MEASURING THE VALUE AND ECONOMIC IMPACTS OF CHANGES IN WATER QUALITY AT GREAT LAKES BEACHES IN MICHIGAN

By

Li Cheng

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Agricultural, Food, and Resource Economics – Doctor of Philosophy

2016

ABSTRACT

MEASURING THE VALUE AND ECONOMIC IMPACTS OF CHANGES IN WATER QUALITY AT GREAT LAKES BEACHES IN MICHIGAN

By

Li Cheng

The objectives of this dissertation are to measure the monetary value of public Great Lakes beaches, then to measure the monetary value and economic impacts of water quality improvements to Great Lakes beaches. The first essay applied all trip data from a general population survey to Michigan adults to estimate the economic value of the public Great Lakes beaches. We found that on average a Michigan adult resident took 3.8 trips to the Great Lakes beaches in the summer of 2011. The seasonal value of access to a public Great Lakes beach ranged from $24.74 to $28.07 per person per trip, which would be reduced to two-thirds of the value if we only used single day trip data. To incorporate water quality attributes, Essay 2 combined trip data (RP) and choice experiment data (SP) to estimate the economic benefits from water quality changes at Great Lakes beaches in Michigan. We first applied a scaling approach to jointly estimate the parameters of attributes in both RP and SP datasets under a unified RUM framework. Different model specifications for common preferences across the data types were tested. The common preference test between the RP and SP data was consistently rejected. Our results provided empirical evidences that the scaling approach is not sufficient to account for differences in the amount of unexplained variance when using RP and SP data together in some applications. With some caveats, we then applied the calibration of SP to RP approach to measure the change in consumer surplus in response to two types of water quality scenarios. We found that water quality improvement impacts Huron south most, Michigan south least; water quality degradation impacts Lake Michigan most, Huron south

least. To measure the economic impacts of Great Lakes beaches, the third essay applied a visitor spending survey to estimate Michigan beachgoers' spending to Great Lakes beaches. An on-site recruitment of beachgoers was conducted at three public beaches in Michigan in 2014. Intercepted beachgoers were asked to take a web survey about their beach activities and their spending of the visits. A sample selection model was used to address potential nonresponse bias problem in the spending data. We found the regional spending of an average beachgoer to Great Lakes beaches ranged from $35.92 to $248.80 in 2014 dollars. Essay 4 integrated the recreation demand system from Essay 2 and spending analysis from Essay 3 to estimate regional variations in economic impacts from trips to Great Lakes beaches in Michigan. We found that the spending by all Michigan beachgoers living in the Lower Peninsula had a total economic impact of direct sales within a region that ranged from $425.87 million to $1,724.1 million per season in 2014 dollars.

.

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to my advisor, Dr. Frank Lupi for his extraordinary guidance and continuous support through my PhD study and dissertation work. His knowledge and experiences of empirical environmental economics impact my research very much. Besides research, I also learned dedication, critical-thinking, skills of communication and presentation from him. I am in debt to him for all his help with my research and my personal development.

Second, my sincere gratitude is given to Dr. Satish Joshi, Dr. Patricia Norris and Dr. Michael Kaplowitz for serving on my committee. Special thanks give to Dr. Joshi for his invaluable support and kind encouragement through my PhD study. I also want to thank Dr. Norris for her valuable comments to help me improve my dissertation. I would also like to thank Dr. Kaplowitz for enlightening me with his practical ideas and valuable comments in terms of policy implications.

Financial support for this dissertation was provided by NOAA Coastal Oceans Program, Michigan Sea Grant and Michigan Applied Public Policy Research Grants, Michigan State University. I am very grateful for all these financial support. In addition, I thank Min Chen, Scott Weicksel for their help with my dissertation; and Dr. Scott Swinton, Debbie Conway, for their administrative help.

Lastly and most importantly, I thank my parents and my sister in China, and my boyfriend Haitao Li. This dissertation is dedicated to them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

1. Motivation

As the largest body of freshwater lakes in the world, the Great Lakes are not only a valuable asset to economic development, but also provide ample recreational opportunities. With 10,210 miles of shoreline, the Great Lakes support a beach-related tourism economy. However, water quality issues have long been a public concern and could deter people from beach recreation in some areas. Some common water quality problems in the Great Lakes include algal blooms, aquatic invasive species (AIS), and bacterial contamination.

The re-emergence of problematic and toxic algal blooms is a severe issue affecting the Great Lakes, and Lake Erie in particular. Lake Erie was declared “dead” by the press in the late 1960s for being chocked with algae. According to NASA in 2011, Lake Erie has again undergone one of the worst algal blooms in decades[1]. In such an event, algal blooms often produce harmful toxins (e.g., *microcystis*) which can lead to beach closures and illness. A related algal problem in the Great Lakes happens when large mats of filamentous green algae (e.g., *clodophera*) break apart and form unsightly mats or even “muck” that fouls beaches (Natural Resources Defense Council, 2009; Verhougstraete et al., 2010).

The issue of aquatic invasive species also directly and indirectly affects beaches, and relates to algae problems. AIS directly affect beaches when they wash up on the shore. For instance, in the 1960’s dead alewives piled up in droves on beaches and or more recently zebra or quagga mussel shells have accumulated on some beaches (Alexander, 2011). AIS issues indirectly

[1] http://earthobservatory.nasa.gov/IOTD/view.php?id=76127&src=iotdrss

affect beaches through their correlation with algal problems. Some invasive species such as quagga and zebra mussels, and potentially in the future Asian carp, have spurred the growth of algae by filtering out plankton and increasing water clarity (NRDC, 2009).

Beach bacterial contamination and resulting beach closures remain a critical water quality issue in the Great Lakes region. In 2012, water quality samples from the Great Lakes region had the highest percentage exceeding EPA's *E. Coli* standards of any area in the Nation (NRDC, 2009). The number of beach closures in the Great Lakes, most of which are due to bacterial contamination, is a growing issue with over 3,000 closure and advisory days annually (Great Lakes Commission, 2009).

The on-going water quality problems may require intervention to protect the water quality of the Great Lakes from additional degradation. Without intervention by governments or other public-spirited organizations, the water quality problems of the Great Lakes will persist and may even become worse, because water quality is a public good which cannot be efficiently allocated by the market. In the case of beach recreation, water quality is same for all beachgoers at the same beach; even if one beachgoer values water quality much higher than the other, they still have to face the same level of water quality, so there is limited incentive for each individual action to provide protection.

In addition to the preceding water quality problems that may require water quality protection, there are also emerging economic incentives that may drive public and policy makers to improve the water quality. Traditionally, the Great Lakes have been used for municipal and industrial water supply, commercial fishing, and transportation, and although all these uses propelled the Michigan economy, some of them have the potential to degrade water quality.

Recently state and local governments are becoming increasingly interested in the "Growing Michigan's Blue Economy" Initiative (Austin & Steinman, 2015), which proposes to develop water-related industries in a clean, healthy, and sustainable way. In light of this possible transition, water quality improvement is crucial for the success of the initiative and the development of "blue" industries. In particular, as beach recreation has always played an important role in outdoor recreation, water quality improvements can directly benefit beach recreation and then contribute to local economy.

