swait and Louviere (1993) developed the method that allows jointly estimated of different data sets, much attention has been focused on combining RP and SP data in order to reduce hypothetical bias which may occur in the latter and improve the accuracy of valuation estimates.

Functional form problems and variable inclusion, which can create a multicolinearity problem, are serious concerns in using the travel cost and hedonic price analysis methods (Azevedo et. al., 2003). Moreover, RP approaches have been based on real behavioral data. If the quality and quantity change of amenities of non-market goods go beyond the experience of a set of respondents or the variation in the data set, models that are based on RP data may not correctly value the new environmental quality and quantity of non-market goods. SP approaches avoid these concerns because respondents may be asked hypothetical questions that are outside the current set of experiences. In addition, in a discrete choice experiment, the use of factorial statistical designs results in orthogonal attribute data, which can avoid the multicollinearity problem (Earnhart, 2001). However, the SP approach has been criticized because it does not rely on actual behavior. When people answer the hypothetical questions, they may not understand or lack experience about the things asked in the questions. Furthermore, they may ignore or downplay their budget constraint when answering the questions (Swait and Louviere, 1993; Louviere et al., 2000). These may cause hypothetical bias in the valuation estimates.

Due to the drawbacks of each approach, economists have begun to combine RP and SP data. The benefits from combining them are as follows. The SP information could provide information of consumer preferences which cannot be observed in the market. Moreover, multicolinearity problems could be reduced by SP data so the attribute effects that were previously unidentified due to multicolinearity problems can now be identified (Adamowicz et. al, 1994). And, the estimators from the jointly estimated model still rely on true parameters because they are also based on real behavior from RP information.

Data description[2]

Data for papers one and two were collected by mailed survey on Oklahoma Lake Use (2007) for travel cost and discrete choice experiments. IRB approval was obtained with approval number AG0734 (the IRB approval letter is in Appendix). Data on travel distances and lake characteristics were compiled from GIS maps from Oklahoma Water Resource Board (OWRB), which was created by Caneday and Jordan (2003), lakes website, and phone interviews with lake managers.

The survey was mailed to 2,000 individuals, who were randomly chosen, in every county of Oklahoma State during fall 2007. A random sample was obtained from Survey Sampling Inc, Fairfield CT stratified across 6 regions of Oklahoma. The survey was first distributed during the last week of September 2007 by mail. Standard Dillman procedures

[2] This part provides detail only about the data set used in the first and the second papers. The detail of data set used in the third paper is provided directly in the third paper.

were used to get the highest possible response rate (Dillman, 2000). The letter, postcard reminder, and follow up letters are provided in Appendix. The survey and cover letter, which explained the importance and objective of this survey, were mailed to 2,000 recipients. Two weeks after the survey was mailed; the postcard reminder was mailed to people who had not responded. Then, two weeks later, the follow up survey with cover letter was mailed individual who had not reply to the survey. Following this method, from 2,000 surveys, 401 were returned. Thirty-nine of them were unusable and allowing for 150 undeliverable surveys due to no forwarding addresses, the net response rate was 19.57 percent. The descriptive statistics of these respondents are shown in Table 1.1.[3] There are two groups of respondents used in this study. The first group is the respondents who have experienced visiting a lake(s) before. This group is later referred as current lake recreationists. The second group of respondents has never visited a lake before, but they answered the discrete choice questions about potential visits. This group of respondents later is referred to as potential lake recreationists. Since the purpose of this paper was to combine RP and SP data, the survey was designed to obtain both types of data.

Revealed preference data

Respondents were asked to report their visitation patterns for single-day trips to 144 public lakes in Oklahoma in 2007. They were also asked to report their activities in lakes as well as features of lakes that are important to them. In order to obtain the effect of water quality on lake recreation demand, the water quality data were gathered from the Beneficial Use Monitoring Program (BUMP) database of OWRB. Other amenity data

[3] 49 respondents who did not answer the discrete choice questions were not used in this study. Therefore, the total number of respondent used is 313.

was collected for each lake including the types and numbers of restrooms, docks, campsites and boat ramps, etc. These amenity data were collected from the lake websites and/ or by phone interview. TransCAD software was used to calculate the distance from each ZIP code to 144 lakes via roads. Then, the distances were expressed as round trip travel cost, which was combined with out-of-pocket expenditure and opportunity cost of time.[4]

Stated preference data

The survey also solicited SP data. Each respondent faced two discrete choice sets which presented possible alternative lake recreational opportunities at differing lake amenity levels and distances. These choice sets were orthogonally designed to estimate the willingness to pay for quality and amenity improvements at a lake similar to the lake respondents most often visited (which they indicated in the RP portion of the same survey). The SP questions elicited lake visitor preferences for lake characteristics, including availability of lake amenities and distance. Six measurable attributes associated with lake recreation experiences of either 2 or 6 levels were determined (Table 1.2). This created $4^3 \times 3 \times 2 \times 6 = 2{,}304$ possible combinations. Each combination was then randomly paired with another combination (Lusk and Norwood, 2005). The third option was stated as the respondents most frequently visited lake as given in the revealed preference data.

[4] The out-of-pocket expenditure was estimated by multiplying distance with \$0.48/ mile, which was reported by AAA 2006, and the opportunity cost of time was calculated as one third of an hourly individual's wage rate time by travel time, which was assumed speed of 50 mile/hour (Haener et. al., 2001, and Boxall et. al. 2003).

Each respondent was asked to answer two experimental choice questions. Each of them contains two options of hypothetical lake choice (Figure 1.1). Because some attributes of the SP question, number of boat ramps, water clarity, and distance, were asked by increasing in numbers, the information from lakes that were most visited by each respondent were used as the base information to adjust the levels of those attributes to be the same as RP data. For example, if Tenkiller Lake was the lake most visited by a respondent, the number of boat ramps in SP question was added to the actual number of boat ramps in Tenkiller Lake. Moreover, the SP questions also asked the number of trips respondents would take given the lake they choose from conjoint choice question. This would allow us to determine the number of trips they would take under the hypothetical situation.

Table 1.1. Descriptive Statistics of Current and Potential Lake Recreationists (Percentage by Category)

Variable	Current Lake User	Potential Lake User
Yearly income		
< 20000	8.20%	14.50%
20000-39999	26.40%	36.64%
40000-59999	21.40%	36.64%
60000-99999	29.70%	15.27%
> 100000	14.30%	14.50%
Age		
< 26	2.75%	0.76%
26-34	10.99%	3.05%
35-49	30.22%	19.85%
50-59	25.27%	33.59%
> 60	30.77%	42.75%
Education level		
< High school	3.29%	2.17%
High school	18.14%	25.83%
Some college/ Vocational school	33.52%	35.01%
College graduate	29.67%	26.51%
Advanced degree	15.38%	10.48%
Gender		
Male	68.70%	50.40%
Female	31.30%	49.60%
Number of respondents	182	131

Table 1.2. Attributes and Levels in the Lake Recreation Discrete Choice Survey

Attribute	Factor Levels
Increase in public boat ramp	None
	1 Boat ramp
	2 Boat ramp
	3 Boat ramp
Campsites	None
	Available
	Available with electric service
Public restroom	None
	Porta-potties/ Pit toilets
	Restroom with flush toilets
	Restroom with flush toilets and showers
Lodge	None
	Available
Water clarity	No improvement
	1 foot increase of water visibility dept from surface
	2 foot increase of water visibility dept from surface
	3 foot increase of water visibility dept from surface
Increase in distance from home	0 miles increase
(one-way)	10 miles increase
	20 miles increase
	30 miles increase
	40 miles increase
	50 miles increase

Figure 1.1. An Example of a Discrete Choice Question

Compared to the lake you most visit, would you choose a lake such as A or B? Or would you choose to stay with the one you currently visit, C? Please choose one.

<table>
<tr><th>Attribute</th><th>Option A</th><th>Option B</th><th>Option C</th></tr>
<tr><td>Increase in public boat ramps</td><td>2 Boat ramp</td><td>1 Boat ramp</td><td rowspan="6">NO CHANGE:
I would rather keep the management of this lake the way it is today</td></tr>
<tr><td>Campsites</td><td>Available with electric service</td><td>Available with electric service</td></tr>
<tr><td>Public restrooms</td><td>Restroom with flush toilets and showers</td><td>Restroom with flush toilets and showers</td></tr>
<tr><td>Lodges</td><td>None</td><td>Available</td></tr>
<tr><td>Water clarity</td><td>1 foot increase of water visibility dept from surface</td><td>No improvement</td></tr>
<tr><td>Increase in distance from home (one-way)</td><td>20 miles increase</td><td>40 miles increase</td></tr>
<tr><td>I would choose (Please check only one)</td><td>☐ A</td><td>☐ B</td><td>☐ C (I would not want either A or B)</td></tr>
</table>

Given your choice above, how many trips per year would you take?

Number of single day trips ☐ same number or ___#less or ___# more

CHAPTER II

PREDICTIVE ABILITY: CAN WE RELY ON THE COMBINED REVEALED AND STATED PREFERENCE MODEL TO PREDICT THE FUTURE BEHAVIOR?

Introduction

The state of Oklahoma has over 300 lakes, more man-made lakes than any other state, with over one million surface acres of water (Oklahoma Tourism and Recreation Department, 2007). Many of these lakes are used for several reasons such as hydroelectric power, flood control, agriculture, and recreation. Some of these water uses can have either negative effect or positive effect on other uses. Water stored in a reservoir at a high level, for instance, could provide benefits for hydroelectric power and recreation; however, it could also reduce the supply of water available for agricultural activities.

Recent conflict over water use between agricultural and recreational uses during periods of prolonged drought in Oklahoma has driven home the need for valuation of non-market benefits of the state's extensive man-made reservoir network for the ongoing state water planning process. Valuing non-market goods using revealed preference (RP) and stated preference (SP) approaches involves tradeoffs in the reliability of the valuation estimate. The RP approach, which is based on actual behavioral data, is often assumed to provide a lower bound for willingness to pay (Louviere, Hensher et al., 2000). However, if the quality and quantity levels of proposed changes in amenities of non market goods go beyond the experience set of respondents, models based on RP data may not be able to predict how respondents prefer new management or quality upgrades (Morikawa, 1994; Hensher et al., 1999; Earnhart, 2001).

The stated preference approach avoids those concerns because researchers can ask hypothetical questions that contain quality and quantity of amenities outside the current set of respondents' experiences. In addition, in choice-based conjoint analysis, thanks to factorial statistical designs, the attribute level results in orthogonal attribute data, thus avoiding multicollinearity problems (Earnhart, 2001). However, the SP approach has been criticized because it does not rely on actual behavior. When people answer the hypothetical questions, they may not understand or lack experience about the things being valued or they may ignore their budget constraint when responding to the survey. These issues may cause bias in the estimators as well as over or under estimates of welfare. Due to the drawbacks of each approach, combining the RP and SP data could provide information on consumer preferences which cannot be observed in the market. Moreover, the attribute effects that were previously unidentified due to multicollinearity can now be identified (Adamowicz etal., 1994).

As welfare measures estimated by combining RP with SP methods have gained attention, consistency tests between both data have shown them to yield different results (e.g. Adamowicz et al., 1997; Whitehead et al., 2000; Earnhart, 2001; and Azevedo et al., 2003). In addition, some studies have focused on the in-sample tests of predictability to measure the benefit gained from combining the RP and SP data (e.g. Adamowicz et al., 1997; Verhoef and Franses, 2003). However, few studies especially in environmental economics, adopt an out-of-sample prediction as a test of gains from combining the RP and SP data. In-sample tests of predictability result in two main concerns in the economic literature. The first concern is that it would not be reliable on unmodelled structural change. Another concern is data mining, i.e., researchers may search for several alternative predictive models to find the model that fits well (Lo and MacKinlay, 1990; Foster et al., 1997). These two problems of in-sample prediction would lead to exaggerated predictive ability (West, 1996; Goyal and Welch, 2003). An out-of-sample prediction method, on the other hand, could avoid the spurious predictive ability because its estimated samples are different from the predicted samples, so structural changes that would not be captured by in-sample prediction would be captured by out-of-sample prediction (Foster et al., 1997). Hence, out-of-sample prediction would likely to provide a more accurate test of gains from combining the RP and SP data than by using in-sample prediction.

In addition, recent research that tested the predictive ability of joint models in the environmental economics literature used only RP data as a holdout sample for prediction of trip numbers (Ben-Akiva and Morikawa, 1990; Haener et al., 2001). However, to my knowledge, there is no study that also uses SP data as a holdout sample to predict trip numbers for testing the accuracy gained from combining RP and SP data. As mentioned by previous research SP data has provided useful information for prediction beyond the current market features (Louviere et al., 2000; Grijalva et al., 2002; Whitehead, 2005). Using SP data along with RP data as holdout samples to predict number of trips taken may provide insight into the predictive performance of the models in terms of actual behavior and future behavior.

In this paper the data used derives from a statewide survey of Oklahomans about lake recreation at 144 public lakes conducted in Fall 2007. The survey elicited information on all public lake trips statewide and also included an orthogonally designed set of discrete choice experiments to estimate willingness to pay for quality and amenity improvements at a lake similar to the lake respondents most often visited.

This paper augments the existing knowledge base of methodology for combining RP and SP data by (1) combining the RP data with SP data to estimate lake recreation demand, (2) comparing the out-of-sample predictive ability of the joint model with the travel cost model and discrete choice model for RP and SP holdout samples. Furthermore, this study will also examine the determinants of lake visitation in Oklahoma. These results will be of interest to individuals involved in non-market valuation seeking information regarding which models could give superior explanation and prediction. In addition, a solid understanding of the factors that affect lake visitation is of interest to policy makers seeking to improve lakes amenities management to match with lake visitor's preference.

Data Description

Data used in this study were collected by a mail survey entitled, "Oklahoma Lake Use" (2007) for travel cost and discrete choice experiments. Data on travel distances and lake characteristics were compiled from GIS maps from Oklahoma Water Resource Board (OWRB), which was created by Caneday and Jordan (2003), individual lake websites, and phone interviews with lake managers.

The survey was mailed to 2,000 individuals, who were randomly chosen, in every county of Oklahoma State during fall 2007. A random sample was obtained from Survey Sampling Inc, Fairfield CT stratified across 6 regions of Oklahoma. The survey was first distributed during the last week of September 2007 by mail. Standard Dillman procedures were used to get the highest possible response rate (Dillman, 2000). The survey with cover letter, which explained the importance and objective of this survey, was mailed to 2,000 recipients. Two weeks after the survey was mailed; the postcard reminder was mailed to people who had not responded. Then, two weeks later, the follow up survey with cover letter was mailed individuals who had not replied to the survey. As a result, 401 surveys out of 2000 were returned. Two hundred and eighteen of them were unusable and allowing for 150 undeliverable surveys due to no forwarding addresses, the net response rate was 10 percent.[5] Descriptive statistics of attribute levels and variables used are given in Table 2.1. Since both revealed and stated preferences data are used, the survey was designed to obtain both types of data.

Revealed Preference Data

Respondents were asked to report their visitation patterns for single-day trips to 144 public lakes in Oklahoma in 2007. They were also asked to report their activities in lakes as well as features of lakes that are important to them. In order to obtain the effect of

[5] Among 218 unusable surveys, 179 of them are actually completed survey, but these respondents have never visited a lake before. Since this study focuses on analyzing current lake users, we dropped these respondents from our sample.

water quality on lake recreation demand, water clarity was used as the proxy for water quality because lake recreationists often identify clear water ashigh quality water as an indicator of lack of contaminants or pathogens and ecosystem health (David et al., 1991; Azevedo et al. 2001; Caneday et al., 2001). Furthermore, detailed information on alternative chemical analysis or indices of water quality was not available statewide. Water clarity information was gathered from the Beneficial Use Monitoring Program (BUMP) database of OWRB (Beneficial Use Monitoring Program Report, 2007).[6] Other amenity data were collected for each lake including the types and numbers of restrooms, docks, campsites and boat ramps, etc. These amenity data were collected from the lake websites and/ or by phone interview. TransCAD software was used to calculate the distance from each ZIP code to 144 lakes via roads by assuming that respondents selected to travel by shortest path (TransCAD, 2008). Then, the distances were expressed as round trip travel cost, which was combined with out-of-pocket expenditure and opportunity cost of time.[7]

<u>Stated Preference Data</u>

The survey solicited SP data. Each respondent faced two discrete choice sets which presented possible alternative lake recreational opportunities at differing lake amenity levels and distance. These choice sets were orthogonally designed with quality and amenity improvements at a lake similar to the lake respondents most often visited (which they indicated in the RP portion of the same survey). The SP questions elicited visitors'

[6] The water quality data used is secchi disk depth that measures the distance under the surface of the water at which the disk is no longer visible.

[7] The out-of-pocket expenditure was estimated by multiplying distance with \$0.48/ mile, which was reported by AAA (2006), and the opportunity cost of time was calculated as one third of an hourly individual's wage rate time by travel time, which was assumed speed of 50 mile/hour (Haener et al., 2001; Boxall et al., 2003).

preferences for lake characteristics, including availability of lake amenities and distance. Six measurable attributes associated with lake recreation experiences of either 2 or 6 levels were determined (Table 1.2). This created $4^3 \times 3 \times 2 \times 6 = 2{,}304$ possible combinations. Each combination was then randomly paired with another combination (Lusk and Norwood, 2005). The third option was stated as the respondent's most frequently visited lake as given in the revealed preference data.

Each respondent was asked to answer two experimental choice questions. Each of them contains two options of hypothetical lakes (Figure 1.1). Because some attributes of the SP questions such as the number of boat ramps, water clarity, and distance, were asked as a quality improvement, i.e. an increase in amenities, the information from lakes that were most visited by each respondent was used as base information to adjust the levels of those attributes to be the same as RP data. For example, if Tenkiller Lake was the lake most visited by a respondent, the number of boat ramps in SP question was added by the actual number of boat ramps in Tenkiller Lake. Moreover, the SP questions also asked the number of trips respondents would take given the lake they choose from discrete choice question. This allows us to determine the number of trips they would take under a hypothetical situation.

Theory and Econometric Models

The conditional logit model is applied to analyze the choice between alternative lakes sites. The conditional logit model is based on a Random Utility Model (RUM) that assumes that lake visitors will choose the option (in this case, a lake) that provides them with the highest utility. However, in reality, the real utility of the respondent could not be specified. Only the indirect utility function of the respondent denoted as V can be observed, and the unobservable part or stochastic component of the utility that is unknown denoted as ε. Therefore, the utility can be represented as following

$$U = V + \varepsilon \tag{2.1}$$

where U is the real utility function. The indirect utility function would be revealed by either examining the respondent's actual behavior or the responses to the discrete choice questions in which the attributes are arguments. Hence, Vcan be expressed as a function of attributes accompanying each alternative

$$V_i = \alpha_i + \boldsymbol{\beta}_\mathrm{k}\mathbf{X}_\mathrm{i}, \qquad \forall_i \epsilon C \tag{2.2}$$

where $\mathbf{X}$ is the vector of k attributes, $\boldsymbol{\beta}$ is a coefficient vector, α is alternative specific constant (ASC), and i is an alternative in choice sets C. The probability that site i will be visited by a respondent is equal to the probability that the utility gained from selecting site i is greater than that from other sites. Let us assume the distribution of the stochastic component is independently and identically distributed (IID) according to the Gumbel random variable, so the probability of choosing choice i among those available (1, 2,…,k) ϵ C can be expressed in closed form as

$$\mathrm{Pr}_i = \frac{exp\left(\mu(\alpha_i + \boldsymbol{\beta}_k\mathbf{X}_i)\right)}{\sum_{k \epsilon C} exp\left(\mu(\alpha_k + \boldsymbol{\beta}_k\mathbf{X}_k)\right)} \tag{2.3}$$

where μ is a scale parameter. The scale parameter in case of single set of data could not be identified, so it is set equal to 1 (Boxall, Englin, and Adamowicz, 2003). From (2.3), the likelihood functions of individual RP and SP models have the following forms

(2.4) $$\text{RP: } L_{RP} = \sum_{n=1}^{N^{RP}} \sum_{P_i \epsilon C_n^{RP}} y_{in}^{RP} lnP_{in}^{RP}(\mathbf{X}_{in}^{RP}|\boldsymbol{\alpha}^{RP}, \boldsymbol{\beta}^{RP})$$

(2.5) $$\text{SP: } L_{SP} = \sum_{n=1}^{N^{SP}} \sum_{P_i \epsilon C_n^{SP}} y_{in}^{SP} \, lnP_{in}^{SP}\left(\mathbf{X}_{in}^{SP}|\boldsymbol{\alpha}^{SP}, \boldsymbol{\beta}^{SP}\right)$$

where y_{in}= 1 if a respondent selects choice i, y_{in} = 0 otherwise, n represents the index of respondents from the RP and SP data, $P_{in}^{RP}(\mathbf{X}_{in}^{RP}|\boldsymbol{\alpha}^{RP}, \boldsymbol{\beta}^{RP})$ and $P_{in}^{SP}\left(\mathbf{X}_{in}^{SP}|\boldsymbol{\alpha}^{SP}, \boldsymbol{\beta}^{SP}\right)$ are the probabilities of a respondent choosing choice i in the RP and SP samples, respectively.