Accordingly, to prevent further degradation of water quality or to improve existing water quality of the Great Lakes will require resources. Because there are only limited funds for competing uses of many natural resources, information on the benefits of water quality protection or improvement are vital in policy makers' efforts to allocate funds and justify funding decisions. Furthermore, inaccurate estimates can undermine the credibility of water quality improvement programs and may cause their untimely failure (EPA, 1989), which emphasizes the need for quality information.

2. Challenges

Although decision makers have an increasing demand on the information, measuring water quality improvements in terms of economic benefits and economic impacts is still challenging. The first challenge lies in the complexity of identifying benefits from water quality improvements (Keeler et al, 2012). Because water quality improvements affect many aspects of human well-being, returns can accrue to recreational use, human health, and commercial use. Failing to consider all the returns will underestimate the benefits. However, as Bockstael, Hanemann and Kling (1987) indicated significant benefits from surface water quality improvements accrue to recreational use, yet little is known about these impacts in the Great Lakes. Thus, we consider recreational beach use, mainly because the Michigan Activity Survey (conducted by Lupi, Kaplowitz, Chen and Weicksel, 2011) found that visiting a beach is more popular than fishing or boating on the Great Lakes.

The second critical challenge lies in the complexity of defining water quality metrics. Water quality is sometimes measured on scales based on a combination of many chemical and biophysical variables in a small sample of water, but it is often difficult to describe overall water quality status in a large waterbody from a large number of variables (Griffiths et al, 2012). Besides, these chemical and biophysical measures may not be directly related to the water quality attributes that people actually perceive and value (Kneese 1968; Keeler et al, 2012). To address this challenge, we utilized water quality attributes that were described by their visual impact and were used in a choice experiment that was further combined with trip data to infer the recreation benefits of water quality improvements from observed behaviors and stated preferences. The water quality attributes were designed to be policy-relevant since they match those that EPA collects through its beach sanitation survey monitoring program (EPA 2008).

The third challenge lies in the lack of substitution effects in recreation demand from water quality changes in most economic impact studies. As Deisenroth, Loomis and Bond (2013) pointed out, most economic impact studies only provide a “snapshot” of an activity’s contribution at a given point in time. However, the economic impacts from water quality changes involve changes of economic demand. In particular, when water quality decreases, human behavior responds and people can choose to visit different sites or to forego visiting at all. Thus, quantifying economic impacts from water quality changes cannot simply rely on a “snapshots” of trips, because failure to account for substitution effects in recreational demand from water quality change results in overestimation of economic impacts (Deisenroth, Loomis & Bond, 2013).

3. Goal

In light of these challenges, the objectives of this dissertation are to measure the monetary value of public Great Lakes beaches, to measure the monetary value of water quality improvements to Great Lakes beaches, to estimate the trip expenditures of recreational beachgoers to Great Lakes beaches, and finally, to estimate the economic impacts of beach recreation and the economic impacts of water quality improvements by establishing the critical linkages between water quality and beach recreation.

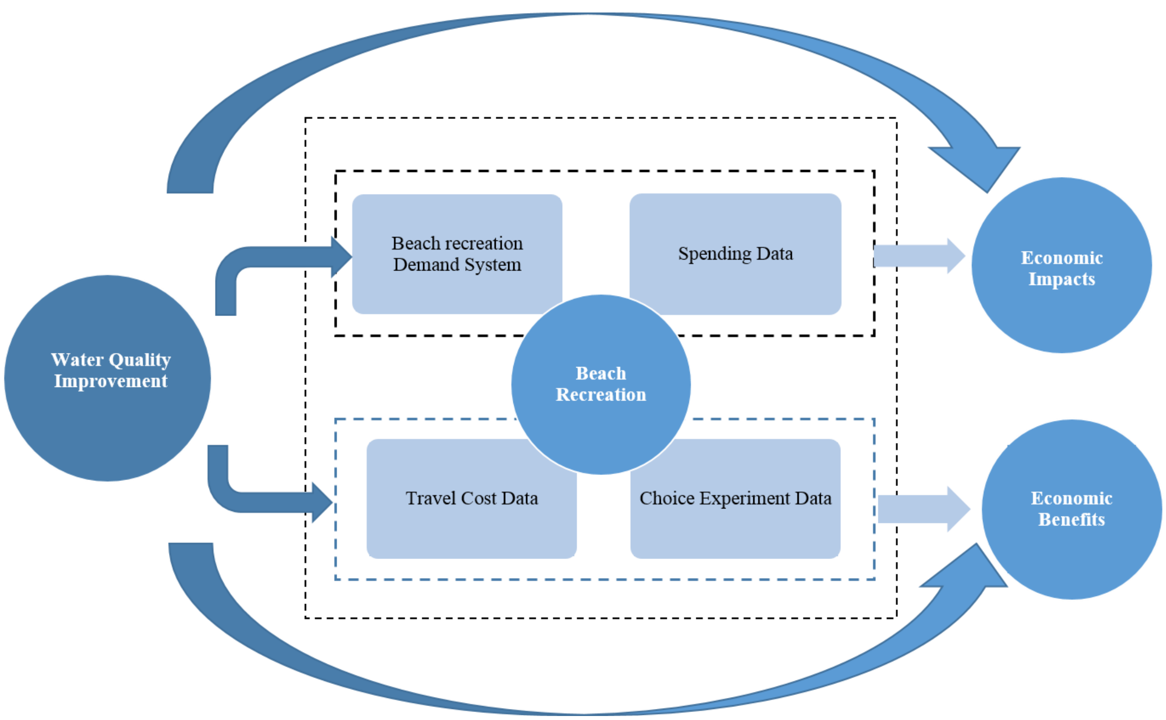


Figure 0.1 Integrated system to measure economic benefits and economic impacts of water quality improvements.

4. Thesis structure

This dissertation consists of four essays.

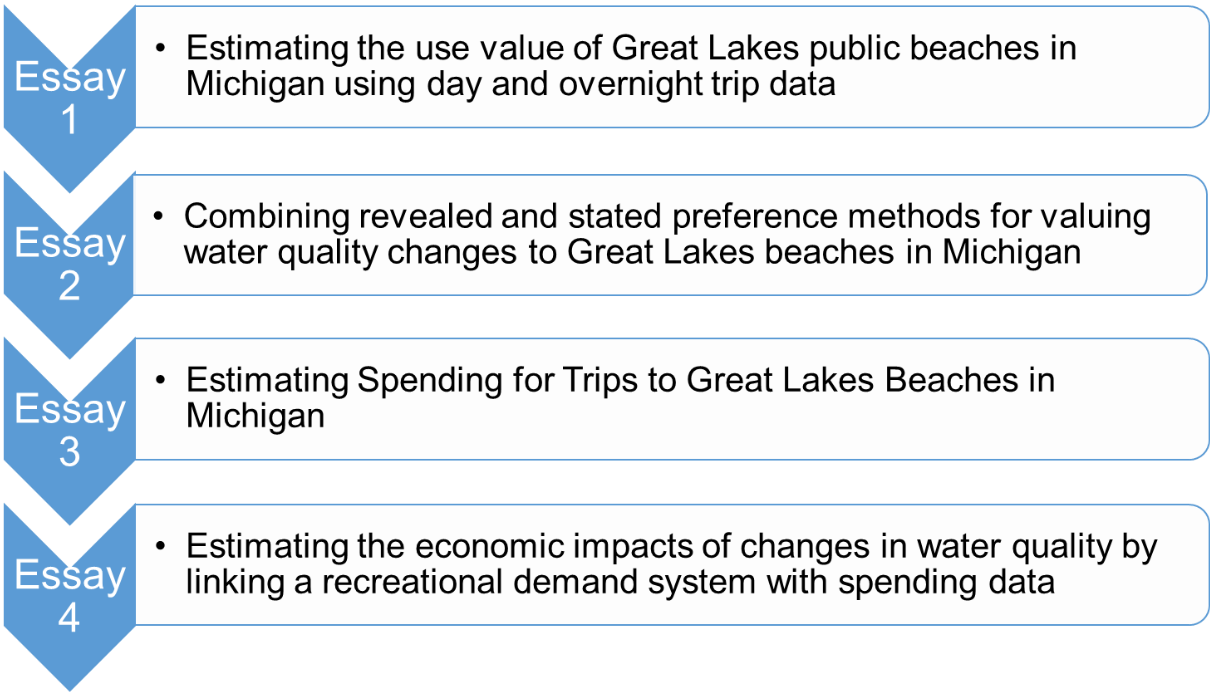


Figure 0.2 Thesis Structure

Essay 1: Estimating the Use Value of Great Lakes Public Beaches in Michigan Using Day and Overnight Trip Data

The first essay applied all trip data from a general population survey to Michigan adults to estimate the economic value of the public beaches on Lake Erie, Lake St. Clair, Lake Huron and Lake Michigan. The trip data was collected from a 2011 Great Lakes Beaches Survey conducted by Michigan State University. Although day trip data is the most widely used for valuing economic value of natural resources, in our trip data set, overnight trips account for a significant portion—around 20% of the total recreation trips. Using all trip data helps to derive the complete recreational demand curve, therefore the estimated results and welfare measures better reflect beachgoing. The economic estimates and welfare measures of this essay provide policy makers and beach managers with a better understanding of the factors determining beachgoers' site selection, as well as the economic benefits associated with access to beaches and for changes in the level of particular beach site attributes. The economic benefit information provided by this essay can also help policy makes with decision making on beach restoration and protection programs.

Essay 2: Combining revealed and stated preference methods for valuing water quality changes to Great Lakes beaches in Michigan

The second essay aims to estimate the economic benefits from water quality improvements at Great Lakes beaches in Michigan. Based on the repeated random utility model (RUM) of recreational demand from Essay 1, Essay 2 explored the possibility of incorporating additional water quality attributes by using joint estimation of revealed preference data and stated preference data. To combine the trip data and choice experiment data from the 2011 Great Lakes Beaches Survey, we first applied a scaling approach to jointly estimate the parameters of attributes in both RP and SP datasets under a unified RUM framework. Different model specifications for common preferences across the data types were tested. The common preference test between the RP and SP data was consistently rejected. Our results provided empirical evidences that the scaling approach is not sufficient to account for differences in the amount of unexplained variance when using RP and SP data together in some applications. With some caveats, we then applied the calibration of SP to RP approach to measure the change in consumer surplus in response to two types of water quality scenarios. The economic benefits of improvement or protection of water quality provided by essay 2 can be used in cost-benefit analyses of water quality program evaluation, damage assessment, and policy making.

Essay3: Estimating spending for trips to Great Lakes Beaches in Michigan

Spending analysis is an essential part of economic impact analysis. Although many coastal states have conducted spending studies for saltwater beaches, few spending studies have addressed Great Lakes beaches. The third essay contributes to this area of study by using a visitor spending survey to estimate Michigan beachgoers' spending to Great Lakes beaches. An on-site recruitment of beachgoers was conducted at three public beaches in Michigan in 2014. Intercepted beachgoers were asked to take a web survey about their beach activities and their spending of the visits. The purpose of essay 3 is to quantify the amount of local spending attributed to beach recreation. Unlike most literature, a sample selection model is used to address potential nonresponse bias problem in the spending data, so that the subsequent estimation of visitors' spending would be more accurate. We further used the estimated spending equation to predict an average beachgoer's spending per trip by using the 2011 Great Lakes Beaches Survey. Finally, we used the predicted trips based on the demand system from Essay 1 to obtain the regional variation of spending from recreation trips to Great Lake beaches.

Essay 4: Estimating the economic impacts of changes in water quality by linking a recreational demand system with spending data

To date, few economic impact studies have linked economic impacts to trip demand functions from a formal recreation demand model, let alone to models for predicting recreational demand from water quality changes. By integrating the recreation demand system from Essay 2 and spending analysis from Essay 3, Essay 4 aims to estimate regional variation in economic impacts from trips to Great Lakes beaches in Michigan. By constructing two types of water quality scenarios, this essay further estimated the changes in economic impacts to the local region when water quality changes. By quantifying the contribution of beaches to the local economy, the results of Essay 4 can help the policy makers and the public to know some of the economic importance of preserving and restoring beaches. Essay 4 also quantified the contribution of water quality improvement to the local economy; therefore, the results can also be useful to water quality restoration and protection programs.

ESSAY 1 Estimating the Use Value of Great Lakes Public Beaches in Michigan Using Day and Overnight Trip Data

1. Motivation

With the longest freshwater coastline in the country, Michigan has abundant public beaches along the Great Lakes' shoreline, which are valuable recreational assets. The 2011 Michigan Activity Study conducted by Michigan State University shows that 58% of Michigan adults visited a beach on the shoreline of the Great Lakes during the summers of 2010 and 2011 (Chen, 2013), which suggests that in those two years about 4 million visitors from all over the state went to beaches along the Great Lakes. Obviously, there is economic value raised from the recreational use of public beaches. Despite the fact that no explicit market price exists, we can still use observed trip behavior to indirectly infer economic values for beach recreation use by using recreation demand models. Accurate measures of use values of public beaches are important for policy makers as they need the economic benefit information to help their decision making in regards to program evaluation, damage assessment, policy making, and environmental legislation.

In spite of the importance of the Great Lakes shoreline, there is very limited information from prior studies on the value of these freshwater beaches. In contrast, the economic value of ocean beaches has been investigated by many researchers. For instance, Deacon and Kolstad (2000) identified 13 relevant studies of the economic value of saltwater beach recreation from the years 1972 to 1984. They reviewed the value of a saltwater beach-day, independent of the high season, to be in the range of \$1.2-\$22.3 in 2011 dollars. King (2002) used a travel cost model to estimate the use value of a beach for San Clemente, CA, reporting values of \$38.82 in the high season and \$3.81 in the low season in 2011 dollars. Lew and Larson (2005) used a RUM recreation demand model to estimate the use value of beaches in San Diego County to be \$36.93 per day in

2011 dollars. Parsons, Massey, and Tomasi (1999) estimated the value of a beach closure to range from \$0.00- \$22.75in 2011 dollars per person per trip across six sites in the Mid-Atlantic region. Bell and Leeworthy (1990) conducted an on-site survey in Florida and estimated a beach day in Florida to be \$76.48 in 2011 dollars for long distance travelers. Pendleton, Kildow and Rote (2006) summarized the use value of beaches in California using meta-analysis, the results of which ranged from \$13.94 to \$80.62 in 2011 dollars. Whitehead et al. (2008) conducted a phone survey of 419 respondents for 17 beaches along the North Carolina coastline. They pooled revealed preference data and stated preference data using the single-site travel cost model and reported the economic values of a beach day from \$104.11 to \$117.22 in 2011 dollars.