When jointly estimating models from two or more data sources, the ratio of scale parameter should be identified. According to Louviere, Hensher, and Swait (2000), the ratio of scale factor is inversely related to the ratio of variance between two data sets. This relationship can be shown as follows:

(2.6) $$\frac{\sigma_{RP}^2}{\sigma_{SP}^2} = \frac{\pi^2/6\mu_{RP}^2}{\pi^2/6\mu_{SP}^2} = \left(\frac{\mu_{SP}}{\mu_{RP}}\right)^2 \quad ,$$

where σ^2 is variance of each data set. Following Louviere et al. (2000), the likelihood function of the pooled data is the sum of the conditional log likelihoods of RP and SP data that is showed as following

(2.7) $$L_{CRPSP} = \sum_{n=1}^{\mathrm{N^{RP}}} \sum_{\mathrm{P_i \epsilon C_n^{RP}}} y_{in}^{RP} lnP_{in}^{RP}\left(\mathbf{X}_{\mathrm{in}}^{\mathrm{RP}}|\boldsymbol{\alpha}^{\mathrm{RP}}, \boldsymbol{\beta}^{\mathrm{RP}}\right) + \sum_{n=1}^{\mathrm{N^{SP}}} \sum_{\mathrm{P_i \epsilon C_n^{SP}}} y_{in}^{SP} \, lnP_{in}^{SP}\left(\mathbf{X}_{\mathrm{in}}^{\mathrm{SP}}|\boldsymbol{\alpha}^{\mathrm{SP}}, \boldsymbol{\beta}^{\mathrm{SP}}, \mu_{SP}\right)$$

where μ_{SP} is the ratio of the scale parameter of SP data to the scale parameter of RP data.[8] Generally, y_{in}^{RP} and y_{in}^{SP} are 0 and 1. However, in the RP data for this analysis, each respondent can visit more than one site in choice set provided in the questionnaire. This may create an overweighting problem for each RP observation since the SP question is considered as one choice set and each respondent provides one response in each choice set. To solve this problem, equation (2.7) is also estimated by weighting the RP log likelihood function. Instead of coding y_{in}^{RP} as 0 and 1, it is weighted by trip proportions, and these proportions also add up to one over each RP choice set (Adamowicz et al., 1997; Haener et al., 2001). For example, if some respondents visited three different lakes, those three lakes will be weighted by one third and the rest of lakes are weighted by zero. In the SP choices, because each SP question is considered as one choice set, y_{in}^{SP} is still coded either 0 or 1 over each choice set. By weighting the data in this manner, the RP and SP observations are given equal weight.

To estimate model parameters, all coefficients of RP and SP are constrained to be equal and a full information maximum likelihood method will be employed to simultaneously optimize equation (2.7) with respect to all parameters.

Predictive Ability Tests

To improve the accuracy of predictability tests, the method of Haener et al. (2001) is applied. Thirty different estimation and holdout samples were randomly drawn from the data sets. However, in a departure from Haener et al. (2001), instead of randomly selecting estimation and holdout samples from the RP data only, the SP data also

[8] In order to find the relative scale factors, we normalize the inclusive value of parameter associated with RP data to unity.

randomly selected as holdout samples. This way there are two sets of holdout samples, which are the RP holdout sample, and the SP holdout sample**.** Each of them is predicted by RP model, SP model, and a combined RP and SP model.

Various predictive ability tests have been developed to measure the predictive accuracy of choice models. Some of them are based on an aggregate level test, while others operate at the individual level. In this study, both the aggregate level test and individual level test are applied.

The aggregate level test used for measuring the accuracy of out-of-sample prediction is the root mean square error (RMSE), which provides an idea of closeness of the prediction. The formula of RMSE is shown as following:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(T_n - \hat{T}_n)^2} \tag{2.8}$$

where N is the number of holdout sample, and T_n and $\hat{T}_n$ are the total numbers of observed trips and predicted trips of individual n, respectively.

Besides using an aggregate level test, an individual level test is also employed, which directly compares the individual's observed and predicted trips. Two individual level tests are used in this study. The first test is overall correlation coefficient between actual and predicted trips, r_i. The second test is the mean of individual correlation coefficient, developed by Haener, Boxall, and Adamoxicz (2001), which relies on the individual-specific correlation coefficients between observed and predicted trips. This test can be presented as follows

$$r_m = \frac{1}{N}\sum_{\mathrm{n=1}}^{\mathrm{N}}\frac{cov(X_{n,P},\ X_{n,O})}{\sqrt{var(X_{n,P})}\sqrt{var(X_{n,O})}} \tag{2.9}$$

where $X_{n,P}$ and $X_{n,O}$ are the vector of predicted and observed trips for individual n, respectively.

To estimate the predictive ability test statistics, each set of holdout sample is used to calculate the vectors of probabilities of choosing lake i for each respondent. After that, individual vectors of probabilities are multiplied by their total trip number to calculate the vector of predicted trip distribution for each respondent. The individual predicted trip distributions are used to calculate the overall correlation coefficient and individual-specific correlation coefficients, equation (2.9). In addition, to calculate the RMSE, the individual predicted trip distributions are summed across all individuals, and compared to the aggregate observed trip numbers.

After conducting these statistical tests for each holdout sample, the predictive ability of each model is ranked in each holdout sample by using 1 for the best prediction model, 2 for the second best prediction, and 3 for the poorest prediction. Then, these ranking are averaged for each model to clarify which models provide the best prediction for RP and SP holdout samples.

Estimation Results

Two sets of models are estimated. Each set contains three different models, which are the combined RP and SP model (CM model), the RP model (RP model), and the SP model (SP model). The first set is an unweighted model, for which the RP log likelihood function is not weighted by trip proportions. The second set is a weighted model, for which RP and SP choice sets are given equal weight. Thirty different estimations are estimated from thirty different estimation samples. To simplify the presentation and get the information about the parameter estimates, the results estimated from entire observation are represented in Table 2.2. Starting with unweighted model, most of coefficients in these three models are consistent with theory and previous research on lake recreation. For example, travel costs in those three models are negative, which implies that given other variables a lake located closer to an individual home has higher chance to be visited than a lake farther away. In addition, lakes with higher attribute quantities and quality such as numbers of boat ramp, availability of flush toilets with shower, and higher water clarity are preferred. For the unique variables of the RP data, the area attributes reveal that lakes located in Northeast, and Southeast regions are more preferred to lakes located in Northwest region, while lake recreationists may consider the quality of lakes located in Southwest region are the same as those located in Northwest region because its coefficient is not statistically significant. Major lakes, for which the surface area is more than 5,000 acres, are also preferred by lake recreationists. Generally, most parameters in the RP and SP models show the similar pattern of preferences across the attributes, except for the availability of campsite and availability of campsite with electricity. However, the effect of each attribute is quite different between these two models. This may imply the differences in the variances of RP and SP data. The RP and SP data are combined to estimate the CM model. The signs of most coefficients are consistent with RP and SP models, but the size of some coefficients, travel cost, boat ramp, and flush toilet, are clearly similar to those from the RP model. The relative scale parameter that takes into account the differences in variances of RP and SP data is also statistically significant. In addition, the value of the relative scale parameter is statistically less than one. This means that the variance of SP data is higher than that from RP data. However, when the equality between the vectors of RP and SP data after taking into account the scale differences is tested, the test find that the coefficient vectors between those two data are significantly different ($\chi^2 = 89.38, df = 9$).

Turning now to weighted model, this set also contains the RP model, SP model, and CM models.[9] Generally, the pattern of preferences across the attributes is similar as that estimated by unweighted model. In the RP model, even though, the estimated results show some different directions of the attributes on preferences when compared to those in unweighted RP model, the differences among these coefficients are not many; especially travel cost, boat ramp, flush-toilet with shower, and water clarity. When the estimated results between the RP and SP models are compared, most coefficients of these two models have similar patterns of preferences across the attributes. In addition, some coefficients of these two models are more similar than those estimated by unweighted model such as travel cost and flush-toilet with shower. In the case of CM model, the coefficients of travel cost and boat ramp are very similar as those from RP model, which is the same pattern as unweighted model. For the relative scale parameter, it reveals that the variances of RP and SP data may be different because it is significantly less than one. However, the test of equality of parameter vectors of these two data after taking into account the difference in variances also reveals that the coefficient vectors between RP and SP data are still significantly different ($\chi^2 = 66.42, df = 9$).[10]

[9] The SP model in this set is the same as that in unweighted model because each SP question is still considered as one choice set.

[10] The same test is also conducted when we estimate thirty estimate samples. We found that this test shows the coefficients vectors of RP and SP data are statistically indifferent for one third of estimate samples, while we found none in unweighted model.

The estimate results of the CM model also show that even when the RP data is weighted by trip proportion, the pattern of parameter estimates is similar to those from unweighted model in terms of both sign of coefficients and the effect of each attribute. This may be due to the fact that on average each respondent visits just about two lakes. Hence, on average the effect of weighting each RP choice set by trip proportion does not have much effect on parameter estimation.

Model Performance: Prediction Tests

The results of aggregate actual trip and hypothetical trip predictions of unweighted model from RP and SP representative holdout samples are represented in Table 2.3. The results of ten lakes are selected due to space limitations. Starting with the RP holdout sample, the prediction results show that the CM model and the RP model provide similar prediction results, and they seem likely to over predict for popular lakes such as Fort Gibson lake, Hefner lake, and Tenkiller lake. However, for lakes with less than 50 trips total in the sample, the performances of these two models clearly improve. In case of the SP model, its prediction performance is generally poorer than that from the CM and the RP model, especially for lakes with total trip less than 100. Turning now to the SP holdout sample, in each SP holdout sample, the hypothetical trip numbers are calculated by summing the total actual trip numbers of the lake most visited by each respondent with the addition trip numbers specified by each respondent under the hypothetical scenario. The performances of each model seem likely to be similar for prediction this holdout sample.

For the aggregate actual trip and hypothetical trip predictions of the weighted model, Table 2.4 reports the prediction results from the same representative holdout samples as Table 2.3. Generally, the results are similar as those represented in Table 2.3. The CM and RP models tend to overpredict trips for popular lakes, but they provide better prediction performances for lakes with total trip numbers less than 50. For the SP holdout sample, the prediction performances of these three models are similar. Furthermore, the prediction results of these models seem likely to be better than those from the unweighted model. However, the predictive ability tests, which are the aggregate level test and the individual level test, could provide clearer information which models provide superior prediction than the results of trip predictions reported in Table 2.3 and table 2.4.

The results of predictive ability tests for RP and SP holdout samples of unweighted models are represented in Table 2.5. Starting with the RP holdout sample, the RMSE shows that the CM model has the lowest error prediction (69.122), while the SP model shows poor performance for the actual behavior prediction, resulting in about eight times higher error predictions than CM model. In addition, the mean rank test clearly shows that the CM model provides the superior predictive performance for almost all RP holdout samples. In the same manner for the individual level prediction tests, the overall correlation coefficient, r_i, and the individual correlation coefficient, r_m, suggest that the CM model generates the best predictive ability, which results in the highest r_i (0.094) and r_m (0.220) values. The mean rank also strictly confirms that the CM model clearly dominates RP and SP models for almost RP holdout samples.

In the case of SP holdout sample, the SP model turns to provide the best prediction performance for aggregate level test (RMSE = 221.734), while the CM model generates the second best predictive ability for this holdout sample. The mean rank also shows that the SP model provides the best aggregate level predictive ability among thirty sets of SP holdout samples. However, the story is totally opposite for individual level prediction tests. The CM model turns to be the best model in term of predictive performance for individual level prediction. The CM model generates the highest r_i and r_m values, which are 0.392 and 0.976, respectively. Furthermore, for the r_i test, the mean rank values also show that the CM model dominates the RP and SP models for almost SP holdout samples. It also completely dominates RP and SP models in every SP holdout samples for r_m test.

The predictive ability tests for the weighted models were also conducted. The results of these tests are represented in table 2.6. The predictive ability for RP holdout sample of each model is similar to the unweighted models. Namely, the CM model still dominates RP and SP models in both aggregate level test and individual level tests. The aggregate prediction error of CM model is lower than that from other two models. In addition, the mean rank clearly shows that the CM model provides the most superior aggregate prediction for almost RP holdout sample. In the case of individual level tests, the mean rank confirms the individual predictive performance of CM model is preferred to RP and SP models for both r_i and r_m tests.

The results of predictive ability of these models in terms of forecasting the future behavior (SP holdout sample) are mixed. The SP model still provides the best aggregate level prediction in this case, while the prediction performance of CM and RP models are similar. However, the CM model becomes the best model for individual level prediction. For r_i test, the CM model provides the highest correlation, and also has the smallest mean rank. In the similar manner, the CM model completely dominates other models in case of r_m test for all SP holdout sample due to the mean rank result.

Discussion and Conclusion

This study finds interesting results in several ways. First, we expected that if the likelihood ratio test rejects combining the RP and SP data, the predictive ability of CM model should be poorer than that from RP and SP models. However, generally, the predictive ability of the CM model actually outperforms RP and SP models in both RP and SP holdout samples. In case of RP holdout sample, the CM model provides superior predictive ability over RP and SP models in both aggregate and individual level tests. Even though in the case of SP holdout sample, the CM model is not the best model for aggregate trip prediction, it clearly dominates other models in case of individual level tests.

In addition, the predictive ability results reveal it is not always true that model based on actual behavior data would predicts well just actual behavior. It could also provide superior predictive ability for the future behavior. This is found by the empirical results from RP and SP models. Given the CM model as the best model for prediction, the RP model clearly outperforms SP model in almost prediction cases. For the RP holdout sample, the RP model absolutely dominates the SP model for all tests. Interestingly, in the case of SP holdout sample, even though the aggregate predictive level of RP model is poorer than that from the SP model, it clearly provides higher predictive performance than SP model in case of individual level test.

However, this does not mean that we can only rely on the RP model, which relies only on actual behavior data, to predict recreationists' behavior due to changes in site management, because the SP data could also provide useful information beyond the current market situation, which could improve the reliability of the model predictive performance. This statement is confirmed by the empirical results that the CM models in both unweighted and weighted cases generate the best predictive performance over the RP and SP models individually.

In conclusion, this study investigation about the predictive ability of combined model's ability to predict both actual behavior and future behavior sheds light on the fact that the combined model (CM) provides the most accurate predictive performance over the individual models. This is due to the fact that the CM model contains both actual behavior (RP data) and future behavior (SP data) information, which offsets the weaknesses of prediction by the individual data sets. The changes in lake recreationist's behavior due to the future changes of lake management that are not currently available in the market can be captured by the SP data. In addition, the model's parameters are not distorted by hypothetical bias because they still rely on real behavior data (RP data). As a result, the CM model would be the best model for prediction both actual and future behavior.

Table 2.1 Descriptive Statistics of Attribute Level and Variables Used

Variable	Definition	Mean
Travel cost	U.S dollar (round trip)	182.47
Number of boat ramp		3.27
Availability of campsite		
No campsite	1 if no campsite, 0 otherwise	31.57%
Campsite	1 if site has campsite, 0 otherwise	66.22%
Campsite with electricity	1 if site has campsite with electricity, 0 otherwise	57.54%
Availability of restroom		
No restroom	1 for no restroom, 0 otherwise	17.35%
Portable toilet	1 if site has portable toilet, 0 otherwise	55.98%
Restroom with flush toilet	1 if site has restroom with flush toilet, 0 otherwise	39.41%
Restroom with flush toilet and shower	1 if site has restroom with flush toilet with shower, 0 otherwise	49.25%
Lodge	1 if site has lodge, 0 otherwise	7.41%
Water clarity	Secchi disk depth measured in foot	2.81
Major lake	1 if major lake, 0 otherwise	15.33%
Lake location		
Northeast region	1 if located in Northeast region, 0 otherwise	39.00%
Southeast region	1 if located in Southeast region, 0 otherwise	30.33%
Southwest region	1 if locate in Southwest region, 0 otherwise	13.00%
Northwest region	1 if located in Northwest region, 0 otherwise	18.78%

Note: Region is geographically indicated by bounds of I-40 and I-35, which divide Oklahoma into four regions.

Table 2.2. Parameter Estimates for Unweighted and Weighted Models of Oklahoma Lake Site Choice Models

Variable	Unweighted Model			Weighted Model		
	CM	RP	SP	CM	RP	SP
Travel Cost	-0.011***	-0.011***	-0.017***	-0.013***	-0.013***	-0.017***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Boat Ramp	0.009***	0.009***	0.209**	0.011**	0.010***	0.209**
	(0.006)	(0.009)	(0.012)	(0.048)	(0.009)	(0.012)
Campsite	-0.024	-0.176	0.740***	0.047	-0.085	0.740***
	(0.757)	(0.503)	(0.002)	(0.593)	(0.503)	(0.002)
Campsite with Electricity	0.063	-0.694	0.904***	0.122	-0.660	0.904***
	(0.401)	(0.108)	(0.000)	(0.153)	(0.108)	(0.000)
Porta-Potties	-0.277***	-0.076	-0.241	-0.367***	-0.087	-0.241
	(0.000)	(0.446)	(0.440)	(0.000)	(0.446)	(0.440)
Flush-Toilet	0.008	0.005	0.933***	0.022	-0.077	0.933***
	(0.909)	(0.960)	(0.000)	(0.796)	(0.960)	(0.000)
Flush-Toilet with Shower	0.321***	1.347***	1.243***	0.267***	1.315***	1.243***
	(0.001)	(0.000)	(0.000)	(0.007)	(0.000)	(0.000)
Lodge	0.016	0.444***	0.255	-0.001	0.507***	0.255
	(0.802)	(0.001)	(0.180)	(0.991)	(0.001)	(0.180)
Water Clarity	0.060***	0.119***	0.171**	0.036	0.116***	0.171**
	(0.002)	(0.000)	(0.045)	(0.158)	(0.000)	(0.045)
Major Lake [a]	1.637***	1.431***		1.836***	1.623***	
	(0.000)	(0.000)		(0.000)	(0.000)	
North East Region	0.505***	0.597***		0.384	0.489***	
	(0.009)	(0.002)		(0.207)	(0.002)	
South East Region	0.352*	0.468**		0.184	0.368**	
	(0.085)	(0.024)		(0.568)	(0.024)	
South West Region	0.063	0.249		-0.344	-0.105	
	(0.781)	(0.283)		(0.358)	(0.283)	
SP Intercept [b]			2.164***			2.164***
			(0.000)			(0.000)
Relative Scale Parameter	0.454***			0.474***		
	(0.000)			(0.000)		
No. of choices	27300	26208	1092	27300	26208	1092
Log-likelihood	-2412.958	-2040.308	-327.953	-1086.879	-725.711	-327.953

Notes: ***, **, and * indicate significant level at 1%, 5%, and 10%, respectively. Numbers in parentheses are p-value.

[a] Lakes with surface area bigger than 5,000 acres are coded to 1 and 0 otherwise.

[b] This intercept represent a dummy variable that equal 1 for option neither A or B and 0 otherwise.