The Great Lakes are unique because they are the largest group of freshwater lakes on Earth. Although much research has been done with ocean beaches, few studies have covered public Great Lakes beaches. Murray, Sohngen and Pendleton (2000, 2001) surveyed 1,587 visitors at 15 Lake Erie beaches on site in 1998. Their result for the economic value of beach advisories suggested that removing one advisory at all beaches is \$2.55 in 2011 dollars per person per trip. However, their study was applied to only 15 beaches on Lake Erie, and may not be representative of beaches on the other areas of the Great Lakes. Using single site demand models, they also estimated the beach use value for Maumee Bay to be \$35.33 per trip and \$21.39 per trip for Headlands in 2011 dollars. Lupi, Kaplowitz, Chen and Weicksel conducted a web survey that covered the general population of Michigan beachgoers (Chen 2013, Weicksel 2012). With detailed single-day trip information to 451 Great Lakes beaches in Michigan, Chen (2013)[2] estimated the value of day trips to public Great Lakes beaches to be somewhere between \$14.25 to \$17.24 in 2011 dollars.

[2] An algebra mistake in the nested logit model of Chen (2013) has been corrected here.

In augmenting Chen's study (2013), this paper aims to estimate the value of a beach day by not only using day trip data but also overnight trip data. Commonly, the most widely used data set for valuing economic value of natural resources is day trip data (Lew and Larson, 2005; Parsons, Massey, and Tomasi, 1999; Murray, Sohngen and Pendleton, 2001). Information on overnight trips is usually excluded, mainly because the majority of trips are day trips, the primary purpose of a day trip usually is recreation, and multiple day trips are often from longer distances. Thus, by excluding multiple day trips, one can substantially reduce the number of choice alternatives, i.e., substitute sites that must appear in each person's choice set, which can dramatically reduce data and computational burdens. However, there are some studies that make use of all trip information to value natural resources (e.g., Hausman, Leonard and McFadden, 1995).

In our Great Lakes beaches data set, although most recreation trips are day trips, overnight trips account for around 20% of the total recreation trips, still a significant portion. Therefore, demand for recreational beach use will be more accurately modeled if all trips are accounted for. The demand curve derived from using only day trip data puts all the weight on the "low priced" trips (i.e. trips with relatively short distances travelled), while omitting the overnight trip data loses some of the "high priced" trips (i.e. trips with longer distances travelled). Since the economic value is derived from the recreation demand, estimating the complete demand curve covering both low priced and high priced trips makes the economic value estimation more accurate.

The remainder of the paper is organized as follows. Section 2 first reviews the theoretical framework, provides some necessary assumptions, and presents the repeated three-level nested logit model. Section 3 describes the Great Lakes Beaches survey, the data set, and the empirical specification of the model. Estimation results and welfare calculations are then presented in Section 4 followed by the conclusions in Section 5.

2. Model

2.1 Repeated Three-level Nested Logit Model

The random utility maximization (RUM) theory (McFadden, 1974) is one of the most popular in recreation demand studies. Within the RUM framework, each recreationist chooses among a set of mutually exclusive sites to visit. It is assumed that the utility the recreationist obtained from his choice is deterministic to the individual but random to the researcher because the researcher does not observe all the factors that influence the individual's choice. The RUM considers a recreationist's utility to be a function of the attributes of the sites. The recreationist's choice implicitly reveals the trade-off between site attributes. If we include travel cost into the site attributes, we can get the implicit value of site attributes in dollar terms.

While the site choice RUM is widely applied when there are many substitute recreation sites available, it does not directly explain the total seasonal number of trips, which is often referred to as seasonal participation. However, by including a "don't go" option in choice sets, the RUM model is easily expanded to allow for repeated choices by a recreationist and in turn can explain total trips per season. The repeated RUM model has been widely used in the recreation demand literature because it combines the recreational site selection and participation decision in a unified framework, which is utility theoretic consistent for welfare analysis (Freeman III, Herriges and Kling, 2014).

In this study, a repeated RUM model is specified as a three-level nested logit model. On a given choice occasion *t*, a Michigan beachgoer *n* has the choice of whether to take a trip or not, which lake to choose, and where to go for the beach. The set of sites that are available to the beachgoer is denoted as the choice set *C*. The decision process can be visualized as choosing among the *M* nests, *M*= {*Trip, No trip*}, among the *L* lakes in the nest *Trip*, *L* = {*Lake Erie, Lake*

St. Clair, Lake Huron, Lake Michigan}, and among the *J* beaches at one of the lakes *l*. The decision tree is illustrated in the figure below:

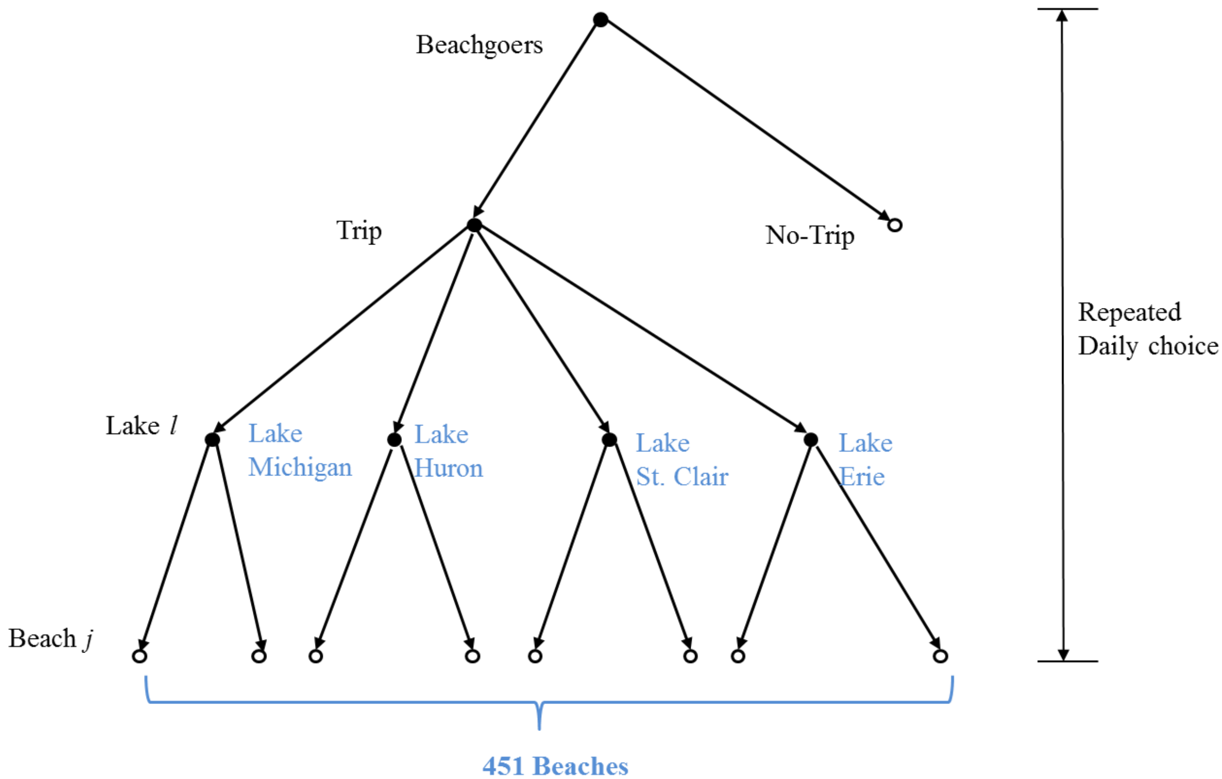


Figure 1-1 Repeated three level decision tree of beach recreation trip.

In the paper, we applied a repeated RUM model to the Great Lakes Beaches trip data, which explains the site choice and recreation demand of trips to Great Lakes beaches in a summer season. Trips are distinguished by Great Lake and beach location. Following Chen (2013) and Morey et al. (1993), we assume the summer season consists of a fixed number of choice occasions (T), during which each beachgoer is assumed to make at most one trip. On every choice occasion, a beachgoer simultaneously decides whether or not and where to go to a beach. In this application, the beach season is defined as the period from Memorial Day weekend to September 30, 2011; with the choice occasion length at 1 day, there are 126 choice occasions in the season. Following the literature (Morey et al 1993), for any sampled respondents, if the sum of their reported day trips and overnight trips in any month exceeds the total number of days in that month, the number of trips is trimmed to the total number of days in that month, i.e., 34 in June, 31 in July, 31 in August, 30 in September[3].

During the indicated summer season, 19,284 trips were taken by the survey respondents, among which 20.92% were overnight trips. To make the best use of the available data, we include both day trip and overnight trip data. We follow Bockstael, Strand and Hanemann (1987) and assume that time spent on the beach is endogenous and therefore not included into the cost of the visit. According to McConnell (1992), when people choose the time to spend on a site as a part of their recreation decisions, on-site time can be ignored in the demand specification and will not bias the demand estimation and welfare analysis of trips, provided the choice of on-site time is endogenous.

We are also aware that beachgoers might have multiple objectives for overnight trips. For example, in addition to visiting a beach, beachgoers may engage in other activities like visiting family or friends, going to state parks, etc. In our survey, we explicitly asked the respondents whether their main purpose of the trip to the first reported beach is for recreation or not. It turns out that 91.08% of respondents reported that their main purpose for a short overnight trip—an overnight trip of less than 4 nights—was recreation. For a long overnight trip, i.e., overnight trip of 4 nights or more, 92.42% of respondents' main purpose was recreation. In order to reflect the fact that not all economic value accrues from beach recreation if there are multiple objectives involved (Yeh, Haab and Sohngen, 2006), we used the above mentioned two percentages as the corresponding weights to adjust the short overnight trips and long overnight trips downward.

---

[3] Less than 0.3% of the observations were trimmed due or exceeding the number of choice occasions in that month. (See Appendix A)

More formally, following Chen (2013) and Morey et al. (1993), the utility that beachgoer *n* derives from choosing alternative *j* from the set *C* is given as (individual subscript *n*, choice occasion *t* is omitted for now to simplify the notation):

$$U_{jlm} = V_{jlm} + \varepsilon_{jlm}, \;\; \forall \, (jlm) \in C$$

The systematic component,V_{jlm}, is observable to researchers and usually it is a function of the attributes of site *j* and the individual's socio-demographic characteristics, while the random term ε_{jlm} captures all the factors unobservable to researchers.

Individuals choose the alternative which generates the highest utility, so the probability that a beachgoer chooses site *j* is a cumulative distribution that depends on the density $f(\varepsilon_{jlm})$. Assume that the joint density function of the random term is given by the first type of generalized extreme value (GEV) distribution for a three-level nest (McFadden, 1978):

$$F(\varepsilon_{jlm}) = exp\left\{-\sum_{m\in M}\left[\sum_{l\in L_m}\left[\sum_{j\in J_{lm}} exp\left(-\frac{\varepsilon_{jlm}}{\lambda}\right)\right]^{\frac{\lambda}{\rho}}\right]^{\rho}\right\}$$

Where

- Beach sites *J*= {1, 2, …, 451};
- Lake alternatives *L*= {*Lake Erie, Lake St. Clair, Lake Huron, Lake Michigan}*;
- Trip alternatives *M*= {*G, No*}; (*G* is short for *Trip*, *No* is short for *No Trip*)

The probability of beach *j* being chosen is given by

$$P_{jlG} = P_{(j|lG)} * P_{(l|G)} * P_G$$

Where $P_{(j|lG)}$ is the conditional probability of choosing a beach *j* given that lake *l* and trip alternative *G* is chosen. $P_{(l|G)}$ is the conditional probability of choosing a lake *l* given a trip alternative *G* is made. P_G is the marginal probability of taking a trip. Then denote the indirect utility of not taking a trip as V_{No}.

The conditional and marginal probabilities are given by:

$$P_G = \frac{\exp(\rho IV_G)}{\exp(\rho IV_G) + \exp(V_{No})}$$

$$P_{(l|G)} = \frac{\exp\left(\frac{\lambda}{\rho} IV_{lG}\right)}{\sum_{k \in L_G}\left[\exp\left(\frac{\lambda}{\rho} IV_{kG}\right)\right]}$$

$$P_{(j|lG)} = \frac{\exp\left(\frac{1}{\lambda} V_{jlG}\right)}{\sum_{i \in J_{lG}}\left[\exp\left(\frac{1}{\lambda} V_{ilG}\right)\right]}$$

The expected utility that each beachgoer receives from the choice of alternatives within each nest is called an inclusive value. IV_G and IV_{lG} are the inclusive values of Trip nest G and sub-Lake nest respectively, where

$$IV_G = \ln\left[\sum_{k \in L_G}\left[\exp\left(\frac{\lambda}{\rho} IV_{kG}\right)\right]\right]$$

$$IV_{lG} = ln\left[\sum_{i \in J_{lG}}\left[\exp\left(\frac{1}{\lambda} V_{ilG}\right)\right]\right]$$

Finally, the unconditional probability of taking a trip to beach *j* is:

$$P_{jlG} = \frac{\exp(\left(\frac{1}{\lambda} V_{ilG}\right) * \left[\sum_{l \in L_m} \left[\sum_{j \in J_{lm}} exp\left(\frac{1}{\lambda} V_{jlG}\right)\right]^{\frac{\lambda}{\rho}}\right]^{\rho-1} * \left[\sum_{j \in J_{lm}} exp\left(\frac{1}{\lambda} V_{jlG}\right)\right]^{\frac{\lambda}{\rho}-1}}{\left[\sum_{k \in L_G} \left[\sum_{i \in J_{kG}} exp\left(\frac{1}{\lambda} V_{ikG}\right)\right]^{\frac{\lambda}{\rho}}\right]^{\rho} + \exp(V_{No})}$$