Table 2.3. An Example of Trip Predictions for the Unweighted Models for RP and SP Holdout Samples

RP Holdout Sample

Model	Fort Gibson	Hefner	Tenkiller	Hudson	Boomer	Arcadia	Wes Watkins	Canton	Atoka	Broken Bow
Total Predicted Trips										
CM	911.8	94.1	180.4	92.9	25.0	31.5	25.2	34.7	3.2	12.3
RP	1329.2	100.3	176.9	82.9	21.6	36.3	23.0	29.2	2.6	26.4
SP	441.9	2.5	23.5	0.9	1.8	31.6	13.1	20.0	0.1	0.9
Total Actual Trips										
	288	149	57	49	41	35	34	27	22	12

SP Holdout Sample

Model	Fort Gibson	Eufaula	Copan	Wes Watkins	Canton	Hefner	Thunderbird	Sooner	Oologah	Okmulgee
Total Predicted Trips										
CM	2932.3	615.4	59.5	63.6	32.8	81.6	38.0	15.7	21.2	10.5
RP	2958.8	589.0	43.4	53.8	31.9	60.2	45.3	13.0	16.7	9.3
SP	2540.6	713.4	28.8	49.7	41.3	53.4	50.4	5.7	15.0	7.1
Total Hypothetical Trips										
	460	136	80	61	40	34	24	20	14	8

Table 2.4. An Example of Trip Predictions for the Weighted Models for RP and SP Holdout Samples

RP Holdout Sample										
Model	Fort Gibson	Hefner	Tenkiller	Hudson	Boomer	Arcadia	Wes Watkins	Canton	Atoka	Broken Bow
Total Predicted Trip										
CM	1041.1	80.1	197.3	89.8	21.4	30.2	24.0	49.8	2.4	10.3
RP	1827.6	91.7	207.1	85.4	13.3	32.0	23.9	38.0	2.5	31.0
SP	441.9	2.5	23.5	0.9	1.8	31.6	13.1	20.0	0.1	0.9
Total Actual Trip										
	288	149	57	49	41	35	34	27	22	12

SP Holdout Sample										
Model	Fort Gibson	Eufaula	Copan	Wes Watkins	Canton	Hefner	Thunderbird	Sooner	Oologah	Okmulgee
Total Predicted Trip										
CM	2928.9	636.6	62.8	62.8	33.8	85.3	38.4	16.3	21.8	10.7
RP	2840.3	603.0	45.8	51.4	32.1	62.7	45.4	12.3	17.0	10.1
SP	2540.6	713.4	28.8	49.7	41.3	53.4	50.4	5.7	15.0	7.1
Total Hypothetical Trip										
	460	136	80	61	40	34	24	20	14	8

Table 2.5. Results of the Predictive Ability Tests over Thirty Sets of Unweighted Models and RP and SP Holdout Samples

	RP Holdout Sample			SP Holdout Sample		
Model	RMSE	r_i	r_m	RMSE	r_i	r_m
			Mean Value of Statistics (Minimum-Maximum)			
CM	69.122	0.094	0.220	256.538	0.392	0.976
	(45.675-93.732)	(0.032-0.168)	(0.171-0.269)	(77.818-438.441)	(0.169-0.816)	(0.971-0.985)
RP	97.939	0.086	0.202	263.683	0.374	0.928
	(67.041-151.430)	(0.024-0.162)	(0.147-0.257)	(71.399-461.949)	(0.145-0.726)	(0.882-0.960)
SP	593.277	0.021	0.048	221.734	0.283	0.890
	(44.143-732.080)	(0.001-0.194)	(0.016-0.260)	(44.973-398.646)	(0.119-0.587)	(0.847-0.920)
			Mean Rank			
CM	1.033	1.033	1.033	2.133	1.267	1.000
RP	2.033	2.033	2.033	2.567	1.800	2.100
SP	2.933	2.933	2.933	1.300	2.933	2.900

Table 2.6. Results of the Predictive Ability Tests over Thirty Sets of Weighted Models and RP and SP Holdout Samples

	RP Holdout Sample			SP Holdout Sample		
Model	RMSE	r_i	r_m	RMSE	r_i	r_m
			Mean Value of Statistics (Minimum-Maximum)			
CM	79.390	0.094	0.218	257.407	0.391	0.973
	(51.504-99.771)	(0.032-0.164)	(0.168-0.258)	(80.680-441.387)	(0.171-0.808)	(0.967-0.986)
RP	121.966	0.084	0.198	261.970	0.370	0.921
	(83.311-180.018)	(0.021-0.151)	(0.143-0.245)	(71.399-449.895)	(0.147-0.740)	(0.836-0.965)
SP	593.277	0.021	0.048	221.734	0.283	0.890
	(44.143-732.080)	(0.001-0.194)	(0.016-0.260)	(44.973-398.646)	(0.119-0.587)	(0.847-0.920)
			Mean Rank			
CM	1.033	1.067	1.067	2.167	1.333	1.000
RP	2.033	2.000	2.000	2.433	1.733	2.200
SP	2.933	2.933	2.933	1.400	2.933	2.800

CHAPTER III

VALUING LAKE RECREATIONAL DEMAND: THE CASE OF TWO-STEP APPROACH WITH TAKING INTO ACCOUNT POTENTIAL LAKE USERS

Introduction

When people decide to take recreational trips, they consider where they want to go and how many times to go. In recent years, many studies focusing on recreational demand have usually applied conditional logit models to study site choice selection. The conditional logit model explains well which sites will be visited by recreationists, but it does not explain how many trips will be taken. The latter issue becomes important because when the quality of amenities in recreational sites changes; both the site selected and the trip numbers will be affected. Count models, such as poisson and negative binomial models, can be used to explain changes in the trip numbers due to changes in destination quality, but they cannot verify the changes in site substitution across recreational sites. Due to the weakness of these models in terms of inability to explain changes in trip numbers and site substitution, some researchers have developed linked site selection models to explain the site selection and number of trips taken by recreationists (Feather et al., 1995; Hausman et al., 1995; Parsons and Kealy, 1995; Parsons et al., 1999). The linked site selection model is based on a two stage estimation. The first stage is site allocation model, and the second stage is the trip number model. The previously mentioned studies of linked site selection models rely on actual behavior, revealed preferences, using information from individuals who actually participated in recreational activities. However, no previous studies have accounted for people who may participate in the future if the quality of recreational sites improves in a way that induces them to take a visit a site even if they had taken no trips before. Ignoring the recreation of potential participants may distort the total benefit gained, in terms of per choice occasion welfare and the number of trips taken, as a result of the quality improvement because the benefits of some individuals are uncounted.

Combining the revealed preference (RP) and stated preference (SP) data reduces the bias from missing potential recreationists who opt in because the SP data can ask hypothetical questions with quality changes to both survey recreational participants or non-recreational participants. This allows us to measure the recreation benefits in terms of increases in numbers of trips taken and per choice occasion with also taking into account the potential new recreationist enticed by improving site quality. Another benefit gained by combining the RP and SP data is the reduction in bias in the estimators, which comes from when respondents are not familiar with the potential site qualities beyond the current situation from SP questions. This anchors the hypothetical behavior (SP) with actual behavior (RP).

However, the standard error obtained from second stage model does not take into account error that appears when generating the predictor from first stage model. This would result in downward biased standard error in the second stage model, which leads to an incorrect hypothetical test. To deal with this problem, the bootstrap technique is applied to calculate robust standard errors for the second stage model. Applying this method using the combined RP and SP data to estimate demand for recreation results in two benefits. First of all, the benefit gained from quality improvements in terms of changing the site selection and numbers of trips taken can be calculated while taking into account potential participants. The second benefit is that the two-step estimation with standard error correction approach provides corrected standard errors, so reliable statistical tests can be imposed.

The data used in this study comes from a statewide survey of Oklahoma lake recreation conducted in Fall 2007. The survey contains both RP and SP data questions. The discrete choice analysis, SP data, provides information on potential behavior of increased or decreased visitation due to quality improvements and changes in price. Because the survey was designed to give visitation changes with potential quality improvements in the stated preference survey, the two step estimation is possible. By combining the RP and SP data of both current and potential lake users, this study propose an estimation method to estimate recreation benefits from site quality improvement that takes into account site allocation, numbers of trips taken, and potential lake recreationists.

Theory Discussion

Consider an individual's demand, shown by Figure 3.1, to take trips to recreational site i, $i = 1, 2, \ldots, j$, which depends on the price and quality of site i. Improvement of a site's quality leads to a forward shift of demand, resulting in an increase in numbers of trip taken, given the price constant. From figure 1, when the quality of recreational site is improved from q^0 to q^1, the individual may decide to take T^1 trips. $D(q^1)$ represents the demand shift resulting from the site's quality improvement. In this case, the consumer surplus gained from site's quality improvement is shown by area a.

In addition, when the site quality is improved, new participants who have never participated in recreation activity at the original quality level may visit the recreational site. These individuals are called as potential recreationists.[11] This situation can be represented by figure 3.2. Before the site's quality improvement, at the price p_0 the potential recreationist's demand for trip is zero. However, after an improvement in the site's quality to q^1, he or she might decide to take T'^1 trips, and consumer surplus gained by this individual is area b.[12]

From these two figures the total consumer surplus gained from current and potential lake recreationists is represented by area *a* plus *b*. These two figures also provide the intuition that if the potential recreationists are ignored, the consumer surplus

[11] There is another group of individuals who would not participate in recreation at any price or quality. These individuals do not participate in recreation for reasons such as health and preference, so they do not receive any consumer surplus (Grogger and Carson, 1991; Haab and McConnell, 1996; Whitehead et al. 2000).

[12] We expect that the numbers of trips taken by potential recreationists after a site's quality improvement should be less than that of the current users.

gained from quality improvement should be downward biased because their benefits gained are ignored.

Data Description

Data used in this study were collected by a mail survey entitled, "Oklahoma Lake Use" (2007) which included information to estimate the travel cost method and a discrete choice experiment. Data on travel distances and lake characteristics were compiled from GIS maps from Oklahoma Water Resource Board (OWRB), which was created by Caneday and Jordan (2003), lake websites, and phone interviews with lake managers.

The survey was mailed to 2,000 individuals, who were randomly chosen, in every county of Oklahoma during fall 2007. A random sample was obtained from Survey Sampling Inc, Fairfield CT stratified across 6 regions of Oklahoma. The survey was first distributed during the last week of September 2007 by mail. Standard Dillman procedures were used to get the highest possible response rate (Dillman, 2000). Two weeks after the survey was mailed; the postcard reminder was mailed to people who had not responded. Then, two weeks later, the follow up survey with cover letter was mailed individuals who had not replied to the survey. As a result, 401 surveys out of 2000 were returned. One hundred and twenty one of them were incomplete and unusable surveys, and allowing for 150 undeliverable surveys due to no forwarding addresses, the net response rate was 15.14 percent.[13] Descriptive statistics of attribute levels and variables used are given in

[13] An unusable survey is a survey that the respondents report as never visiting lake before and also selected choice C (want to do as they stated in RP question) in discrete choice questions indicating no preferences for change.

Table 2.1. Since revealed and stated preferences data are used, the survey was designed to obtain both types of data.

Revealed Preference Data

Respondents were asked to report their visitation patterns for single-day trips to 144 public lakes in Oklahoma in 2007.[14] They were also asked to report their activities at the lakes, as well as the features of lakes that were important to them. In order to obtain the effect of water quality on lake recreation demand, water clarity was used as the proxy for water quality. Water clarity data was gathered from the Beneficial Use Monitoring Program (BUMP) database of OWRB (Beneficial Use Monitoring Program Report, 2007).[1] Other amenity data were collected for each lake including the types and numbers of restrooms, docks, campsites and boat ramps, etc. These amenity data were collected from the lake websites and/ or by phone interview. TransCAD software was used to calculate the distance from each ZIP code to 144 lakes via roads by assuming that respondents selected to travel by shortest path (TransCAD, 2008). Then, the distances were expressed as round trip travel cost, which was combined with out-of-pocket expenditure and opportunity cost of time.[2]

Stated Preference Data

The survey solicited SP data. Each respondent faced two discrete choice sets which presented possible alternative lake recreational opportunities at differing lake amenity levels and distance. These choice sets were orthogonally designed with quality and

[14] The survey also provides choice for people who have never visited lake before. Even though, these people have never visited lake before, they are also asked to answer the discrete choice questions.

amenity improvements at a lake similar to the lake respondents most often visited (which they indicated in the RP portion of the same survey). The SP questions elicited visitors' preferences for lake characteristics, including availability of lake amenities and distance. Six measurable attributes associated with lake recreation experiences of either 2 or 6 levels were determined (Table 1.1). This created $4^3 \times 3 \times 2 \times 6 = 2{,}304$ possible combinations. Each combination was then randomly paired with another combination (Lusk and Norwood, 2005). The third option was stated as the respondent's most frequently visited lake as given in the revealed preference data.

Each respondent was asked to answer two experimental choice questions. Each of them contains two options of hypothetical lakes (Figure 1.1). Because some attributes of the SP questions such as the number of boat ramps, water clarity, and distance, were asked as a quality improvement, i.e. and increase in amenities, the information from lakes that were most visited by each respondent was used as the base information to adjust the levels of those attributes to be the same as RP data. For example, if Tenkiller Lake was the lake most visited by a respondent, the number of boat ramps in SP question was added by the actual number of boat ramps in Tenkiller Lake. Moreover, the SP questions also asked the number of trips respondents would take given the lake they choose in the discrete choice question. This allows us to determine the number of trips they would take under a hypothetical situation.

Empirical Model

The Site Choice Selection

An individual' decision to visit lakes is modeled using the random utility model. The idea of random utility model is that an individual would visit lake i if his/ her utility gained from visiting lake i is higher than or equal that from either other lakes or not visit any lakes;

(3.1) $$U_i(\mathbf{q}_i\ , p_i\ ) \geq U_j(\mathbf{q}_j\ , p_j\ );\ \forall j \in C$$

where U_i is the utility of visiting site i and U_j is the utility of visiting any other site in choice set C, which also includes the option to not visit any lakes. Also, $\mathbf{q}$ and p are the vector of lake quality and cost of visiting lake, respectively. Let us assume that the indirect utility function consists of two components, which are an observed component (V_i), and an unobserved component (ε_i). The probability of visiting site i is

(3.2) $$Pr(i) = \Pr\left(V_i + \varepsilon_i \geq V_j + \varepsilon_j;\ \forall j \in C\right)$$

The observed component would be observed by either respondent's actual behavior or from the hypothetical responses in which the attributes are arguments. If the distribution of the stochastic component is independently and identically distributed (IID) according to Gumbel random variable, the probability of choosing choice i among those available in choice set C can be expressed in closed form as

(3.3) $$Pr(i) = \frac{\exp(\mu V_i)}{\sum_{j \in C} \exp\left(\mu V_j\right)}$$

where μ is scale parameter. Commonly, in the case of a single set of data, the scale parameter cannot be identified, so it is usually set equal to 1. However, for at least two data sets pooled together, the scale parameter can be identified.

In this study, two data sets, the current lake user data and the potential lake user data, are pooled together. In addition, each data set also contains two types of data, which are RP and SP data. Therefore, differences of variances between data sets and also between types of data are possible. Three scale parameters to capture these differences are constructed. The first scale parameter calibrates the difference of variances between the current and potential lake user data by normalizing the scale parameter of current lake user data to 1. The second scale parameter captures the differences of variances between RP and SP data of current lake user data set by setting the scale parameter of RP data equal to 1. The differences between RP and SP data for potential lake user data sets are also possible, so the third scale parameter is constructed to clarify these differences by normalizing the scale parameter of RP data to unity. To show how scale parameters are allowed to vary, the variables for equation (3.3) is rewritten as

$$Pr_{ns^m}(i) = \frac{\exp(\mu_{s^m} V_i)}{\sum_{j \epsilon C} \exp(\mu_{s^m} V_j)} \tag{3.4}$$

Equation (3.4) represents the probability of choosing site i of individual n for trip scenarios s^m, where s is current lake user data or potential lake user data, and m represents RP or SP data. The log likelihood function to be maximized then becomes

$$L = \prod_{n \epsilon N} \prod_{s^m \epsilon S^m} Pr_{ns^m}(i) \tag{3.5}$$

However, in the current lake user RP data, each respondent can visit more than one site in each choice set provided in the questionnaire. This may create an overweighting problem for the RP observations since other RP and SP data are considered as one choice set and each respondent provides one response in each choice set. To solve this problem, equation (3.5) is weighted by weighting current lake user RP data by the trip proportions, and these proportions also add up to one over each RP choice set (Adamowicz et al., 1997; Haener et al., 2001). For example, if some respondents visited three different lakes, those three lakes will be weighted by one third and the rest of lakes are weighted by zero. By weighting the data in this manner, all observations are given equal weight.

The Number of Trips Taken Model

In addition to the decision to choose lakes to visit, respondents also decide how many trips they would take given the current and hypothetical lake attribute's quality. To create the trip number model and linked site choice selection model, the models in this study is followed the Hausman et al. (1995). The linkage between the site choice selection model and trip number model is calculated by dividing expected utility from visiting each lake (calculated from site choice selection model) by the absolute value of the travel cost coefficient from the site choice selection model. This variable reflects the per trip consumer surplus from visiting each lake.[15] Later, this variable is called a per trip consumer surplus.[16] [17]

[15] Hausman et al. (1995) claimed that by using the per trip consumer surplus in the second-stage model, their linked model is theoretical consistency with two-stage budgeting process. However, Smith (1997) and Herriges et al. (1999) argued that this consistency would only hold in cases where extremely assumptions are maintained. Herriges et al. also suggest that a Kuhn-Tucker model may be better in case of utility

The trips to a lake are given as count data, so a count model such as poisson model or negative binomial model is appropriate. Because each data set for each respondent contains more than one choice of lake, the random effects model is employed to take into account the heterogeneity among individuals. Assume that $T_{is^m n}$, the number of trips taken to lake i by individual n in a particular trip scenario s^m, is draw from the Poisson distribution with mean $\lambda_{is^m n}$

$$\Pr(T_{is^m n} = t_{is^m n}) = \frac{\exp\left(-\lambda_{is^m n}\right)\lambda_{is^m n}^{t_{is^m n}}}{t_{is^m n}!} \tag{3.6}$$

where

$t_{is^m n} = 0, 1, 2, \ldots; s =$ current and potential lake user data; $m =$ RP and SP data.

$\lambda_{is^m n}$ depends on the per trip consumer surplus, demographics of respondents, and lake activities and is as follows

$$ln\lambda_{is^m n} = ln\mu_{is^m n} + u_n = \theta CS_{is^m n} + \varphi A_n + \vartheta w_{in} + u_n \tag{3.7}$$

$CS_{is^m n}$ is the per trip consumer surplus of individual n taking trip to lake i in trip scenario s^m; A_n is a vector of respondents demographics; w_{in} is a vector of activities at lake i of individual n; u_n is the random effect for individual n. The variable u_n allows trip variation across individuals that cannot be explained by independent variables. The estimate parameters are $\theta, \varphi,$ and ϑ. The distribution of trips $t_{is^m n}$ can be either poisson distribution or negative binomial distribution depending on its mean and variance. If mean and variance are equal, the poisson distribution is appropriate. However, if

consistency but the estimation is difficult, especially in the case of a large number of available recreational sites.

[16] In the Hausman et al. (1995) paper, they reversed sign of this variable to the negative and called it as price index. However, to prevent confusion, we prefer to use it as the per trip consumer surplus.

[17] This consumer surplus is based on the assumption of indifference in Willing bounds, which means that the areas under Marshallian demand curves are as close of the more exact areas under Hicksian demand curves (Haab and McConell, 2002).

variance exceeds mean, overdispersion, the negative binomial distribution is preferred. This will be tested in the estimation process.

To combine current and potential lake user data sets, for which each data set contains RP and SP data, structural changes in trip demand between these data may exist. Hence, we create dummy variables to account these structural changes. The first dummy variable, which captures the structural change between current and potential lake users, is set equal to 1 if the observation comes from potential lake user, and equal to 0 otherwise. This dummy variable captures the parallel demand shift between current and potential lake users. In addition to the difference between current and potential lake user data, each data set also contains RP and SP data, so the structural changes among these data may be possible. Therefore, another set of dummy variables are created. The first dummy variable of this set is the SP dummy variable, which is set equal to 1 if the data is SP and 0 otherwise. In addition, the SP data of potential lake user may be different from that of current lake user. To account for the structural change between these two data, the second dummy variable of this set is included to the model by interacting the SP dummy variable with potential lake user dummy variable.