The unconditional probability of *not* taking a trip to any beach is:

$$P_{No} = \frac{\exp(V_N)}{\left[\sum_{k \in L_G} \left[\sum_{i \in J_{kG}} exp\left(\frac{1}{\lambda} V_{ikG}\right)\right]^{\frac{\lambda}{\rho}}\right]^{\rho} + \exp(V_{No})}$$

Then the expected maximum utility for each choice occasion, or the inclusive value of each individual *n*, can be obtained as:

$$IV = ln\left\{\left[\sum_{k \in L_G} \left[\sum_{i \in J_{kG}} exp\left(\frac{1}{\lambda} V_{ikG}\right)\right]^{\frac{\lambda}{\rho}}\right]^{\rho} + \exp(V_{No})\right\}$$

Let *T* denote the total number of choice occasions, called the beach season. Let $\mathrm{y_{nt}} = 1$, if person *n* visited beach on occasion *t*, and $\mathrm{y_{nt}} = 0$, otherwise. To simplify the notation for probability expressions, individual *n* at time *t* will be noted after the comma in the subscript of the probability.

The log-likelihood function for this sample is:

$$LL_{beach}^{RP} = \sum_{n=1}^{N} \sum_{t=1}^{T} \left[\sum_{l \in L_G} \sum_{j \in J_{kG}} w_n * y_{nt} * \ln\left(P_{jlG,nt}\right) + w_n * (1 - y_{nt}) * \ln(P_{No,nt})\right]$$

where w_n is the weight of person n, which consists of 3 components(Appendix A). The first one is the sample weight, aiming to correct for sampling strata and possible non-representativeness of the sample (see Chen, 2013, Appendix C). The second one is the downward weight to correct for multiple purposes for overnight trips, which is 91.08% for short overnight trips and 92.42% for long overnight trips, respectively. The final one is the weight used for correcting for adjusted trip counts. In our web survey, after respondents finished their trip log section, we summarized the number of each type of trip they reported into a table, then verified whether the numbers in the table sound correct to them or not. Only 3.59% of the total samples reported “No” in this trip verification question. For the person who reported “No”, they were given a new table and asked to correct the number of trips they took. For each type of trip, less than 1% of sample changed their number of trips. We used the ratio of the first reported number of trips to the changed number as the weight to correct for the trip adjustments. For instance, if a person first reported 20 for the total number of day trips, then changed to 10 after the verification question, we apply 10/20=0.5 to weight the monthly trip number of day trips. Similarly, we used the same method to correct for the downward adjustments. The final weight w_n is the product of the 3 components.

Another more complicated part of the sample is called the “incomplete” sample (Morey, 1993), as there is only partial information on the alternatives chosen. In the trip data, some people only reported the nearest town or city to the beach, so we do not know the exact beach but only an aggregated area for their visit. For trips with partial information, Chen (2013) grouped the 451 beaches into 80 groups based on the characteristics and distance of the reported beach to the nearest town or city. We applied the same approach to handle trips with partial information. Denote the grouped area as $\boldsymbol{a}$, then the log-likelihood function for this “incomplete” sample is:

$$LL_{group}^{RP} = \sum_{n=1}^{N}\sum_{t=1}^{T}\left[\sum_{l\in L_G}\sum_{j\in \boldsymbol{a}} w_n * y_{nt} * \ln\left(P_{jlG,nt}\right) + w_n * (1 - y_{nt}) * \ln(P_{No,nt})\right]$$

That is, it is the sum of the probabilities of visiting the individual sites within area ***a***. To illustrate what happens with the grouped areas, note that the probability of rolling a one or two on a six-sided die is simply the probability of rolling a one plus the probability of rolling a two since the events are independent. Thus, the probability of visiting a site in area ***a*** is the sum of the site probabilities in the area. Finally, we have some reported beaches which were unknown to researchers, as the way they were reported does not allow researchers to either locate the exact beach or aggregate the beaches into groups. However, we do know if a respondent has taken a trip, so the unconditional probability P_G was applied to the unknown-beach samples.

$$LL_{unknown}^{RP} = \sum_{n=1}^{N}\sum_{t=1}^{T}\left[w_n * y_{nt} * \ln\left(P_{G,nt}\right) + w_n * (1 - y_{nt}) * \ln(P_{No,nt})\right]$$

The log-likelihood function for all the samples in the trip data is:

$$LL^{RP} = LL_{beach}^{RP} + LL_{group}^{RP} + LL_{unknown}^{RP}$$

2.2 Predicted Trips

Once we get the estimated parameters from maximizing the log-likelihood function, we can predict individuals' unconditional probabilities of taking trips to a specific beach and probabilities of taking trips. Specifically, for individual n, in the given beach season, the predicted total number of trips is:

$$\hat{Y}_{G,n} = \sum_{t=1}^{T}\hat{P}_{G,nt}$$

The predicted total number of trips to beach *j* at Lake *l* in the beach season is:

$$\hat{Y}_{jlG,n} = \sum_{t=1}^{T} \hat{P}_{jlG,nt}$$

If a beach closed or the water quality attributes changed, the change in predicted total number of trips is:

$$\Delta\hat{Y}_{G,n} = \sum_{t=1}^{T} \hat{P}_{G,nt}(scenario) - \sum_{t=1}^{T} \hat{P}_{G,nt}(status\ quo)$$

Similarly, the change in predicted total number of trips to beach *j* at Lake *l* is:

$$\Delta\hat{Y}_{jlG,n} = \sum_{t=1}^{T} \hat{P}_{jlG,nt}\,(scenario) - \sum_{t=1}^{T} \hat{P}_{jlG,nt}\,(status\ quo)$$

2.3 Welfare Measures

For valuation, one needs to measure the change in consumer surplus in response to a particular policy. According to McFadden (1973) and Small and Rosen (1981), the welfare change can be calculate as the change of expected maximum utility, i.e. the change of inclusive value, divided by the marginal utility of income.

$$cs_{nt} = \frac{\widehat{IV}_G(senario) - \widehat{IV}_G(status\ quo)}{-\hat{\beta}_{tc}}$$

For individual *n*, the seasonal welfare change will be the sum of all consumer surplus changes in each choice occasion *t:*

$$CS_n = \sum_{t=1}^{T} cs_{nt}$$

The weighted average seasonal value per person is:

$$\overline{CS} = \frac{\sum_{n=1}^{N} w_n * CS_n}{\sum_{n=1}^{N} w_n}$$

It is sometimes convenient for comparison to other literature to normalize these seasonal value to the change in trips. There are two ways to normalize the weighted average seasonal value per person by per trip. One is to divide the value by the weighted average total change in trips to all sites (i.e., total trip changes)

$$\overline{\overline{CS}}_G = \frac{\overline{CS}}{\Delta\bar{Y}_G} = \frac{(\sum_{n=1}^{N} w_n * CS_n)/(\sum_{n=1}^{N} w_n)}{\sum_{n=1}^{N} w_n * \Delta\hat{Y}_{G,n}/(\sum_{n=1}^{N} w_n)}$$

Another is to divide the value by the weighted average change in trips to beach *j* on lake *l*.