In addition, differences in data sets may also affect the slope of demand for trips. Hence, each dummy variable is multiplied by the per trip consumer surplus to capture this effect. The demand for a lake recreation model that allows us to pool the RP and SP data of current and potential lake recreationists together is derived by adding and interacting these dummy variables into the mean $\lambda_{is^m n}$. The modified trip demand model can be shown as follows

$$(3.8)\quad ln\lambda_{is^m n} = ln\mu_{is^m n} + u_n$$

$$= \theta CS_{is^m n} + \varphi A_n + \vartheta w_{in} + \gamma_1 D_1 + \gamma_2 D_2 + \gamma_3 (D_1 * D_2) +$$

$$+\gamma_4 (D_1 * CS_{is^m n}) + \gamma_5 (D_2 * CS_{is^m n}) + \gamma_6 (D_1 * D_2 * CS_{is^m n}) + u_n$$

where D_1 and D_2 are the potential lake user dummy variable and the SP dummy variable, respectively.

Including these dummy variables with their interaction effects requires us to test several hypotheses. The first hypothesis to be tested is that if there is no structural change at all between current lake and potential lake user's data sets, then $\gamma_1 = \gamma_4 = 0$. The second hypothesis is that it is possible that the structural changes in trip demand are parallel shifts so that $\gamma_1 \neq 0$ and $\gamma_4 = 0$. Besides the structural change between current and potential lake users, it is also possible that the structural changes between RP and SP data may occur. Hence, the hypothesis $\gamma_2 = \gamma_5 = 0$ is also tested for both parallel and slope changes between RP and SP data. On the same manner, if only the parallel shift occurs then $\gamma_2 \neq 0$ and $\gamma_5 = 0$. The final set of hypotheses to be tested is the SP data of potential lake user. These tests test for the structural changes between entire SP data and the SP data of potential lake user. If there is no structural change between these two data sets, $\gamma_3 = \gamma_6 = 0$. If structural change is just parallel shift, then $\gamma_3 \neq 0$ and $\gamma_6 = 0$. However, the standard errors estimated from the second step model, trip taken model, are incorrect because they do not take into account the errors from first step model, site choice selection model. Ignoring this problem would result in wrong hypothetical test results.

Two approaches can deal with this problem of incorrect standard errors in the trip numbers model. The first is the Full Information Maximum Likelihood (FIML) approach.

The second is a two-step approach. The FIML would yield consistent estimators and asymptotically correct estimates of standard errors for second-step model, the number of trips taken model, in this study only if the joint distribution of errors between first and second steps models is defined correctly. However, previous research shows that sometime it is difficult to identify the appropriate joint distribution of errors (Hubbell et al., 2000; Greene, 2003; Starbuck et al., 2004). Two-step approach, on the other hand, does not require joint-density function for the errors, and it would also yield the consistent estimates of second-step model parameters. However, the standard errors of the second-step model are miscalculated because the errors occurred in the first-step model are not taken into account in the second-step model. This problem causes incorrect statistical tests. Murphy and Topel (1985) developed a standard error correction approach that can provide corrected standard error for second-step model. However, it is difficult to implement, especially for panel data models (Martina and Neha, 2007). As shown by Petrin and Train (2002), Pinar and Train (2003), and Martina and Neha, (2007) a bootstrap technique can substitute for the Murphy-Topel method to correct the standard errors in the second step model, so this technique is applied to correct the standard error in the second step model.

Welfare Estimation

From the two-stage approach presented above, the welfare changes from lake quality improvement that take into account per choice welfare change and change in trip numbers can be calculated. When the quality of a lake is improved, the per trip consumer surplus, which is measured from the site choice selection model, for each lake also changes. Since the consumer surplus reveals the per choice (also per trip) welfare changes due to the quality improvement, it also impacts the number of trips taken, changes in consumer surplus calculated from the difference of trip taken before and after quality improvement could reflect the welfare changes that account for per trip welfare and number of trip taken changes (Hausman et al. 1995; Parsons et al. 1999). The measurement of this welfare change can be represented as follows[18]

$$(3.9) \qquad \Delta W_H = \int_{CS_0}^{CS_1} \exp\,(X\beta + \theta CS + u_n) dCS$$

$$= \frac{1}{\theta}(T^1 - T^0)$$

where X is the vector of variables in the number of trip model, shown in equation (3.8); CS_1 and CS_0 are the consumer surplus after and before lake quality improvement, respectively. T^1 and T^0 are the predicted trip numbers after and before lake quality improvement, respectively.[19]

In addition to make this welfare calculation results strong, welfare analysis proposed by Bockstael et al. (1987) is also applied, which can be employed with the two-stage model presented above (Parsons et al., 1999).[20] This welfare analysis starts with calculating per trip welfare changes measured from site choice selection model, which can be shown as follows[21]

[18] This welfare measurement is based on the assumption of small income effect, which makes the Hicksian demand function for trips is approximately the same as the Marshallian demand function represented by equation (3.8) (Hausman et al. 1995).
[19] In case of potential lake user, the predicted trip number before lake quality improvement, T^0, is set to zero.
[20] They found that the welfares estimated from Hausman et al. (1995) and Bockstael et al. (1987) methods were similar.
[21] This formulation assumes the marginal utility of income is constant, so the Marshallian demand is approximately equal to the Hicksian demand.

$$\Delta W_{PT} = -\frac{1}{\alpha}(I^1 - I^0) \tag{3.10}$$

where α is the coefficient of travel cost from site choice selection model and I^1 and I^0 are the expected maximum utilities after and before lake quality changes, which are also calculated from site choice selection model.

Equation (3.10) is then multiplied by the average of predicted trip numbers before and after lake quality improvement, which are estimated from the count model, equation (3.8). The welfare change that include changes of per choice occasion welfare and number of trips is

$$\Delta W_P = \frac{(T^1 + T^0)}{2} * \Delta W_{PT} \tag{3.11}$$

To compare the welfare measurement between the combined RP and SP model that includes both current and potential lake users and the combined RP and SP model using only current lake users, the set of combined current lake user RP and SP models are also employed. For this case, the first stage model contains only the relative scale parameter between RP and SP current lake user data. In addition, only the current lake user RP/SP dummy variable and its interaction with consumer surplus shown in (3.8) are included in the second stage model.

Estimation Results

The first stage estimation results are shown in Table 3.2. The FCP model is the first stage model of the combining current and potential lake user case, while the FCO is the first stage model of the current lake user only case. Starting with the FCP model, most of coefficients in this model are consistent with theory and previous lake recreation studies. Lakes located closer to an individual's home have a higher chance of being visited than those further away. In addition, lakes with higher quantity amenities such as numbers of boat ramps, availability of flush toilets and flush toilets with showers seems to attract lake recreationists more than those with fewer of these amenities. Lakes with higher water clarity are also preferred by lake recreationists. For the unique variables of the RP data, the variable, type of lake, reveals that major lakes, which have water surface area more than 5,000 acres, are also preferred by lake users. However, the comparison between regional variables and the no visits to lake(s) option, which is selected by all potential lake recreationists in RP question, reveals the "no visit" option provides higher utility than visiting a lake option represented by the lake region locations. This surprise result is shown by the negative sign and significant level of North East (NE), South East (SE), South West (SW), and North West (NW) variables. It may be due to the fact that the no visit choice is uniformly chosen by all potential lake users, while the region variables, specified by lake locations, are not uniformly selected by current lake recreationists. This may result in insignificant impact of region factor on choosing lake to visit, shown by similar coefficient values of each region variable. This will be investigated in the first stage model of the current lake user only case (FCO model).

To combine the RP and SP data from current lake user and potential lake user, the relative scale parameters that take into account the differences in variances of these data are estimated. *Sigma1* represents the relative scale parameter of current lake user data and potential lake user data, while *Sigma2* and *Sigma3* represent the relative scale parameters of RP and SP data of current lake user and potential lake user, respectively. The estimate results confirm the differences in variances of these data sets due to the statistical significance of the coefficients on *Sigma1*, *Sigma2*, and *Sigma3*.[22]

Turning now to the FCO model, generally the pattern of preference across attributes of FCO model is similar as that estimated by FCP model. Namely, most of coefficients have the same signs and are also significant like the FCP model. However, the size of coefficients in FCO model is generally bigger than those obtained from FCP model. This could imply that the current lake users may be more sensitive to changes in lake attributes than potential lake users. Because the FCO model contains only RP and SP data of current lake users, only one relative scale parameter is estimated. The value of the relative scale parameter in this case is also statistically significant, which confirms the difference in variances of current lake user RP and SP data.[23] Because this model contains only data from current lake user, the *NE*, *SE*, and *SW* locations are compared to the *NW* location. The result is interested, which is no regional variables are statistical significant. This confirms the expectation in FCP model that region factor may not be the key factor for current lake recreationists to make their visiting decision.

After estimating the first stage models, then using each respondent origin with current lake condition (RP data) and hypothetical lake condition in discrete choice question (SP data), the per trip consumer surplus for each lake obtained by each respondent is computed. This per trip consumer surplus is used as a linkage variable between site choice selection model and the number of trips taken model. Other

[22] We also test for the equality of parameter vectors of these data sets after taking into account the relative scale parameters. The test reveals that the coefficient vectors among these data sets are significantly different ($\chi^2 = 130.587, df = 9$).
[23] We also conduct the same test as we did for the FCP model. Similarly, the test reveals that the coefficient vectors between RP and SP data of current lake user are significantly different ($\chi^2 = 25.477, df = 9$).

explanatory variables included in the random effects negative binomial model are the activity engaged in at the lake, income, and dummy variables that capture the structural changes among these data sets.[24]

Table 3.3 represents the estimation results of number of trip taken models for both combining current and potential lake users and only current lake user data sets. Starting at the combined current and potential lake user data set, two models are estimated, SCP1 and SCP2. The SCP1 is the most general model, in which the intercepts and slopes of demand for trip are allowed to vary across current and potential lake users as well as their RP and SP versions. The coefficient of per trip consumer surplus has the expected sign and is statistically significant.

Only fishing, swimming, and picnicking activities seem to have significant impact on the numbers of trip taken, while other activities reported in the table may have no effect on the numbers of trip taken.[25] Only individuals who have annual income higher than \$60,000 tend to significantly take more trip than those who have annual income less than \$60,000.

Then to test whether the current and potential lake user's data represent the same underlying behavior, the hypothesis that $\gamma_1 = \gamma_2 = 0$ is tested. The test reveals that these coefficients are jointly significantly different from zero ($\chi^2 = 46.25, df = 2$), suggesting that the trip demand of potential lake users is different from that of current lake user. However, individual tests show that the structural change may be just parallel

[24] We run descriptive statistics to check whether the mean and variance are equal. The descriptive statistics clearly present that variance exceeds mean, so the random effect negative binomial is preferred for our data.
[25] Activities in lakes reported in the table are compared to other activities.

shift only due to the statistically insignificance of γ_2. As expected, due to the negative sign of γ_1, at the same level of lake quality, potential lake users seem likely to take fewer trips than current lake users. A similar test between RP and SP data sets is also conducted. This test tests whether the SP data represent the same trip demand model as the RP data, $\gamma_3 = \gamma_4 = 0$. The result is also similar to the previous test. Namely, the joint test reveals the structural change between these two data sets ($\chi^2 = 90.84, df = 2$), and both γ_3 and γ_4 are individually statistically significant, indicating that both intercept and slope structural changes may exist between RP and SP data. The final test conducted for this model is whether the trip demand model of SP data of potential lake user is different from that of entire SP data. This test tests whether $\gamma_5 = \gamma_6 = 0$. Similar to the previous tests, this joint test shows that the structural changes of SP data of potential lake user may also occur ($\chi^2 = 6.50, df = 2$). However, this structural change may be just a parallel shift only thanks to insignificant of γ_6.

Based on these test results, a model without interaction between *Pot_ dummy* and *Cons surplus* and between *Pot_dummy*SP dummy* and *Cons surplus* is estimated. However, the entire SP trip demand is allowed to shift and change shape with the interaction variable. This model is shown as SCP2 in Table 3.3. In addition, all coefficients of these variables, which capture the differences among trip demand behavior of current and potential lake users, are statistically significant. As expected, the coefficient of *Pot_ dummy* is still negative as in SCP1 model, indicating that potential lake user would take fewer trips than current lake user given the same lake quality. The *SP dummy* and *Pot_dummy*SP dummy* variables indicate similar behavior of current and potential lake users. These variables show that current and potential lake users tend to take more trips when they answer the SP question. This is not surprise results because choices in the SP question have at least one lake's amenity improved from its current condition.

From the significance of these variables, the current and potential lake user's data, which each contains RP and SP data, can be combined after the structural changes between them are calibrated. For other variables, as can be seen, most of their signs and statistical significance change only slightly when compared to the SCP1 model. Therefore, the estimated results of this model are not discussed again.

To compare the welfare changes from improving lake quality improvement, a demand for trip model with current lake users only is also estimated. The estimation result of this model is represented by SCO column in Table 3.3. Most of coefficients have similar pattern as those from SCP2 model. However, the size of coefficient is generally larger than those obtained from SCP2 model. The coefficient of *Cons surplus* variable, for example, is 0.011, which is significantly larger than that of SCP2 model, 0.005. This confirms that the difference in trip demands between current and potential lake users exists. In addition,a joint test result clearly indicates the difference underlying behavior between RP and SP data of current lake user ($\chi^2 = 1450.99, df = 2$). Individual level tests also confirm the parallel shift and slope change of SP behavior of current lake user.

From the SCP2 and SCO models results, the demand for trips is different among current and potential lake users and also for their RP and SP data. The changes in demand for trips are mixed. The demand for trips changes in both parallel shift and slope change for current lake user, while only parallel shift exists for the potential lake user.

Welfare Measures

After having the estimates results from both site choice selection model and numbers of trip taken model, these results is used to calculate the welfare changes due to an increase in 1 foot of water clarity. Two sets of welfare changes are estimated. The first set is the mean per trip welfare, which is calculated from the site choice model. The second set is the annual welfare, which is calculated from the trip demand model. Starting with the mean per trip welfare, Table 3.4 contains two sets of results, the combined current and potential lake users and current lake user only. In addition, each set has the welfare measures for two conditions; current lake condition and an increase in 1 foot of water clarity. The sample mean of lake condition is used to calculate welfare estimates in case of current lake condition. For an increase in 1 foot of water clarity case, water clarity attribute is increased to1 foot above the current mean value, while other attributes are the same as current lake conditions. The differences in welfare estimates between minor lake and major lake are also allowed. The results clearly show that the per trip welfare estimates from a model that combines current and potential lake users are significantly larger than those from current lake users only. As expected, major lake is valued higher than minor lake in both cases. In addition, improving water clarity by an increase in 1 foot of water visibility would increase per trip (per choice) welfare about \$10 to \$13 per trip (\$2007 USD).

Turning now to the annual welfare estimates, to calculate the annual welfare changes from an increase in 1 foot of water clarity for these two data sets, the means per trip consumer surplus in Table 3.4 are plugged in to the SCP2 and SCO model to calculate the predicted trip numbers of current water clarity and improving water clarity conditions.

These results are represented by Table 3.5. Starting with the SCP2 model, there are two sets of predicted trips number for improving water clarity, which are the predicted trip numbers of current and potential lake users. Each set also contains minor and major lakes predicted trip numbers. The predicted trip numbers of current lake user are predicted using per trip consumer surplus of an increase in 1 foot of water clarity and per trip consumer surplus of current water clarity. When the water clarity is improved by 1 foot of water visibility, the predicted trip numbers of minor and major lakes are 2.495 and 5.805, respectively. In case of current water clarity, the predicted trip numbers of minor and major are 2.372 and 5.518. This results in 0.123 and 0.287 increase in trip numbers for minor and major lakes of current lake user due to an increase in 1 foot of water clarity. These changes of trip numbers are then used to calculate the annual welfare changes due to an increase in water clarity. The HLM is the Hausman et al. (1995) welfare measure shown in equation 3.9, while BHK is the Bockstael et al. (1987) welfare measure represented in equation 3.11. For the minor lake, the HLM and BHK give very similar results of annual welfare changes, which is about $25. Similarly, these two welfare measures techniques also provide almost the same results for major lake, which yields an increase in annual welfare about $58. The similarity of welfare results calculated from these two techniques was also found by Parson et al. (1999).

In case of potential lake user, because they have not had experience visiting these lakes before and this information was not shown in the survey, so we assume that types of lake would not affect their visiting decision. From this reason, the predicted trip numbers of minor and major lakes for potential lake user are restricted the same by dropping the coefficient of major lake out from trip calculation. The predicted trip number of potential lake users is then calculated by using the same per trip consumer surplus as current lake user case, but calibrating the predicted trip number by the coefficient of *Pot_ dummy* variable, γ_1. This ends up with the trip number about 0.378 for potential lake user after the water quality improved by 1 foot increase of water visibility. Because the trip numbers at current water clarity is zero for potential lake user, this predicted trip number is used as the change in trip numbers. In contrast to the current lake user's case, the HLM and BHK techniques provide significant different welfare change results. The annual welfare change from HLM technique is about $77, while that from BHK technique is just $4.

Changes in trip numbers and annual welfare by using only current lake recreationist data are also calculated. The model used to estimate trip numbers in this case is SCO model. The pattern to estimate the predicted trip numbers is similar as that from the combined current and potential lake user's case. Namely, the per trip consumer surplus of an increase in 1 foot of water clarity and current water clarity are calculated from the FCO model. These two consumer surpluses then are included in the SCO model to calculate the predicted trip number after and before the water clarity change. The predicted trip numbers in case of an increase in 1 foot of water clarity are 0.127 for minor lake and 0.301 for major lake. The final trip calculation of SCO model is the predicted trip for current water clarity. The predicted trip numbers are 0.114 and 0.270 for minor and major lakes, respectively. Then these predicted trip numbers are used to calculate the seasonal welfare changes by HLM and BHK techniques. The results are surprising in which the welfare changes are very small in both techniques, which are about $1 and $3 for minor and major lakes. This may be due to the small amount of consumer surplus generated from the first stage model, which results in the very small predicted trip numbers as well as their changes after the lake clarity improvement.

As expected, per trip and annual welfares calculated from the combined current and potential lake recreationist’s model are generally larger than those from current lake recreationist model. Per trip welfares calculated from the combined model are generally almost three times larger than those from the current lake user model. Similarly, the combined model also significantly generates larger annual welfare changes than those from the current lake user model, even in case of potential lake users. In addition, the most important benefit of the combined model may be the fact that it can capture the welfare gained from the potential lake user after the lake water clarity improvement. This would prevent the downward bias of the total welfare estimation due to ignoring the benefit gained of potential lake users who could become participants when the lake quality is improved. However, the potential lake user’s welfare estimates could be biased in either downward or upward directions if the potential lake users do not react to quality improvement as they state in survey questions.

Conclusions

This paper shows how the current and potential lake user’s data, which each contains revealed and stated preferences data, can be combined to use in the linked site choice model, which could measure the annual welfare changes due to changes in recreational site amenities. The stated preference data allows estimating welfare changes beyond the range of historical and current quality variation. This study also states that new lake users attracted by improving in lake quality should be included in estimating recreation demands because some non-participants could become participants when the higher lake quality is introduced. Failure to include participants who become participants after the site quality improves results in biases in welfare estimation.

The empirical models suggest that structural change between current and potential lake user for trip demand exists. In addition, the structural change of trip demand also occurs among revealed and stated preference data of current and potential lake users. A significant shift and change in the shape of trip demand occurs for the current lake user, while only a significant shift in demand for trips exists for potential lake user. In addition to compare the annual welfare changes due to the lake quality improvement, the linked site choice selection model with current lake user data only is also estimated. The empirical results are similar as those from the combined current and potential lake user model in which the demand for trip shifts and changes in shape between RP and SP data.

These models then are used to calculate the welfare changes due to an increase of 1 foot of water clarity. The annual welfare changes calculated from the combined model are significantly larger than those obtained from the current lake user model even in case of potential lake user. In addition, the combined model can also capture the annual welfare change from the potential lake user, which cannot be generated by the current lake user model. This would be the most benefit generated by the combined model and this also shows that ignoring potential participants results in a downward bias of annual welfare measures because the benefits gained by potential participants are ignored.

In term of policy implications, only an increase in 1 foot of water visibility may not be enough to attract Oklahoman potential lake users to take trips to lakes because the predicted trip number from this improvement is actually lower than 1. To attract them, an increase in water clarity more than 1 foot and/ or improvement of other lake amenities such as restroom with flush toilet and shower might entice potential lake users to participate in lake-based recreation activities.