$$\overline{\overline{CS}}_{jlG} = \frac{\overline{CS}}{\Delta\bar{Y}_{jlG}} = \frac{\sum_{n=1}^{N} w_n * CS_n\,/(\sum_{n=1}^{N} w_n)}{\sum_{n=1}^{N} w_n * \Delta\hat{Y}_{jlG,n}/(\sum_{n=1}^{N} w_n)}$$

3. Survey and Data

3.1 Survey

The data comes from the Great Lakes Beaches Survey[4], which was conducted by Lupi, Kaplowitz, Chen and Weicksel in 2011 and 2012. The Great Lakes Beaches Survey was a statewide general population survey, the procedure consisted of two stages: a short screener survey, and then a web survey. First, in order to identify beachgoers, the screener survey was mailed to 32,230 Michigan adults who were randomly drawn from a Michigan driver's license list. To reduce potential self-selection bias, the screener survey covered a broad range of indoor and outdoor

[4] See Min Chen (2013), Scott Weicksel (2012) for additional details regarding the survey sampling and implementation.

leisure activities, among which there was only one screening question for Great Lakes beach recreation. Respondents who answered that they had visited a Great Lakes beach during two summers in 2010 and 2011were invited to take a follow-up web survey.

The web survey asked respondents for detailed monthly trip information on three types of trips from Memorial Day weekend to September 30, 2011: day trips (lasting a day or less), short overnight trips (less than four nights), and long overnight trips (four nights or more). In addition to trip information, respondents were asked for more detailed questions on up to two randomly selected trips, such as date, main purpose of the trip, etc. Specifically, for the short overnight trip, if respondents went to more than one beach on a trip, they were asked only to report the beach where they spent the most time on the trip. For the long overnight trips, respondents were asked to report the beaches on which they spent the most/second most/third most amount of time. We use the beach where they spent the most time as the destination for this paper.

3.2 Data

In the mail survey dataset with 9,591 observations, 3,838 indicated they did not visit any Great Lakes beaches in 2010 or 2011, so they are defined as “nonusers” for beach recreation. The 5,737 respondents that indicated they had visited a Great lakes beach were invited to the web survey. There were 3,196 people who responded to the web survey resulting in a response rate for the web survey of about 59%. In the demographic section of the survey, respondents were asked if they were the person to whom the web survey was addressed or if they were another household member or “someone else”. To maintain consistent demographic information, we only kept the respondents to whom the web survey was addressed, which left us 2,537 effective respondents from the web survey.

Following the definition in Shonkwiler and Shaw (1996), we define the respondents of the web survey who took at least one trip to Great Lakes beaches from Memorial Day weekend to September 30, 2011 as “users”, and those who had taken trips to Great Lakes beaches before but did not take any trip during the indicated season as “potential users”. Including the nonusers from the screener survey and users and potential users from the web survey, the effective sample size is 6,375. Among the 6,375 observations, there were 3,838 nonusers who had not taken any trips to Great Lakes beaches before, 1,894 users who took at least one trip, and 643 potential users with no trip during the indicated beach season. Specifically, Table 1-1 shows the number of users and potential users for three types of trips. Day trips were the most common with 1538 users. 607 users took at least one short overnight trip, and 543 users took at least one long overnight trip. Fewer users took overnight trips.

Table 1-1 The number of users and potential users for different types of trips

	Day trip	Short overnight trip	Long overnight trip	All trips
Users	1538	607	543	1894
Potential users	999	1920	1994	643
Total	2537	2537	2537	2537

Table 1-2 provides descriptive statistics for the sample of 6,375 respondents used in this paper. Demographic information of the users and potential users is collected from the web survey as it is the most recent, though we also collect demographic information in the screener mail survey, which is used for missing data imputation (see Appendix B for missing data imputation). Demographic information of the nonusers can only be collected from the screener survey. The characteristics of users which are most different from those of the potential users and nonusers are income, employment status, race and gender. We would expect that beachgoers are more likely to take a trip if they are employed full-time, not retired, white males with higher income.

Table 1-2 Demographic characteristics of effective samples [5]

	All	Users	Potential Users	Nonusers
Sample size	6375	1894	643	3838
Age (Mean)	46.5	44.2	44.96	49.5
Income (Mean, $1000)	73.1	82.6	80.1	60.9
Education Years (Mean)	14.4	14.9	14.6	13.8
Male (%)	48.6	48.6	45.6	49.7
White (%)	86.3	91.9	88.4	80.1
Employed Full-Time (%)	47.1	53.9	47.9	40.1
Retired (%)	23.7	18.1	22.1	29.9
Children under 17 (%)	32.5	35.4	33.8	29.2

The choice set is composed of reported beaches on Lake Erie, Lake St. Clair, Lake Huron and Lake Michigan. In other words, the choice set does *not* include the following beaches: reported beaches that are on Lake Superior/inland lakes, reported beaches that are out of Michigan, and three beaches that do not have length information. After matching the reported beaches to the Michigan DEQ beach database, the choice set for each individual is comprised of 451 beaches (Figure 1-2).

[5] Demographic statistics are weighted by sample weights.

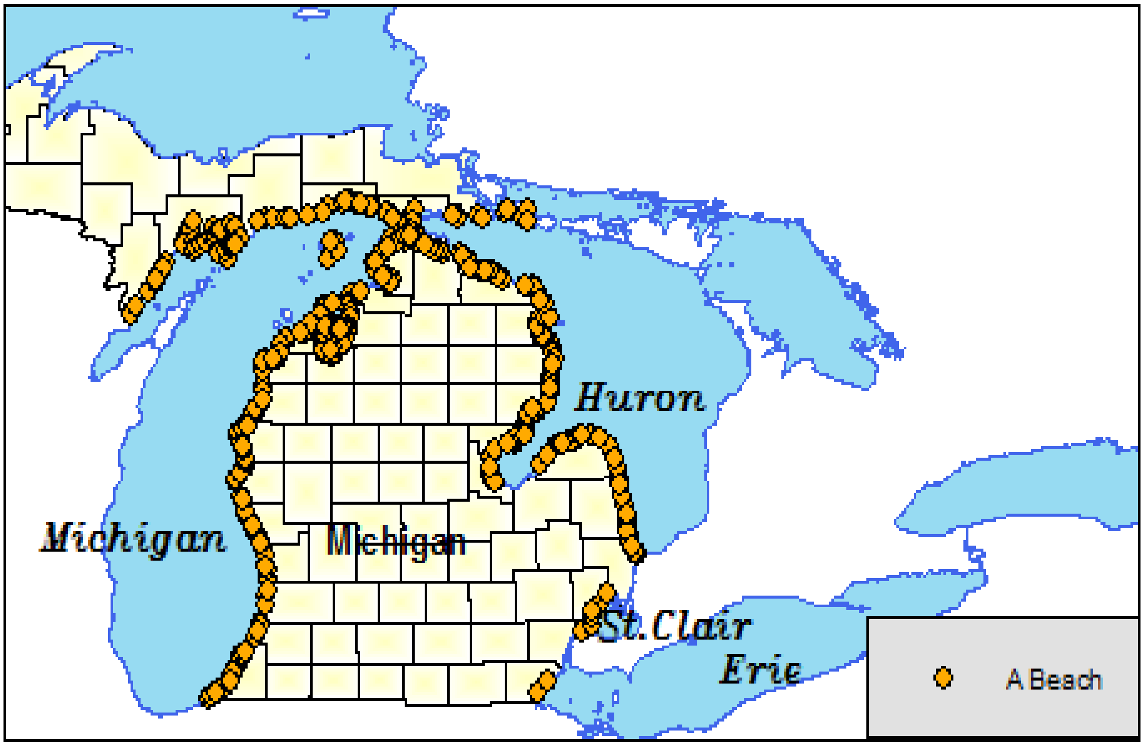


Figure 1-2 The 451 public great lakes beaches in the choice set

The trip data consists of self-reported trips to Great Lakes beaches, including day trips, short overnight trips, and long overnight trips. Although the majority of trips are day trips, the short overnight and the long overnight trips are 13.38% and 7.54% of the total trips, respectively (Table 1-3). Added together, the total number of overnight trips is 4,035, taking 20.92% of the total trips.