Even though, the model could verify the behavior of potential lake user as well as their welfare changes due to lake quality improvement, to predict their trip numbers and calculate the welfare changes, this study assume that the potential lake users would behave as they answer in the survey questions when the lake quality is improved. The welfare estimates from this study could be biased in either downward or upward directions if the potential lake users do not react to quality improvement as they state in survey questions, i.e. there is hypothetical bias. Therefore, future research should investigate this issue (Norwood et al., 2007). One way to do is to collect data from these participants by stated preference questions with predictable quality changes. Then collecting the revealed behavior of these participants again after the quality changes happen. A comparison of stated behavior before the quality change with revealed behavior after the quality change would provide some evidence whether potential lake users really react to the site quality changes as they state in the survey.

Table 3.1 Descriptive Statistics of Attribute Level and Variables Used

Variable	Definition	Mean
Travel cost	U.S dollar (round trip)	177.94
Number of boat ramp		3.31
Availability of campsite		
No campsite	1 if no campsite, 0 otherwise	31.36%
Campsite	1 if site has campsite, 0 otherwise	65.92%
Campsite with electricity	1 if site has campsite with electricity, 0 otherwise	57.25%
Availability of restroom		
No restroom	1 for no restroom, 0 otherwise	17.23%
Portable toilets	1 if site has porta-potties toilet, 0 otherwise	55.94%
Restroom with flush toilet	1 if site has restroom with flush toilet, 0 otherwise	39.40%
Restroom with flush toilet and shower	1 if site has restroom with flush toilet with shower, 0 otherwise	49.06%
Lodge	1 if site has a lodge, 0 otherwise	7.41%
Water clarity	Secchi disk depth measured in foot	2.75
Major lake	1 if major lake, 0 otherwise	15.27%
Lake location		
Northeast region	1 if located in Northeast region, 0 otherwise	37.84%
Southeast region	1 if located in Southeast region, 0 otherwise	29.21%
Southwest region	1 if located in Southwest region, 0 otherwise	11.95%
Northwest region	1 if located in Northwest region, 0 otherwise	16.60%
Consumer surplus		-622.16
Activity in lake		
Fishing	1 if fishing, 0 otherwise	57.54%
Boating	1 if boating, 0 otherwise	46.40%
Sightseeing	1 if sightseeing, 0 otherwise	41.82%
Picnicking	1 if picnicking, 0 otherwise	43.18%
Swimming	1 if swimming, 0 otherwise	42.42%
Yearly Income		
< 20000	1 if yearly income less than 20000, 0 otherwise	8.23%
20000-39999	1 if yearly income between 20000-39999, 0 otherwise	29.30%
40000-59999	1 if yearly income between 40000-59999, 0 otherwise	21.08%
60000-99999	1 if yearly income between 60000-99999, 0 otherwise	25.73%
> 100000	1 if yearly income higher than 100000, 0 otherwise	15.65%
Potential lake user	1 if potential lake user, 0 otherwise	41.85%

Note: Region is geographically indicated by bounds of I-40 and I-35, which divide Oklahoma into four regions.

Table 3.2. First Stage Model Results of FCP and FCO Models

Variable	FCP	FCO
Travel cost	-0.005***	-0.012***
	(0.001)	(0.001)
Boat ramp	0.006	0.012**
	(0.004)	(0.005)
Campsite	0.179***	0.166
	(0.048)	(0.108)
Campsite with electric	0.181***	0.281***
	(0.047)	(0.106)
Porta-potties	0.022	-0.213*
	(0.053)	(0.114)
Restroom with flush toilet	0.261***	0.191*
	(0.052)	(0.102)
Restroom with flush toilet and shower	0.343***	0.495***
	(0.060)	(0.124)
Lodge	0.012*	0.195*
	(0.006)	(0.100)
Water clarity	0.054***	0.128***
	(0.016)	(0.031)
Major lake	1.141***	1.554***
	(0.124)	(0.175)
NE	-3.160***	0.438
	(0.228)	(0.311)
SE	-3.434***	0.226
	(0.253)	(0.331)
SW	-3.333***	-0.385
	(0.258)	(0.396)
NW	-3.771***	
	(0.298)	
SP ASC	0.351***	0.680***
	(0.066)	(0.150)
Sigma1	0.309***	
	(0.030)	
Sigma2	0.236***	0.544***
	(0.038)	(0.094)
Sigma3	0.918***	
	(0.233)	
Log likelihood	-1387.762	-1030.523
No. of Observation	42178	26400

Note: *** and * indicate significant level at 1% and 10%, respectively. *Sigma1* refers to the relative scale parameter of current and potential lake user data. *Sigma2* refers to the relative scale parameter of RP and SP data of current lake users. *Sigma3* refers to the relative scale parameter of RP and SP data of potential lake users.

Table 3.3. Second Stage Model Results of SCP1, SCP2, and SCO Models

Variable	SCP1	SCP2	SCO
Cons surplus	0.005***	0.005***	0.011***
	(0.000)	(0.000)	(0.001)
Activities in lake			
Fishing	0.179***	0.176***	0.242***
	(0.064)	(0.064)	(0.075)
Boating	-0.054	-0.053	-0.093
	(0.058)	(0.059)	(0.069)
Sightseeing	0.090	0.090	0.099
	(0.059)	(0.059)	(0.069)
Picnicking	0.119**	0.117*	0.010
	(0.060)	(0.060)	(0.072)
Swimming	0.113*	0.112*	0.049
	(0.065)	(0.065)	(0.073)
Income			
20000-39999	0.031	0.028	-0.012
	(0.103)	(0.104)	(0.137)
4000-59999	0.120	0.118	0.143
	(0.111)	(0.111)	(0.139)
60000-99999	0.313***	0.309***	0.396**
	(0.111)	(0.112)	(0.134)
> 100000	0.344***	0.345***	0.421***
	(0.126)	(0.126)	(0.148)
Pot_ dummy (γ_1)	-1.544***	-1.887***	
	(0.399)	(0.269)	
*Cons Surplus*Pot_dummy* (γ_2)	0.001		
	(0.001)		
SP dummy (γ_3)	0.425***	0.399***	3.150***
	(0.106)	(0.103)	(0.084)
*Cons Surplus*SP dummy* (γ_4)	-0.002***	-0.002***	-0.008***
	(0.000)	(0.000)	(0.001)
*Pot_dummy*SP dummy* (γ_5)	0.935**	1.456***	
	(0.431)	(0.271)	
*Cons Surplus*Pot_dummy*SP dummy* (γ_6)	0.003		
	(0.002)		
Constant	-2.690***	-2.655***	-5.473***
	(0.161)	(0.161)	(0.168)
Log likelihood	-5953.249	-5955.395	-4946.629
No. of Observation	42718	42718	26400
LR test	29.150***	29.18***	19.460***

Note: ***, **, and * indicate significant level at 1%, 5%, and 10%, respectively. Numbers in parentheses are standard errors, which are calculated by 1,000 bootstrap repetitions. LR test tests which models between random effect negative binomial model and pooled negative binomial model is appropriate. The null hypothesis is the pooled negative binomial is preferred.

Table 3.4. Mean Per-trip Welfare Estimate for FCP and FCO Models

Water clarity	Minor lake	SE	Major lake	SE
	FCP (Current and potential lake users)			
Current condition	$133.478 ($62.849-$204.107)	36.036	$304.956 ($202.414-$407.497)	52.318
1 foot increase	$143.766 ($70.439-$217.093)	37.412	$315.244 ($210.438-$420.049)	53.473
	FCO (Current lake user only)			
Current condition	$21.932 (-$25.820-$69.684)	24.364	$104.511 ($44.227-$164.795)	30.758
1 foot increase	$32.422 (-$17.972-$82.816)	25.712	$115.001 ($52.307-$177.695)	31.987

Note: The numbers in parentheses indicate the 95% confidence intervals of per-trip welfare, which are calculated using the Delta method.

Table 3.5. Mean Annual Welfare Estimates and Changes in Trips due to an Increase in 1 foot of Water Visibility for SCP2 and SCO Models

Water clarity	Minor lake	SE	Major lake	SE
SCP2 (Current and potential lake users data combined)				
Current user				
Change in mean trips	0.123		0.287	
Change in welfare				
HLM (equation 3.9)	$25.034	1.049	$ 58.240	2.440
	($22.979-$27.090)		($53.457-$63.022)	
BHK (equation 3.11)	$25.040	7.173	$58.252	16.688
	($10.980-$39.099)		($25.545-$90.961)	
Potential user				
Change in mean trips	0.378		0.879	
Change in welfare				
HLM (equation 3.9)	$76.822	3.219	$ 178.719	7.488
	($70.513-$83.131)		($164.043-$193.396)	
BHK (equation 3.11)	$3.891	1.115	$9.053	2.594
	($1.706-$6.077)		($3.970-$14.137)	
SCO (Current lake user data only)				
Change in mean trips	0.013		0.031	
Change in welfare				
HLM (equation 3.9)	$1.262	0.080	$2.998	0.190
	($1.105-$1.419)		($2.625-$3.371)	
BHK (equation 3.11)	$1.264	0.304	$3.001	0.722
	($0.667-$1.860)		($1.585-$4.417)	

Note: The numbers in parentheses indicate the 95% confidence intervals of per-trip welfare, which are calculated using the Delta method. HLM and BHK represent the Hausman et al. and Bockstael et al. annual welfare measures, respectively.

Figure 3.1. Trip Demand for Current Recreationists at Current and Improved Site's Quality

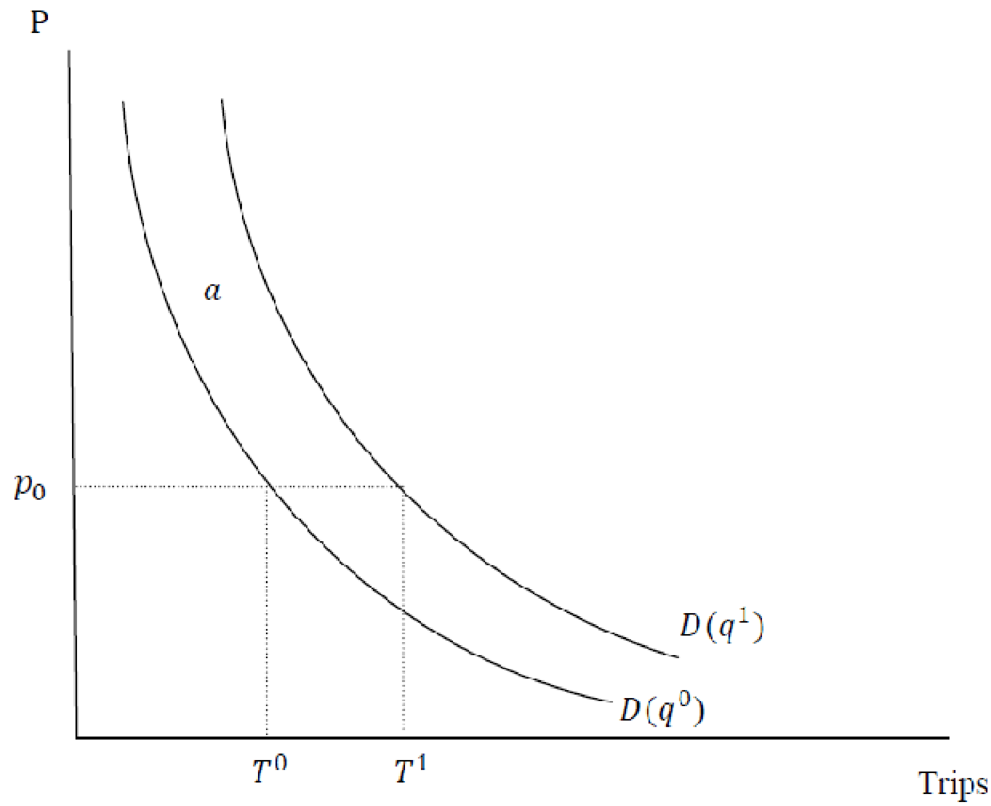
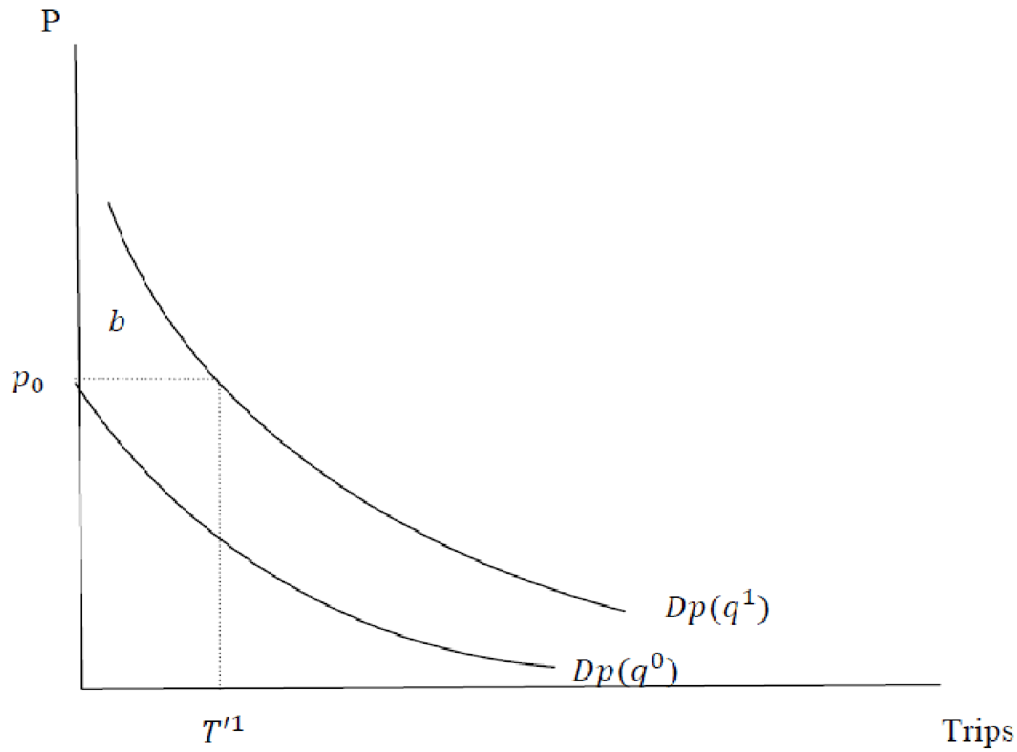


Figure 3.2. Trip Demand for Potential Recreationists at Current and Improved Site's Quality



CHAPTER IV

ESTIMATING DEMAND FOR URBAN FISHERIES MANAGEMENT: AN ILLUSTRATOPM OF CONJOINT ANALYSIS AS A TOOL FOR FISHERIES MANAGERS

Introduction

As the population becomes increasingly urbanized across the United States, angling participation and fishing license sales have declined (U.S. Fish and Wildlife Service, 2007). According to the U.S. census in 2000, 79% of the U.S. population and 72% of anglers live in metropolitan areas (U.S. Census Bureau, 2000; U.S. Department of Interior, 2002). Compared to rural residents, however, urbanites are less likely to participate in angling (U.S. Department of Interior, 2002). For example, in 2005, the U.S. census showed that 63.3% of Oklahoma's population lived in metropolitan areas, up from 61% in 1990 (Barta et al. 2007). Although Oklahoma's population has been increasing, the number of angling licenses has not increased proportionately (Summers, 2008). The nation-wide decreased involvement in fishing and hunting creates a disconnect between people and nature (ASA and AFWA, 2007) that can lead to reduced support for wildlife management or conservation issues (Kellert and Westervelt, 1983; Siemer and Knuth, 2001; Schramm and Dennis, 1993). Furthermore, state conservation agencies depend on fishing and hunting license sales to fund conservation and maintenance of wildlife areas (Noble and Jones, 1999; ASA and AFWA, 2007).

Urban fisheries may hold the key to reversing declines in angling participation, maintaining state budgets for wildlife management, and waning concern for natural resource conservation and the environment (See Eades et al., 2008 for an AFS-collected volume on Urban Fisheries). Urban dwellers have high opportunity costs for their time, meaning angling opportunities need to be placed close to these urbanites or they will continue to pursue activities that require a lower time commitment (Hunt and Ditton, 1996; Fedler, 2000; ASA andAFWA, 2007). However, quality fishing opportunities that are strategically placed can recruit lapsed anglers back into participation (Fedler, 2007).

The interests and factors associated with satisfaction of anglers in urban fisheries can differ from those of rural anglers (Arlinghaus and Mehner, 2004). Relatively little is known about what urban anglers value. Therefore, part of building a successful urban fishing program is assessing angler needs and interests so that fishing opportunities and amenities can be tailored to what the urban anglers value most (Balsman and Shoup, 2008). Different groups of anglers may have different interests (Hunt and Ditton, 1997), so amenities may need to vary by location to meet demand of a diverse urban population (Hunt and Ditton, 1997; Toth and Brown, 1997). For example, anglers who fish with family members place importance on physical amenities such as picnic tables, restrooms, and camping facilities whereas solo anglers place more importance on their ability to catch fish (Hunt and Ditton, 1997). The cost of stocking fish and maintaining park amenities can be expensive. With limited funds, state agencies need to consider the cost-benefit tradeoffs of different options for maintaining urban fishing programs.

Environmental and publicly provided goods, such as urban parks, are not commonly valued in the market place, making justification of specific management changes difficult to quantify on a cost-benefit basis. Therefore, tools are needed that allow managers to better assess the costs and benefits of multiple management options.

The objective of this study is to illustrate the use of a non-market valuation technique, called discrete choice analysis, for assessing the effect of different management or quality variables on demand for urban fisheries using three fishing ponds that are a part of the Close-to-Home-Fishing Program (CTHFP) in the Oklahoma City, OK metropolitan area. As we will illustrate, having relative values for potential management changes for a fishery helps to inform fishery managers about what characteristics the anglers value most, allowing them to better serve the anglers' interests and to justify the costs of implementing these changes.

Methods

Study site

Begun in 2002, Oklahoma's CTHFP was designed by the Oklahoma Department of Wildlife Conservation (ODWC) to provide Oklahoma metropolitan residents "quality fishing within a neighborhood-based fishing program by focusing on angler desires, use and benefits, and by implementing management techniques on urban ponds" (Gilliland, 2005). Over a dozen lakes and ponds in the Oklahoma City Metro area are included in this program. Three of these, Kid's Lake North (8 ha), Dolese Youth Pond (8 ha), and South Lake Park East (1.2 ha) were chosen for this study because they had well established adult fish populations at the time the study began. While Kid's Lake North and Dolese Youth Pond have been open for fishing in the program since 2002, the smaller South Lake Park East was recently renovated and was not opened to fishing until a stable adult population was established in Spring 2006. These ponds have established populations of sunfish *Lepomis spp.*, largemouth bass *Micropterus salmoides* and other Centrarchids, and are regularly stocked with channel catfish *Ictalurus punctatus*. Dolese Youth Pond is also stocked with rainbow trout *Oncorhynchus mykiss* during the winter. To fish these urban ponds, anglers must hold a state fishing license (USD $20) and a city fishing license, which is $3 daily or $15/year.[26] All largemouth bass caught must be released, but the state allows a bag limit of up to six catfish with no size limits (ODWC, 2008). For all other species, the CTHFP follows the state-wide regulations.

Survey design

Conjoint analysis is a marketing tool for analyzing consumers' demand for multi-attribute goods, in this case, a recreational experience. Many factors influence an angler's preferences for recreational sites. Therefore, conjoint analysis is an ideal tool for analyzing angler preference because it provides a framework for estimating demand for different combinations of potential qualities of the site such as docks, restrooms, bag limits, and size and type of fish stocked. Data for conjoint analysis can be obtained using a survey technique that presents respondents with a set of choices. Each option represents a potential management scenario with varied attributes (e.g., Figure 1). The respondent is asked to pick one of several options on each set of choices according to his or her preferences about that set of attributes for that bundle of characteristics of the site

[26] Oklahoma State Fishing licenses are $20 for individuals 18-63 years old; 16-17 year olds can buy a $5 license; and seniors over 64 years of age can buy a $6 lifetime license. Seniors and children 16 and under are exempt from purchase of a city fishing license.

and the angling experience at the location. Price can be included as one of the attributes to elicit a willingness-to-pay, from which the implicit marginal prices of the attributes can be estimated (Haab and McConnell, 2002; Baarsma, 2003; Freeman, 2003). In this study, a discrete choice experiment was used whereby the respondent chooses his/her preferred option, rather than ranking the options, a similar method called conjoint ranking; both techniques are subsets of conjoint analysis. Surveys to estimate recreation demand for angling and hunting have been used for the past twenty years. Some of these survey techniques have used actual data on trips (revealed preference data) such as travel cost models (Parsons, 2003). However, when it is difficult to survey for quality differences at many sites while controlling for unknown differences, the travel cost method may not yield clear answers. In these cases, stated preference methods such as contingent valuation (Loomis, 2006) and conjoint analysis (discrete choice and ranking experiments) have the advantage of eliciting preferences when management scenarios are hypothetical (Freeman, 2003).

Discrete choice experiments have been widely used to value environmental amenities, several of which focus on demand for outdoor recreation such as river flow (Adamowicz et al., 1994), caribou moose viewing and conservation (Adamowicz et al. 1988), rock climbing (Hanley et al., 2001), and waterfowl hunting (Mackenzie, 1990). This approach is particularly well suited to determining the values of urban anglers for alternative hypothetical management approaches, allowing managers to better evaluate the cost-benefit of the different options (Freeman, 2003). Fisheries economists have begun to examine anglers preferences for management alternatives using conjoint techniques, but as yet not in the urban setting or in conjunction with an on-site creel survey (Gillis and Ditton, 2002; Aas et al., 2000). While the effort required to collect such data can be expensive by mail or on site, choice set surveys can easily be added to traditional creel surveys that may already be planned, thus allowing for collection of needed data at little or no additional cost or effort.[27]

Data for this study were collected as part of a larger study that assessed fish stock size, growth rates, and mortality in combination with a creel survey conducted to ascertain angler demographics, catch and harvest data, and level of satisfaction. The creel survey was conducted from September 2006 to August 2008 at the three ponds using a roving creel clerk design. A convenience sample of all individuals on site was used. In this case, the majority of anglers on site were interviewed, resulting in a 97% response rate. Anglers were asked basic demographic information and trip characteristics for their current fishing trip. Each angler was also presented with two conjoint choice sets. Each choice set for potential management at the pond had three options for which the third was always the status quo at the pond where the angler was interviewed. Table 4.1 lists the seven pond attributes and their associated levels, which were used to construct the survey options. The seven attributes chosen as discrete management options included the size of catfish stocked, the length limit on catfish taken, the type of fish stocked, the availability of a fishing dock, the availability of restrooms, and an

[27] Interviewing anglers on site is known to result in avidity bias, i.e., responses have a disproportionate representation of users who frequently use the fishery because they are more likely to be interviewed. Thomson (1991) showed that this was likely to inflate expenditures and trip estimates per person in travel cost surveys on site relative to mail surveys. We are unaware of any such studies conducted for conjoint-choice studies, but it is likely that avid anglers' preferences may differ from those of other anglers and that estimates of visitation numbers based solely on the creel survey will also be upward biased. This is a shortcoming of using a creel and economic survey together, but the direction and ordering of preferences is unlikely to be affected. Furthermore, while conducting the conjoint on-site with planned creel activities may be cost effective, it does not provide insight on whether non-anglers might decide to participate should the proposed hypothetical management changes occur.

increase in the annual fishing license. The choice sets were orthogonally designed to eliminate collinearity between choices. Seven measurable attributes with CTHFP experiences of either 2, 3, or 4 levels were included in total. This created 2^4 X 3^1 X 4^2 = 768 possible combinations (management scenarios). Each combination was then randomly paired with another combination (option B). A third combination (option C) represented the status quo or no change. Norwood and Lusk (2005) demonstrated that using a random assignment of profiles from the full factorial performs well in terms of efficiency of the willingness to pay estimates. Respondents were asked to compare three alternatives simultaneously and to choose one of them. Figure 4.1 gives an example of a choice set. The survey design of randomized choice sets was created by generating a full factorial combination of all attribute levels and randomly assigning each potential combination with a different random combination in Excel 2007, but can be done in automated routines such as FACTEX in SAS (SAS Institute, 2004). A full survey design may also be created in SAWTOOTH™ software.

Econometric Model

A random Utility Model (RUM) was used to estimate the likelihood of respondent choice (Train, 2003). This model was assumed that when asked to choose between options A, B and C, respondents chose the option that gives them the highest utility (a measure of welfare or happiness). This condition is represented by

(4.1) $$U_j > U_k$$

where U is the respondent's utility. A respondent will select option j over k only if (4.1) holds for all k not equal to j, (i.e., the chosen option always gives him or her the highest satisfaction).

However, the real utility of the respondent is not known. Only the indirect utility function of the respondent denoted as V can be observed; the unobservable part of the utility that is unknown is denoted as ε. Therefore, the utility can be represented as

$$U = V + \varepsilon \tag{4.2}$$

The indirect utility function can be observed by using the answers to the discrete choice questions in which the attributes are arguments. Therefore, V can be expressed as a function of policy attributes accompanying each alternative. Therefore, the utility can be represented as

$$V_j = \mathbf{X\beta}, \qquad \forall j \in C \tag{4.3}$$

where $\mathbf{X}$ is the vector of policy attributes, $\boldsymbol{\beta}$ is a vector of unknown coefficients, and j is the alternative in choice C. For simplicity, V_j is assumed to be linear in X, so

$$V_j = \beta_0 + \beta_1(S_j) + \beta_2(L_j) + \beta_3(T_j) + \beta_4(B_j) + \beta_5(F_j) + \beta_6(PR_j) + \\ + \beta_7(FR_j) + \beta_8(P_j) \tag{4.4}$$

where S is the size of catfish stocked; L is the length limit of catfish taken; T is the type of fish stocked in ponds; B is the bag limit for catfish per day; F represents whether a fishing dock is provided in ponds; PR is a dummy variable for portable toilets; FR is the dummy variable for restrooms with flush toilets and running water; and *P* is the increase in the yearly city license fee dollars. $\beta's$ are the parameters to be estimated, and β_0 is the alternative specific constant which captures the effect in utility of a respondents selecting option C, the status quo, more often than options A and B in the sample (i.e., this measures if there is a status quo bias among respondents). This model 1 is referred as the basic model because it does not include any interaction terms (i.e., all ponds and demographic groups are considered the same with respect to the attributes that lead to utility).

In addition to the model specified in (4.4), two separate models were used to test for differences in preferences at different ponds and among different demographic groups as distinguished by income, race, age, and having children. By including ponds and demographic characteristics as an interaction terms for each attribute, the impacts are allowed to vary among respondents with different demographics. The first pond model is specified as:

$$V_j = \beta X + \delta(X * POND) \tag{4.5}$$

where *X* is a vector of variables specified in (4.4). *POND* is a vector of ponds, which is made up of Dolese Pond and South Lake Park, with Kid's Lake Park as the reference pond. Model 3, the Interaction Model also includes demographic characteristics. The interaction model is as follows:

$$V_j = \beta X + \delta(X * POND) + \alpha(X * Y) + \gamma(X * AGE) + \theta(X * CHILD) \tag{4.6}$$

where X is a vector of variables specified in (4.4). Y is a vector of income, which is separated into five categories. *AGE* is a vector of age, which is separated to four dummy variables that are coded to 1 if the individual is in that age group and zero otherwise. *CHILD* is a dummy variable to represent households with children, which is coded to one if a respondent has at least one child living in the household and zero otherwise.

From the choices made in each of two choice sets presented to an angler, a conditional logit model was employed to estimate equations (4.4), (4.5) and (4.6). The models estimate the probability that management option j would be chosen given option k was an alternative, where $j \neq k$. Then a basic model (equation (4.4)) with interaction terms for two of the three ponds (Equation 4.5) was estimated, with the third pond (Kid's Lake North) serving as the reference pond. This allowed us to see if there were pond-specific differences in the factors associated with utility. Pond Model 2 is reported in table 4.3 as an abbreviated model, reported after the full set of interactions were run. In addition to test the differences among demographic groups, equation (4.6) was estimated and reported as the interaction model in table 4.3. In order to estimate the willingness to pay, the pond attribute coefficients ($\beta_1 \ldots \beta_7$) were divided by the negative of the coefficient of city license fee ($-\beta_8$) for each model (Train, 2003). The delta method in Stata 10 (STATA, 2007) was used to compute the significance of the willingness to pay methods because both the pond attributes and fee coefficient have different standard errors. The resulting value is the marginal value or price (2008 U.S. Dollars, the same units used for the options for the city license fee) the respondent was willing to pay for that attribute. The conditional logit model was estimated in Stata, but would also be estimable in SAS MDC or any number of other statistical software packages (STATA, 2007; SAS, 2004).

Results and Discussion

A total of 568 respondents filled out the discrete choice questions. Descriptive statistics of these respondents are given in Table 4.2. The largest group of anglers ranged in age from 31-45 years old (35.56%). Forty one percent of anglers reported their household income as greater than $50,000 per year[28]. The majority of respondents were non-Hispanic white (72.68%). Other minority racial and ethnic groups observed included Asian, Hispanic, African-American and American Indian.[29]

The estimated valuation models provided coefficient estimates (β values) for all variables tested. Interpretation of these models is similar to linear multiple regression in that only coefficients (β) that are significantly different from zero should be considered, coefficients that have significant negative values indicate respondents were less likely to choose an option containing that attribute, and coefficients with significantly positive values indicate respondents were more likely to choose an option containing this variable. By dividing the estimated value of the coefficient by the negative of the fee coefficient, the marginal willingness to pay or value for an attribute such as larger catfish stocked can be computed. For example, the value of an increase in the bag limit for catfish was

[28] A separate study shows that the percentage of survey respondents at lower income households are underrepresented in the study compared to those in the general public using information on U.S. postal zip codes . In 2007, the median household income in Oklahoma County was $41,598 (U.S. Census, 2000).
[29] Soliciting information on ethnic and racial identity can be problematic. In this study, the most basic distinction was made between non-Hispanic/non-Latino whites versus all other ethnic and minority groups to illustrate that race and ethnicity has an effect on preferences even when controlling for differences in income. Often a more detailed analysis of both race and ethnic backgrounds should be considered. Policy managers interested in specifically comparing racial and ethnic composition to census data often use the census categories available at the www.census.gov or by the Office of Management and Budget (2000). Additional guidance on treatment of race in survey design is available in Stanfield and Dennis (1993).

negative ($\beta_{3/}$ -β_8 = 0.044/-(-0.192)=$-0.79; Train, 2003). This is interpreted, for example, as the willingness to accept the imposition of a 2 catfish per day limit if the yearly license fee were $0.79 less per year. If the bag limit were raised to 4 more catfish per day, the fee would have to be $1.58 lower per year to make anglers willing to accept it. In this case the angler dislikes the management change as shown by the negative value and must be compensated by a lower fee to make him or her as satisfied with the angling experience before the change. In this study, fee increases or decreases could be accomplished by changing the price of the required fishing license. The alternative specific constant for the status quo was statistically significant and negative for all three models, indicating that there is a specific preference against maintaining the status quo of management at the lakes. This means that on average, respondents for these two models were significantly likely to choose any option A or B that proposed changes in the management of the lakes in those models.

In the basic model that included all observations, all of the variables except the length limit on catfish significantly affected willingness to pay (Table 4.3).[30] Respondents were more likely to choose an option that had larger stocked catfish, a fishing dock, portable toilets, and restrooms with flush toilets. Anglers significantly preferred restrooms with flush toilets (β = 0.73) to portable toilets (β = 0.19)[31]. However, both facility choices were preferred to none, as theoretically expected. Respondents were significantly less likely to choose options with only catfish stocked (other species may

[30] Each of the 568 respondents saw 2 choice sets with 3 options resulting in n=3408 observations in the model.
[31] Within a single model, the coefficient that is higher and significant it is preferred on average to another significant coefficient.

still be present, but would not be maintained through stocking), more liberal daily bag limits on catfish, and higher license fees.

Table 4.4 gives the average angler's dollar values (willingness to pay) for each change in the level of attributes by model. An increase in stocked catfish size from 8 inches to 12 inches was valued at $0.23 per angler per year. Having only catfish stocked, rather than a mixture of bass, bluegill, and catfish available at a pond was worth $0.23 less per year on average, meaning diversifying the pond stock should be worth that much per year for an angler. A preference for a diverse fishery has been found in other community fishing surveys (Hunt and Ditton, 1996). Increasing the daily bag limit on catfish takes away $0.79 value for every 2 two additional fish an angler can keep. Anecdotal evidence from the creel clerks indicates that anglers perceived that higher bag limits would lower their own fishing success, rather than result in more catfish to take home. Other studies have similarly found anglers generally are highly supportive of bag restrictions (Hardin et al., 1987; Reed and Parsons, 1999; Edison et al., 2006), including anglers in urban environments (Hunt and Ditton, 1996). Furthermore, support for bag limits is higher from anglers in more densely populated areas (Edison et al., 2006), as would be the case in this study. Having a fishing dock would increase an angler's willingness to pay for an annual license by $1.28. Anglers were willing to pay $3.81 annually for flush toilets compared to having no restroom facilities (WTP for portable toilets was not significant when calculated using the delta method). Within the model, these results may be interpreted relatively to mean that the highest willingness to pay for a management change is for flush toilets followed by construction of a fishing dock.

Because choice-based analysis allows for estimated marginal values of each attribute, the individual willingness to pay may be seen as a relative measure of the benefit to the angler. However, the researcher's choice of payment mechanism, (e.g., higher license fees versus something such as a property tax increase) can affect an individuals' willingness to pay or even result in protest responses as shown in the contingent valuation method literature (Champ et al., 2003). In this case, pretesting did not result in protest bids and changes in license fees offered the most realistic payment vehicle. These results also allow for cost-benefit analysis of these options, even if there is no intention to actually raise license fees to cover the cost. For example, if park managers have actual or estimated visitation rates from the creel survey, the total value per year of having portable toilets could be computed (willingness to pay x number of anglers who buy the city license) and compared to the rental and maintenance rate for portable toilets (after adjusting for sampling bias). Unfortunately, Oklahoma City does not keep records on how many licenses are sold per year at this time, so such an analysis is not possible in this study. Furthermore, while useful, such multiplication would still provide a lower bound on value because new users might be attracted to visit a site by improvements.[32] A multi-year benefit-cost analysis that allows for catalogued benefits and costs in each year of occurrence , which are then discounted to the present could be performed for extensive infrastructure or biological improvements that have longer project lifetimes (Boardman et al., 2006 provides an excellent text).

[32] A limitation of using the discrete choice survey only on site with the creel survey is that potential users are not surveyed. At an additional cost, potential users might be surveyed by mail or internet (telephone surveys are ill-suited for the visual display needed for the choice sets).

Results from Pond Model 2 showed that willingness to pay for some pond management attributes significantly varied by pond for bag limits and flush toilets (Table 4.3). Willingness to pay differences were estimated by interacting each attribute with Dolese and South Lake Pond dummy variables for these two management attributes in order to test which attributes were significant (P-value<0.10) and should be included in a more limited model. The reported Pond Interaction Model 2 was then estimated and reported (Table 4.3), including only these two interactions that were significant at the 10 percent level in the initial model. Compared to Kid's Lake North and South Lake Park, anglers at Dolese Pond were more likely to choose scenarios with higher bag limits (β =0.05; P-value<0.10; table 4.3), but overall users at Dolese had an insignificant willingness to accept value when computed using the delta method perhaps because the interaction variable was only marginally significant at the 10 percent level (Table 4.4). By lake, table 4.4 shows that users at the other two lakes had a negative \$0.22 value per increase in the bag limit by 2 catfish per day. In addition, users at both Dolese (β =1.00) and South Lake Park (β =0.54) were willing to pay more for flush toilets than users of Kid's Lake North (table 4.3), and the computed willingness to pay for flush toilets was significant at \$3.42 (P-value<0.10; table 4.4)and \$5.60 (P-value<0.10; Table 4.4) per year at South Lake and Dolese ponds respectively, and insignificant at Kid's Lake North pond (Table 4.4). Kid's Lake North anglers did not have a significant willingness to pay, possibly because of its secluded location and the high relative abundance of trees and brush at the site that provide cover to anglers in lieu of facilities. The differences in willingness to pay among users of each of these ponds illustrates that while preferences may be similar in direction across a management program, the magnitude and specific issues may differ.

Interaction Model 3 shows that anglers' preferences differ by demographic attributes such as minority status, income, and age at the close to home ponds (Table 4.3). Compared to Caucasian anglers, minority anglers were less likely to choose options with higher fees for licenses, but since this coefficient proved insignificant (β = -0.10; P-value>0.10; Table 4.3), we may only interpret the willingness to pay values in the table 4.4 which proved significant using the delta method. In Table 4.4, estimates showed minorities had significantly lower willingness to pay for Fishing docks ($1.20 compared to $1.82 for non-Hispanic whites); Portable toilets ($0.90 compared to $1.36 for non-Hispanic whites); and flush toilets ($1.24 and $2.06 less at South Lake Park and Dolese Pond, respectively; Table 4.4). Minority anglers had to be compensated less for accepting increases in bag limits on catfish at Kid's Lake and South Lake Park ponds (-$0.48 for a 2 catfish increase) than non-Hispanic whites (-$.72 per 2 fish increase) at the same ponds (Table 4.4). This reduced willingness to pay higher fees would be expected to reduce the values estimated for all tested attributes, because willingness to pay for each attribute is calculated by dividing the estimated coefficient for each attribute by the estimated coefficient for the fee. Households with children were slightly less averse to catfish stocking than the average household (-$6.19 compared to $-6.57; Table 4.4). For households with children, variety in the angling activity may not be as important as simply catching a fish

Age and Income also affected willingness to pay for attributes such as the size of catfish stocked and bag limits for catfish (Table 4.3 and 4.4). Compared to anglers over

60 years of age, the reference age group, anglers in the 31-45 group were less likely to choose scenarios with larger catfish (β=-0.06; table 4.3) and more likely to choose higher bag limits (β=0.11; Table 4.3), but Table 4.4 shows that the computed willingness to pay values were insignificant. Anglers less than 30 years of age were more likely to choose higher bag limits (β=0.10), but the willingness to pay values also proved insignificant. Willingness to pay only differed significantly from the reference group (over 60 years of age) at Kid's Lake Park and South Lake Park for 46-60 year olds who had a negative $0.52 value per given increase in size of fish stocked per year (increase in the size of catfish stocked from 8 to 12 inches, Table 4.4).[33] These results suggest that all groups under 60 might be more harvest oriented. Alternatively, older anglers may feel that higher bag limits would decrease their own individual probability of catching fish.

Although anglers on average were willing to pay for larger stocked catfish when grouped as a whole, the levels varied by income. Anglers with households earning between $30,000 and 50,000 were willing to pay $0.70 per year more for an increase in the size of catfish stocked compared to the lowest income bracket (below $20,000 per year). Responses from anglers in the top three income categories suggest than an increase in catfish bag limits by two at Kid's Lake and South Lake Park ponds lowered the value of their license (Table 4.4). At these two ponds, anglers in the $30,000-50,000, $50,001-100,000, and greater than $100,001 household income brackets find their licenses devalued by -$0.62, -$.50, and -$1.50 per year per increase in catfish bag limit per day. Thus, the wealthiest anglers were most opposed to bag limit increases, perhaps

[33] It is important to remember that willingness to pay for interacted categories is computed by adding the interaction term coefficient to the coefficient for the base and dividing by the fee coefficient and then taking the negative of the result to obtain the categorical willingness to pay (WTP(46-60*Bag Limit)=-(-0.14+.0.04)/-0.19). The delta uses the sandwich standard error to compute the significance of the WTP estimate.

because of higher opportunity costs of time to fish elsewhere. Finally, anglers in households with an income over 100,000 were slightly less opposed to bag limits at Dolese (WTP=-$1.20 per 2 catfish increase in bag limit per year) than at the other two ponds. Anglers in households earning under $30,000 per year have no significantly different willingness to pay values by species stocking, bag limit, and length limit (Table 4.4), an interesting result because urban fishing programs are often intended to serve less affluent city dwellers who lack the means and opportunity to fish elsewhere (Botts, 1984). However, as the demographic data shows, individuals between 18 and 45 had negative values for increases in the size of catfish stocked but when the results were tested by income, value increases in stocked catfish size, particularly for those in the middle income bracket of $30,000-50,000 annually. Little is published on the opinions of anglers by age group or income level. This study results suggest differences in opinion among individuals of different age and income should be considered in future studies.

Conclusions

The results of this discrete choice survey show that anglers are willing to pay for increases in management effort such as larger stocked catfish and increases in variety in fish stocked in CTHFP ponds. However, relatively speaking, anglers are more willing to pay more for physical amenities such as docks and restroom facilities in the urban setting. This is consistent with other studies that suggest amenities are critical to the success of an urban fishing program (see review by Balsman and Shoup, 2008). This study also found that anglers do not desire increased bag limits, but are less opposed at Dolese Youth Park. Support for bag restrictions has been observed in other angler surveys, both in urban fishing programs (Hunt and Ditton 1996) and non-urban settings (Hardin et al., 1987;

Reed and Parsons, 1999; Edison et al., 2006). This study results suggest Oklahoma City anglers differ in this respect. Length limits were insignificant in all models, suggesting this type of regulation is not a high priority for anglers. Length limits are imposed in this fishery to maintain a longer average length of fish captured than what would likely occur without the limit (reproduction is not a consideration in this put-grow-take fishery). In terms of providing fishing experiences for minority households, the results show that these groups are willing to pay less on average for most management improvements. Therefore, if an urban program is designed to target these groups, it is important to seek funding from sources other than increased license fees because higher costs may discourage these demographic groups from participating (Balsman and Shoup, 2008).

Given that detailed valuation estimates of anglers' demand for different attributes of managed angling sites can be time consuming and costly, managers may be tempted to use value estimates from another site for a new location, something called benefits transfer. Unfortunately, as the pond interaction variables show, such assumptions may miss variation among users. Meta-analysis of multiple studies of value for characteristics might be used to adjust welfare measures of value, but given the lack of studies on urban fisheries, site-specific research is still necessary. Furthermore, the goals of a different project may diverge from those of the CTHFP, which looked at physical amenities for direct use by anglers, i.e., manmade docks, restrooms, and stocking rates. Furthermore, adjustments for study methodology would need to be made (Johnston et al., 2006). However, this study illustrates that valuation data for current users can easily be obtained in conjunction with creel surveys that may already be planned. Using the methods described in this study, these data can be used to determine marginal willingness to pay for management options or amenities, but obtaining the proper expertise for designing the choice experiment is key. Once obtained, the marginal willingness to pay for a change can be readily compared to the increased cost of implementing the change using benefit-cost analysis. If the willingness to pay is greater than the cost, then the agency should implement the change, if the budget allows. Another potential benefit of using conjoint choice in other settings is the ability to value other non-market values besides recreational use, such as angler's willingness to pay for species preservation, maintenance of in-stream flows for wildlife, or improved water quality.

Table 4.1. Site Attributes and Levels

Attribute	Attribute Levels
Size of Catfish Stocked	8 inches
	12 inches
Length Limitation of Catfish taken	None
	12 inches
Bag Limit for Catfish per day	4
	6
	8
	10
Type of Fish in Ponds	Bass, Bluegill, and Catfish
	Catfish Only
Fishing Dock	None
	1 Open dock
Restroom	None
	Portable toilets
	Restroom with flush toilets and running water
Increase in the yearly city license	$0
	$2
	$4
	$6

Table 4.2. Descriptive Statistics of Attribute Level and Angler Respondents Using the Close-to-Home-Fishing Ponds in the Oklahoma City Metro Area (2006-2008).

Variable	Definition	Mean
Size of Catfish Stocked	8 if 8 inch, 12 if 12 inch	9.33
Length Limitation on Catfish taken	1 if 12 inch, 0 if no limit	0.34
Catfish Only	1 if catfish only, 0 if bass, bluegill, and catfish	0.38
Bag Limit for Catfish per day		3.92
	0 if no limit	
	2 if 2 catfish limit	
	4 if 4 catfish limit	
	6 if 6 catfish limit	
Fishing Dock	1 if 1 open dock available, 0 if none	0.35
Restroom		
None (use as base)	1 if no restroom, 0 otherwise	0.23
Portable Toilets	1 if having portable toilets, 0 otherwise	0.22
Restroom with Flush Toilets	1 if having restroom with flush toilets, 0 otherwise	0.22
License Fee		2.03
	0 if no increase in the yearly city license	
	2 if $2 increase in the yearly city license	
	4 if $4 increase in the yearly city license	
	6 if $6 increase in the yearly city license	
Age		
Less than 31	1 if age is less than 31, 0 otherwise	23.39%
31-45	1 if age is between 31 to 45, 0 otherwise	35.56%
46-60	1 if age is between 46 to 60, 0 otherwise	25.18%
More than 60 (use as base)	1 if age is more than 60, 0 otherwise	15.87%
Yearly Income		
< 20001 (use as base)	1 if yearly income is less than 20001, 0 otherwise	13.38%
20001-30000	1 if yearly income is between 20001 to 30000, 0 otherwise	16.02%
30001-50000	1 if yearly income is between 30001 to 50000, 0 otherwise	25.52%
50001-100000	1 if yearly income is between 50001 to 100000, 0 otherwise	38.38%
> 100000	1 if yearly income is higher than 100000, 0 otherwise	6.69%
Minority	1 if minority, 0 if white	
White (non-Hispanic)		72.68%
Minority	Asian, African-American, American Indian, other race or Hispanic (ethnicity)	27.32%
Child	1 if there are children in household, 0 if none	43.66%

Note: Mean is based on a total of 568 collected surveys, 39 respondents chose not to report their annual income.

Table 4.3. Conditional Logit Regression Results

Variable	Basic Model 1	Pond Model 2	Interaction Model 3
Size of Catfish Stocked	0.04*	0.05**	0.06
	(0.02)	(0.02)	(0.062)
Length Limitation on Catfish taken	0.12	0.06	0.15
	(0.08)	(0.08)	(0.13)
Bag Limit of Catfish per day	0.15***	-0.05*	-0.14**
	(0.03)	(0.03)	(0.06)
Catfish Only	-0.05**	-1.21***	-1.27***
	(0.021)	(0.11)	(0.11)
Fishing Dock	0.25***	0.33***	0.35***
	(0.09)	(0.09)	(0.10)
Portable Toilets	0.19*	0.25**	0.26**
	(0.10)	(0.12)	(0.12)
Restroom with Flush Toilet	0.73***	0.18	0.20
	(0.11)	(0.22)	(0.23)
License Fee	-0.19***	-0.21***	-0.19***
	(0.02)	(0.02)	(0.02)
Pond Interactions			
Dolese*Bag Limit		0.05*	0.06**
		(0.02)	(0.03)
South Lake Park*Flush Toilet		0.54**	0.51*
		(0.27)	(0.29)
Dolese Pond*Flush Toilet		1.00***	0.98***
		(0.26)	(0.27)
Minority Interaction			
Minority*License Fee			-0.10
			(0.06)
Age Interactions			
0-30*Size of Catfish			-0.02
			(0.06)
31-45*Size of Catfish			-0.06
			(0.04)
46-60*Size of Catfish			-0.01
			(0.04)

Note: Standard errors are in parenthesis, and they are calculated using sandwich estimator of variance. Asterisks indicate significance as follows: P<0.10*, P<0.05**, P<0.01***.

Table 4.3. Conditional Logit Model Results (Cont.)

Variable	Basic Model 1	Pond Model 2	Interaction Model 3
0-30*Bag Limit			0.10**
			(0.05)
31-45*Bag Limit			0.11**
			(0.06)
46-60*Bag Limit			0.04
			(0.06)
Income Interaction			
>20k-30k*Size of Catfish			-0.11
			(0.08)
>30k-50k*Size of Catfish			0.08
			(0.07)
>50k-100k*Size of Catfish			0.01
			(0.06)
>100k*Size of Catfish			0.05
			(0.10)
>20k-30k*Bag Limit			0.07
			(0.06)
>30k-50k*Bag Limit			0.02
			(0.05)
>50k-100k*Bag Limit			0.04
			(0.05)
>100k*Bag Limit			-0.15**
			(0.07)
Child Interaction			
Size of Catfish*Child			0.04*
			(0.02)
Catfish Only*Child			0.07
			(0.05)
Length Limit*Child			-0.18
			(0.17)
ASC	-0.606***	-0.87***	-0.83***
	(0.207)	(0.22)	(0.28)
Log Likelihood	-1,452.101	-1376.39	-1349.90
No. of Observation	3,408	3,408	3,408

Note: Standard errors are in parenthesis, and they are calculated using sandwich estimator of variance. Asterisks indicate significance as follows: P<0.10*, P<0.05**, P<0.01***.

Table 4.4. Willingness to Pay (WTP; U.S. $ 2008) by Management Attribute of Anglers Using the Close-to-Home-Fishing Program ponds in the Oklahoma City Area (2006-2008)

WTP	Basic Model 1	Pond Model 2	Interaction Model 3						
					Annual Household Income				
	Mean	Mean	Mean	Minority	>20k-30k	>30k-50k	>50k-100k	>100k	Children in Household
Size of Catfish Stocked	**$0.23**	**$0.25**	$0.30	$0.20	-$0.29	**$0.70**	$0.35	$0.54	$0.51
Length Limitation on Catfish taken	$0.62	$0.30	$0.79	$0.52					-$0.12
Catfish Stocked only	**-$0.23**	**-$5.75**	**-$6.57**	**-$4.34**					**-$6.19**
Bag Limit for Catfish/day	**-$0.79**								
Kid's Lake		**-$0.22**	**-$0.72**	**-$0.48**	-$0.36	**-$0.62**	**-$0.50**	**-$1.50**	
South Lake Park		**-$0.22**	**-$0.72**	**-$0.48**	-$0.36	**-$0.62**	**-$0.50**	**-$1.50**	
Dolese Pond		-$0.03	-$0.43	-$0.28	-$0.07	-$0.32	-$0.21	**-$1.20**	
Fishing Dock	**$1.28**	**$1.58**	**$1.82**	**$1.20**					
Portable Toilets	$0.98	**$1.19**	**$1.36**	**$0.90**					
Flush toilets	**$3.81**								
Kid's Lake		$0.84	$1.05	$0.70					
South Lake Park		**$3.42**	**$3.67**	**$2.43**					
Dolese Pond		**$5.60**	**$6.11**	**$4.05**					

Note: Delta method was applied to clarify the significant level of WTP. Values for which P<0.10 are indicated in bold.

Table 4.4. Willingness to Pay (Cont.).

WTP	Interaction Model 3 (Cont.)		
	Age Group (years)		
	<31	31-45	46-60
Size of Catfish Stocked	$0.18	-$0.02	-$0.24
Length Limitation on Catfish taken			
Catfish Stocked only			
Bag Limit for Catfish/day			
Kid's Lake	-$0.18	-$0.14	**-$0.52**
South Lake Park	-$0.18	-$0.14	**-$0.52**
Dolese Pond	$0.12	$0.15	-$0.22
Fishing Dock			
Portable Toilets			
Flush toilets			
Kid's Lake			
South Lake Park			
Dolese Pond			

Note: Delta method was applied to clarify the significant level of WTP. Values for which P<0.10 are indicated in bold.

Figure 4.1. Example of a Discrete Choice Set for Management Options used to Assess Angler Willingness to Pay for Management Options at Close-to-Home-Fishing Program Ponds in the Oklahoma City Metro Area from 2006 – 2008.

Below you will find three management scenarios being considered to improve the close to home fishing program. Please choose **one** of the following options below.

Attribute	Option A	Option B	Option C
Size of Catfish Stocked	8 inches	8 inches	**NO CHANGE:** I would rather keep the management of this pond the way it is today and not pay any increase in the city license fee.
Length Limit on Catfish taken	No length limit	12 inches	
Bag Limit for Catfish per day	4	10	
Type of Fish in the Pond	Catfish only	Catfish only	
Fishing Dock	None	1 open dock	
Restrooms	Porta-potties	None	
Increase in the yearly city license (Dollar)	$2 increase	$2 increase	
I would choose (Please check only one)	☐**A**	☐ **B**	☐ **C** (I would not want either A or B

REFERENCES

Aas, O., W. Haider, and L. Hunt, 2000. "Angler responses to potential harvest regulations in a Norwegian sport fishery: a conjoint-based choice modeling approach." *North American Journal of Fisheries Management* 20: 940-950.

Oklahoma Tourism and Recreation Department 2007. "About Oklahoma: General Information." Available: http://www.travelok.com/about/index.asp#01. (January, 2008).

Adamowicz, W. L. 1988. "Behavioral Implications of Nonmarket Valuation Models." *Canadian Journal of Agricultural Economics* 36:929-939.

Adamowicz, W., J. Louviere, and M. Williams. 1994. "Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities." *Journal of Environmental Economics and Management* 26:271-92.

Adamowicz, W.L., J.J. Louviere, and M. William. 1997. "Perceptions versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation." *Journal of Environmental Economics and Management* 32: 65-84.

Arlinghaus R. and T. Mehner. 2004. "A Management-Orientated Comparative Analysis of Urban and Rural Anglers Living in a Metropolis (Berlin, Germany)." *Environmental Management* 33:331-344.

ASA&AFWA (American Sportfishing Association and the Association of Fish and Wildlife Agencies). 2007. Lifestyles and license buying habits of America's anglers: a five-year examination of U.S. license buyers. Southwick Associates, Inc. Available: http://www.rbff.org/uploads/Research_section/Related_Research/ASA-AFWA_License_Analysis_Summary_Report_2006.pdf. (March 2009).

Azevedo, C.D., J.A. Herriges, and C.L. Kling. 2001. "Valuing Preservation and Improvements of Water Quality in Clear Lake." Staff Report 01-SR 94, Center for Agricultural and Rural Development, Iowa State University.

_________. 2003. "Combining Revealed and Stated Preferences: Consistency Tests and Their Interpretations." *American Journal of Agricultural Economics* 85(3): 525-537.

Baarsma, B. E. 2003. "The Valuation of The Ijmeer Nature Reserve using Conjoint Analysis." *Environmental and Resource Economics* 25:343-356.

Balsman, D. M. and D. E. Shoup. 2008. "Opportunities for urban fishing: developing urban fishing programs to recruit and retain urban anglers." *American Fisheries Society Symposium* 67:31-40.

Barta, S., L. O'Connell, J. Wiliams, G. Doekson, B. Whitacre, and M. Woods. 2007. Economic conditions and trends in rural Oklahoma. (11th Edition) Oklahoma Cooperative Extension Service, Oklahoma State University. Available: www.rd.okstate.edu/. (September 2009).

Ben-Akiva, M., and S.R. Lerman. 1985. *Discrete choice analysis: Theory and application to travel demand*. Cambridge: MIT Press.

Ben-Akiva, M., and T. Morikawa. 1990. "Estimation of Switching Models from Revealed Preferences and Stated Intentions. *Transportation Research* 24(6): 485-495.

Beneficial Use Monitoring Program 2007. Oklahoma Water Resources Board 2007. Online. Available: www.owrb.ok.gov/quality/mornitoring /bump.php. (January 2008.)

Boardman, A.E., D.H. Greenberg, A.R. Vining, and D.L. Weimer. 2006. Cost-benefit analysis: concepts and practice. Pearson Education, New Jersey, U.S.A.

Bockstael, N.E., W.M. Hanemann, and C.L. Kling. 1987. "Estimating the Value of Water Quality Improvements in a Recreational Demand Framework." *Water Resource Research* 23 (5): 951-960.

Botts, L. 1984. Symposium summary. In L. A. Allen, eds. *Urban fishing symposium proceedings*. 284-288. American Fisheries Society, Bethesda, Maryland.

Boxall, P.C., J. Englin,and W.L. Adamowicz. 2003. "Valuing Aboriginal Artifacts: A Combined Revealed-Stated Preference Approach." *Journal of Environmental Economics and Management* 45 (2): 213-230.

Burlingame, M. N., and C. S. Guy, 1999. "Diversity among anglers in Kansas: A focus on channel catfish anglers." *American Fisheries Society Symposium* 24:427-433.

Caneday, L. 2000. "First Year Report: Development of Information to Evaluate Customer Requirements for Recreation Facilities and Services in the Corps of Engineers, Tulsa District." Working Paper, Leisure Study, College of Education, Oklahoma State University.

_________. 2001. "Oklahoma Statewide Comprehensive Outdoor Recreation Plan: The Public Recreation Estate." Oklahoma Tourism and Recreation Department Division of Planning and Development. Available: http://www.otrd.state.ok.us/rd/Scorp-2001.pdf. (April, 2010).

Caneday, L., and D. Jordan. 2003. "State Park Visitor Survey: 2002-2003." Working Paper, Leisure Study, College of Education, Oklahoma State University.

Caneday, L., D. Jordan, P. Brown, T.J. San Diego, K. Smith, and K. Fink. 2007. "A Second Century of Outdoor Recreation in Oklahoma: 2007 Statewide Comprehensive Outdoor Recreation Plan." Working Paper, Oklahoma Tourism and Recreation Department.

Champ, P.A., K.J. Boyle, and T.C. Brown. 2003. *A primer on nonmarket valuation*. Dordrecht: Kluwer Academic Publishers.

David, G.M., C.M. Angela, and G.F. Croker. 1991. "Water Clarity Criteria for Bathing Waters Based on User Perception." *Journal of Environmental Management* 33: 285-299.

Dillman, D.A. 2000. *Mail and Internet Surveys: The Tailored Design Method*. 2nd Edition, New York: John Wiley & Sons.

Dunlap, R. E., and R. B. Heffernan. 1975. "Outdoor recreation and environmental concern: an empirical examination." *Rural Sociology* 40(1):18-30.

Eades R.T. , J. W. Neal, T. J. Lang, K. M. Hunt, and P. Pajak, editors 2008. Urban and Community Fisheries Programs: Development, Management, and Evaluation. American Fisheries Society, Symposium 67, Bethesda, Maryland.

Earnhart, D. 2001. "Combining Revealed and Stated Preferences Methods to Value Environmental Amenities at Residential Locations." *Land Economics* 70(1): 12-29.

_________. 2004. "Time is Money: Improved Valuation of Time and Transportation Costs." *Environmental and Resource Economics* 29(2): 159-190.

Edison, T. W., D. H. Wahl, M. J. Diana, D. P. Philipp, and D. J. Austen. 2006. "Angler Opinion of Potential Bluegill Regulations on Illinois Lakes: Effects of Angler Demographics and Bluegill Population Size Structure." *North American Journal of Fisheries Management* 26:800-811.

Feather P., D. Hellerstein, and T. Tomasi. 1995. "A Discrete Count Model of Recreational Demand." *Journal of Environmental Economics and Management* 29: 214-227.

Fedler, A. J. 2000. Participation in boating and fishing: a literature review. Recreational Boating and Fishing Foundation. Alexandria, Virginia.

_________. 2007. Take me fishing: a summary of key learnings from state marketing programs. RBFF. Available: http://www.rbff.org/uploads/Research_section/Program_ Evaluation _Files/Key_Learnings_Report_-_FINAL.pdf. (March 2009).

Foster, F.D., T. Smith, and R.E. Whaley. 1997. "Assessing Goodness-of-fit of Asset Pricing Models: The Distribution of the Maximum R^2." *Journal of Finance* 52(2): 591-607.

Freeman, A.M. 2003. *The Measurement of Environmental and Resource Values: Theory and Methods*. Resources for the Future, Washington, DC.

Gilliland, E. R. 2005. Oklahoma Close-to-Home Fishing Program. Oklahoma Department of Wildlife Conservation, Federal Aid in Sport Fish Restoration, Project F-54-D-3, Performance Report, Oklahoma City.

Gillis,K.S. and R B. Ditton. 2002. "A Conjoint Analysis of U.S. Atlantic Billfish Fishery Management Alternatives." *North American Journal of Fisheries Management* 22: 1218-1228.

Goyal, A., and I. Welch. 2003. "Predicting the Equity Premium with Dividend Ratios." *Management Science* 49(5): 639-654.

Greene, W.H. 2003. *Econometric Analysis*, 5th. ed. New Jersey: Prentice Hall.

Grijalva, T.C., R.P. Berrens, A.K. Bohara, and W.D. Shaw. 2002. "Testing the Validity of Contingent Behavior Trip Responses." *American Journal of Agricultural Economics* 84(2): 404-414.

Grogger, J. T., and R.T. Carson. 1991. "Models for Truncated Counts." *Journal of Applied Econometrics* 6: 225-238.

Haab, T. C., and K. E. McConnell. 1996. "Count Data Models and the Problem of Zeros in Recreation Demand Analysis." *American Journal of Agricultural Economics* 78: 89-102.

_________. 2002. *Valuing Environmental and Natural Resources,* Cheltenham, UK: Edward Elgar.

Haener, M.K., P.C. Boxall, and W.L. Adamowicz. 2001. "Modeling Recreation Site Choice: Do Hypothetical Choices Reflect Actual Behavior?" *American Journal of Agricultural Economics* 83(3): 629-642.

Hanley, N. 1989. "Valuing Non-Market Goods Using Contingent Valuation." *Journal of Economic Surveys* 3: 235-252.

Hanley, N., and C.L. Spach. 1993. *Cost Benefit Analysis and the Environment*, UK: Edward Elgar Publishing Company.

Hanley, N., S. Mourato, and R. E. Wright. 2001. "Choice Modeling Approaches: A superior alternative for Environmental Valuation?" *Journal of Economic Surveys* 15:435-462.

_________. 2003. "Estimating the Monetary Value of Health Care: Lessons from Environmental Economics." *Health Economics* 12(1): 3-16.

Hardin, S., J. D. Rayburn, and G. Posnansky. 1987. "Attitudes, practices and preferences of licensed freshwater anglers in Florida." *Proceedings of the Annual Conference Southeastern Association Fish and Wildlife Agencies* 41:168–177.

Hausman J. A., G. K. Leonard, and D. McFadden. 1995. "A Utility-Consistent, Combined Discrete Choice and Count Model Assessing Recreational Use Losses due to Natural Resource Damage." *Journal of Public Economics* 56: 1-30.

Hensher, D.A., J.J. Louviere, and J.D. Swait. 1999. "Combining Sources of Preference Data." *Journal of Econometrics* 89(1): 197-221.

Herriges, J.A., C.L. Kling, and D.J. Phaneuf. 1999. “Coner Solution Models of Recreation Demand: A Comparison of Competing Frameworks.” In J.A. Herriges, and C.L. Kling, eds. *Valuing Recreation and the Environment*. 163-197. Cheltenham, UK: Edward Elgar.

Hubbell, B. J., M. C. Marra, and G. A. Carlson. 2000. “Estimating the Demand for a New Technology: BT Cotton and Insecticide Policies.” *American Journal of Agricultural Economics* 82: 118-132.

Hunt, K. A., and R. B. Ditton. 1996. A Statewide Perspective on Community Fishing in Texas. Texas A&M University Human Dimensions of Fisheries Research Laboratory, Report HD-606, College Station, Texas.

_________. 1997. “The Social Context of Site Selection for Freshwater Fishing.” *North American Journal of Fisheries Management* 17(2):331-338.

Johnston, R. J., E. Y. Besedin, and M.H. Ranson. 2006. “Characterizing the effects of valuation methodology in function-based benefits transfer.” *Ecological Economics* 60:407-419.

Kellert, S. R., and M. O. Westervelt. 1983. Children’s Attitudes, Knowledge and Behaviors Towards Animals. U.S. Fish and Wildlife Service, Washington DC: U.S Government Printing Office.

Kjaer, T. 2005. “A Review of the Discrete Choice Experiment with Emphasis on its Application in Health Care.” Health Economics Papers 2005:1, University of Southern Denmark. Available: http://www.sdu.dk/~/media/52E4A6B76FF340C3900EB41CAB67D9EA.ashx. (January 2009).

Lenard, V., and D.D. Badger. 1979. “Impact of Recreation Use at Two New Lakes in Oklahoma: Kaw and Hugo.” OSU Profession Paper P-610, Oklahoma Agricultural Experiment Station.

Lo, A.W., and A.C. MacKinlay. 1990. “Data-Snooping Biases in Tests of Financial Asset Pricing Models.” *Review of Financial Studies* 3(3): 431-467.

Loomis. J. 2006. “Use of Survey Data to Estimate Economic Value and Regional Economic Effects of Fishery Improvements.” *North American Journal of Fisheries Management* 26:301–307.

Louviere, J.J., D.A. Hensher, and J.D. Swait. 2000. *Stated Choice Methods: Analysis and Application*, Cambridge: Cambridge University Press.

Lusk, J.L., and F.B. Norwood., 2005. “Effect of Experimental Design on Choice-Based Conjoint Valuation Estimates.” *American Journal of Agricultural Economics* 87(5): 771-785.

Mackenzie, J. 1990. “Conjoint Analysis of Deer Hunting.” *Northeastern Journal of Agricultural and Resource Economics* 10:109-117.

Martina, V., and K. Neha. 2007. “Can Voluntary Pollution Prevention Programs Fulfill their Promises? Further Evidence from the EPA’s 33/50 Program.” *Journal of Environmental Economics and Management* 53: 180-195.

Morikawa, T. 1994. “Correcting State Dependence and Serial Correlation in the RP/SP Combined Estimation Method.” *Transportation* 21(2): 153-165.

Murphy, K.M., and R.H. Topel. 1985. “Estimation and Inference in Two-Step Econometric Models.” *Journal of Business & Economic Statistics* 4(3): 88-97.

Noble, R. L., and T. W. Jones. 1999. “Managing fisheries with regulations.” In C. C. Kohler, and W. A. Hubert, eds. *Inland fisheries management*, 2nd edition. 455-480. American Fisheries Society, Education Section and Fisheries Management Section, Bethesda, Maryland.

Norwood, B., J. Lusk, and T. Boyer. 2007. “Forecasting Hypothetical Bias: A Tale of Two Calibrations.” In *Experimental Methods in Environmental Economics*, Cherry, T.L., S. Kroll, and J.F. Shogren, eds. 447-465. Routledge, NY.

ODWC (Oklahoma Department of Wildlife Conservation). 2008. "Close to Home" fishing Opportunities. Available: http://www.wildlifedepartment.com/fishregs/closetohome.htm. (January 2009).

Oklahoma Tourism and Recreation Department. 2007. “About Oklahoma: General Information.” Available: http://www.travelok.com/about/index.asp#01. (January 2008).

Parsons, G. 2003. “The Travel Cost Model.” In P. A. Champ, K. J. Boyle, and T. C. Brown, eds. *A Primer on Nonmarket Valuation*. 269-330. Massachusetts: Kluwer Academic Publishers.

Parsons G. R., and M. J. Kealy. 1995. “A Demand Theory for Number of Trips in a Random Utility Model of Recreation.” *Journal of Environmental Economics and Management* 29: 357-367.

Parsons, G.R., P.M. Jakus, and T. Tosami. 1999. “A Comparison of Welfare Estimates from Four Models for Linking Seasonal Recreational Trips to Multinomial Logit Models of Site Choice.” *Journal of Environmental Economics and Management* 38: 143-157.

Petrin, A., and K. Train. 2002. Omitted Product Attributes in Discrete Choice Models. Available: http://elsa.berkeley.edu/wp/train1202.pdf. (August, 2009).

Pinar, K. M., and K. Train. 2003. Standard Error Correction in Two-Stage Estimation with Nested Samples. Available: http://emlab.berkeley.edu/users/train/mtrevision.pdf. (September, 2009).

Quist, M. C., J. S. Tillma, M. N. Burlingame, and C. S. Guy. 1999. "Overwinter Habitat Use of Shovelnose Sturgeon in the Kansas River." *Transactions of the American Fisheries Society* 128:522-527.

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

_________. 1998. "Beyond the Crucial Experiment: Mapping the Performance Characteristics of Contingent Valuation." *Resource and Energy Economics* 20: 197-206.

RBFF (Recreational Boating & Fishing Foundation). 2007. MN marketing program: executive summary. RBFF. Available: http://www.rbff.org/uploads/Research_section/Program_Evaluation_Files/MN_Marketing_Program_Executive_Summary_-_FINAL.pdf. (March 2009).

Reed, J. R., and B. G. Parsons. 1999. "Angler Opinions of Bluegill Management and Related Hypothetical Effects on Bluegill Fisheries in Four Minnesota Lakes." *North American Journal of Fisheries Management* 19:515–519.

Ryan, M., and S. Wordsworth. 2000. "Sensitivity of Willingness to Pay Estimates to the Level of Attributes in Discrete Choice Experiments." *Scottish Journal of Political Economy* 47(5): 504-524.

SAS Institute. 2004. SAS/STAT user's guide, version 9.1.3. SAS Institute, Cary, North Carolina.

Schramm, H. L., Jr., and J. A. Dennis. 1993. "Characteristics and Perceptions of Users and Nonusers of an Urban Fishery Program in Lubbock, Texas." *North American Journal of Fisheries Management* 13:210-216.

Siemer, W. F., and B. A. Knuth. 2001. "Effects of Fishing Education Programs on Antecedents of Responsible Environmental Behavior." *The Journal of Environmental Education* 32(4):23-29.

Smith, V. K. 1997. Combining Discrete Choice and Count Data Models: A Comment. Manuscript Duke University.

Stanfield, J.H. and R.M. Dennis. 1993. *Race and Ethnicity in Research Methods*. Newbury Park: Sage Publications, Inc.

Starbuck, C. M., S. J. Alexander, R. P. Berrens, and A. K. Bohara. 2004. "Valuing Special Forest Products Harvesting: A Two-Step Travel Cost Recreation Demand Analysis." *Journal of Forest Economics* 10: 37-53.

StataCorp. 2007. Stata Statistical Software: Release 10. StataCorp LP. College Station, TX.

Summers, G. 2008. Using Geographic Information Systems to Create and Analyze Spatial Relationships from Oklahoma Angler Data. Oklahoma Department of Wildlife Conservation, Federal Aid in Sport Fish Restoration, Project F-50-R-14, Final Performance Report, Oklahoma City.

Swait, J., and J. Louviere. 1993. “The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit-Model.” *Journal of Marketing Research* 30(3): 305-314.

Thomson, C.J. 1991. “Effects of the Avidity Bias on Survey Estimates of Fishing Effort and Economic Value.” *American Fisheries Society Symposium* 12:356-366.

Titenberg, T., and L. Lewis. 2009. *Environmental Economics and Policy*. New Jersey: Prentice Hall.

Toth, J. F., Jr., and R. B. Brown. 1997. “Racial and Gender Meanings of Why People Participate in Recreational Fishing.” *Leisure Sciences* 19:129-146.

Train, K. E. 2003. Discrete choice methods with simulation. Cambridge University Press, Cambridge, U.K.

TransCAD. 2008. *TransCAD Transportation Planning Software: Release 5.* Newton, MA: Caliper Corporation.

U.S. Census Bureau. 2000. American factfinder. U.S. Census Bureau. Available: http://factfinder.census.gov/servlet/DTTable?_bm=y&-geo_id=01000US&-ds_name=DEC_2000_SF1_U&-_lang=en&-mt_name=DEC_2000_SF1_U_P002&-format=&-CONTEXT=dt. (March 2009).

U.S. Census Bureau. 2009. U.S. Census Bureau: State and County QuickFacts.Oklahoma County. Available: http://quickfacts.census.gov/qfd/states/40/40109.html (February 2010).

U.S. Department of the Interior (U.S. Fish and Wildlife Service and U.S. Department of Commerce, U.S. Census Bureau). 2002. 2001 national survey of fishing, hunting, and wildlife-associated recreation. U.S. Government Printing Office, Washington, D.C.

U.S. Fish and Wildlife Service. 2007. 2006 national survey of fishing, hunting, and wildlife-associated recreation national overview. U.S. Government Printing Office, Washington D.C.

Verhoef, P., and P. Franses., 2003. "Combining Revealed and Stated Preferences to Forecast Customer Behavior: Three Case Studies." *International Journal of Market Research* 45(4): 467-474.

West, K.D., 1996. "Asymptotic Inference about Predictive Ability." *Econometrica* 64(5): 1067-1084.

Whitehead, J.C. 2005. "Environmental Risk and Averting Behavior: Predictive Validity of Revealed and Stated Preference Data." *Environmental and Resource Economics* 32(3): 301-316.

Whitehead, J.C., T.C. Haab, and J.C. Huang. 2000. "Measuring Recreation Benefits of Quality Improvements with Revealed and Stated Behavior Data." *Resource and Energy Economics* 22(4): 339-354.

Wilde, G.R., and R.B. Ditton. 1994. "A Management-Oriented Approach to Understanding Diversity Among Largemouth Bass Anglers." *North American Journal of Fisheries Management* 14: 34-40.

APPENDICES

APPENDIX A. INSTITUTIONAL REVIEW BOARD LETTER

Oklahoma State University Institutional Review Board

Date:	Wednesday, September 05, 2007
IRB Application No	AG0734
Proposal Title:	Oklahoma Lakes Survey 2007
Reviewed and Processed as:	Expedited

Status Recommended by Reviewer(s): Approved Protocol Expires: 9/4/2008

Principal
Investigator(s

Tracy Boyer
312 Ag Hall
Stillwater, OK 74078

The IRB application referenced above has been approved. It is the judgment of the reviewers that the rights and welfare of individuals who may be asked to participate in this study will be respected, and that the research will be conducted in a manner consistent with the IRB requirements as outlined in section 45 CFR 46.

☒ The final versions of any printed recruitment, consent and assent documents bearing the IRB approval stamp are attached to this letter. These are the versions that must be used during the study.

As Principal Investigator, it is your responsibility to do the following:

1. Conduct this study exactly as it has been approved. Any modifications to the research protocol must be submitted with the appropriate signatures for IRB approval.
2. Submit a request for continuation if the study extends beyond the approval period of one calendar year. This continuation must receive IRB review and approval before the research can continue.
3. Report any adverse events to the IRB Chair promptly. Adverse events are those which are unanticipated and impact the subjects during the course of this research; and
4. Notify the IRB office in writing when your research project is complete.

Please note that approved protocols are subject to monitoring by the IRB and that the IRB office has the authority to inspect research records associated with this protocol at any time. If you have questions about the IRB procedures or need any assistance from the Board, please contact Beth McTernan in 219 Cordell North (phone: 405-744-5700, beth.mcternan@okstate.edu).

Sincerely,

Sue C. Jacobs

Sue C. Jacobs, Chair
Institutional Review Board

APPENDIX B. FIRST COVER LETTER

Name and Address of addressee

September x, 2007

Dear

Would you do us a favor?

I am writing to ask you to help in a study of recreational lakes in Oklahoma. This study examines how lakes are used and what factors influence people's selection of lakes to visit.

We are contacting a random sample of residents from every county in the state to ask whether they visit lakes in Oklahoma, how often, and why.

Your participation will require several minutes to complete the enclosed questionnaire. Results from the survey will help Oklahoma agencies such as the Oklahoma Water Resources Board and Oklahoma State Parks manage and protect our lake resources. Even if you do not visit Oklahoma lakes, your response to the survey will help us understand why you have not visited the lakes and improve your satisfaction with them.

Your answers will remain completely confidential, and no individual's answers can be identified. Your information will be stored securely and will be available only to persons conducting the study. No reference will be made on written reports which could link you to the study. After this study is completed, your name will be deleted and never connected to your answer in any way. This survey is voluntary. There are no known risks associated with this survey which are greater than those ordinarily encountered in daily life. Your answers will help us very much to share your lake visiting experience. If for some reason you prefer not to respond, please let us know by returning the blank questionnaire in the enclosed stamped envelope.

If you have questions about your rights as a research volunteer, you may contact Dr. Sue

C. Jacobs, IRB Chair, 219 Cordell North, Stillwater, OK 74078, 405-744-1676 or irb@okstate.edu.

Thank you very much for helping with this important study.

Sincerely,

Tracy Boyer
Assistant Professor
Tracy.Boyer@okstate.edu

APPENDIX C. POSCARD REMINDER

In the last two weeks, a questionnaire seeking your opinion about Oklahoma Lakes was mailed to you.

If you have already completed and returned the questionnaire to us, please accept our sincere thanks. If not, please do so today. We are especially grateful for your help because it is only by asking people like you to share experiences that we can understand why people decide to visit or not visit lakes in state of Oklahoma. If you did not visit any lakes recently your response is still important and we'd appreciate answers to questions 1 and 14-25!

If you did not receive a questionnaire, or if it was misplaced, please call us at (405) 744-6169 or email us at Tracy.boyer@okstate.edu, and we will get another one in the mail to you.

Tracy Boyer
Assistant Professor

Department of Agricultural Economics
Oklahoma State University
Stillwater, OK 74078

APPENDIX D. SECOND COVER LETTER

Name and Address of addressee

September x, 2007

Dear X

I am writing a second time to ask you to help in a study of recreational lakes in Oklahoma. This study examines how lakes are used and what factors influence people's selection of lakes to visit. If you have already sent in your survey back, thank you, you do not need to complete another survey. If you perhaps lost your previous survey, we really would appreciate if you could take a few minutes of your time to do the survey this time.

We are contacting a random sample of residents from every county in the state to ask whether they visit lakes in Oklahoma, how often, and why. It is very important to us that we have as many responses as possible, because your opinion is important. Even if you do not visit Oklahoma lakes, your response to the survey will help us understand why you have not visited the lakes and improve your satisfaction with them. Results from the survey will help Oklahoma agencies such as the Oklahoma Water Resources Board and Oklahoma State Parks manage and protect our lake resources.

Again, we assure you that your answers will remain completely confidential, and no individual's answers can be identified. Your information will be stored securely and will be available only to persons conducting the study. No reference will be made on written reports which could link you to the study. After this study is completed, your name will be deleted and never connected to your answer in any way. This survey is voluntary. However, your answers will help us very much to share your lake visiting experience. If for some reason you prefer not to respond, please let us know by returning the blank questionnaire in the enclosed stamped envelope.

If you have any questions or comments about this study, feel free to contact Dr. Tracy Boyer by telephone or email at (405) 744-6169 or Tracy.boyer@okstate.edu, You may also write to us at the address on the letterhead.

Thank you very much for helping with this important study.

Sincerely,

Tracy Boyer
Assistant Professor

VITA

Phumsith Mahasuweerachai

Candidate for the Degree of

Doctor of Philosophy

Thesis: ESSAYS ON DEMAND FOR WATER-BASED RECREATION IN OKLAHOMA

Major Field: Agricultural Economics

Biographical:

Personal Data: Born in SiSaKet, Thailand, on February 8, 1976, Son of Sun and Chuleeporn Mahasuweerachai.

Education: Completed the requirements for the Doctor of Philosophy in Agricultural Economics at Oklahoma State University, Stillwater, Oklahoma in May, 2010.

Completed the requirements for the Master of Economics at National Institute of Development Administration, Bangkok, Thailand in October, 2003.

Completed the requirements for the Bachelor of Economics at Khon Kaen University, Khon Kaen, Thailand in October, 1998.

Experience: Oklahoma State University, Department of Agricultural Economics, Research Assistant, Teaching Assistant
Khon Kaen University, Department of Economics, Lecturer

Professional Memberships: Agricultural and Applied Economics Association

Name: Phumsith Mahasuweerachai Date of Degree: May, 2010

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: ESSAYS ON DEMAND FOR WATER-BASED RECREATION IN OKLAHOMA

Pages in Study: 120 Candidate for the Degree of Doctor of Philosophy

Major Field: Agricultural Economics

Scope, Method of Study, and Findings: Even though the demand for lake recreation in Oklahoma has increased continuously since the 1950s, few recent studies have analyzed the demand for lake recreation as well as their welfare effects from lake use in term of recreation. This study proposes to explore what factors influence lake recreation demand and how changes in these factors affect Oklahoma lake recreationists. Discrete choice analysis and travel cost techniques are applied in this study.

The first essay explores the benefit gained from combining revealed and stated preferences data (RP and SP data) to explain and predict current and future behavior of lake recreationists. Overall, this study found that the model that uses RP and SP data together provides the best explanation and prediction of lake recreationists' current and future behavior.

In addition, the second paper also develops a linked site choice model that combines current and potential lake recreationists' data to calculate the annual welfare changes due to lake quality changes. The idea of this essay is that new lake users attracted by improvements in lake quality should be included in estimating recreational demands because some current non-users could be attracted by higher lake quality. Failure to include this group of people results in biased annual welfare estimates. This study found that the model that pools current and potential lake recreationists' data provides significant larger annual welfare estimates than that which estimates from current lake users only. Furthermore, the combined current and potential lake user model can also capture the annual welfare changes from potential lake recreationists, which cannot be generated by the model that uses only current lake user data.

The third and final essay of this study focuses on anglers' preferences for the Close-to-Home-Fishing Program (CTHFP). Generally, anglers are willing to pay for increases in management effort such as larger stocked catfish and increased variety in fish stocked in the CTHFP ponds. However, relatively speaking, anglers are more willing to pay more for physical amenities such as fishing docks and restroom facilities.

ADVISER'S APPROVAL: Dr. Tracy A. Boyer