Table 1-3 The number of trips for three types of beaches

	Matched beaches	Grouped beaches	Unknown beaches	All beaches
Day trips	8519	5382	1348	15249
Short overnight trips	699	1482	400	2581
Long overnight trips	154	246	1054	1454
Total trips	9372	7110	2802	19284

As beachgoers reported their trip log information, some beachgoers reported the beach name, some of them reported the name of the nearest town or city to the beach, and some others skipped those two questions or the way they reported does not allow us to locate their trips. Therefore, there are three types of beaches in our choice set. The first one is called "matched beaches", namely we can match their reported beach names to the Michigan DEQ database. The second type is called "grouped beaches", as we do not know the exact beach but the area, the beaches are then grouped or aggregated into an area. The third type is named as "unknown beaches" because we do not know the beach or even the area but only know that they had taken a trip or not in the beach season. Trips from the "grouped beaches" and "unknown beaches" take 51.4% of the total trips, with each taking 36.87% and 14.53% respectively.

3.3 Econometric Model Specification

Following Chen (2013), in occasion *t*, the indirect utility for individual *n* obtained from visiting beach *j* at Lake *l* is:

$$V_{jlt} = \beta_{tc} * travel\ cost_{jl} + \beta_l * \log\left(beach\ length_{jl}\right) + \omega_t * temperature_{jlt} + \omega_{cd} * closure\ days\ of\ 2010_{jl} + \omega_r * regional\ dummies_{jl}$$

In particular, the computation of travel cost also follows Chen (2013) as:

$$Travel\ cost = round\ trip\ travel\ distance * \$0.2422\ per\ mile + round\ trip\ travel\ time * (annual\ income/2{,}000) * (1/3)$$

where travel cost is the sum of driving cost and time cost. Round trip travel distance and round trip travel time are calculated in PC Miller software. The cost of driving calculated is $0.2422 per mile, based on data from the 2011 AAA report. The opportunity cost of an hour is approximated using one third of the hourly wage, which is annual income divided by 2,000.

Regarding the regional dummies, as shown in Figure 1-3, we divided Michigan into 7 regions *R*, and *R={UP Peninsula, LP Northeast, LP Mid-East, LP Southeast, LP Northwest, LP Mid-West, LP Southwest}*. We set the *UP Peninsula* as the baseline, then the other 6 regions turned into *regional dummies*.

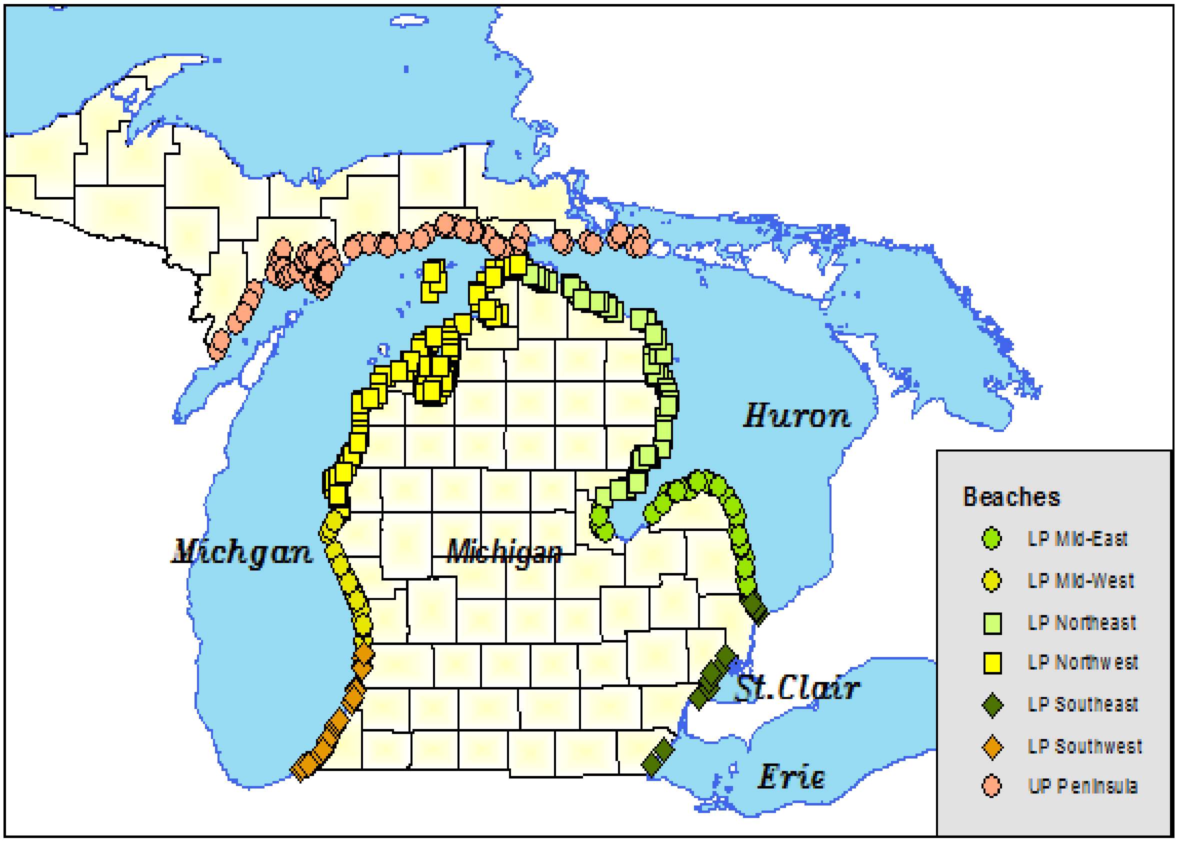


Figure 1-3 The 451 public great lakes beaches by region.

The indirect utility for individual *n* who chose *not* to take a trip is:

$$V_{No} = \gamma_{male} * male + \gamma_{age} * age + \gamma_{white} * white + \gamma_{edu} * edu + \gamma_{fulltime} * Fulltime + \gamma_{retire} * Retire + \gamma_{under17} * under17 + constant$$

Table 1-4 reports descriptive statistics for site attributes in the indirect utility V_{jlt} and individual characteristics in V_{No}.

Table 1-4 Descriptive statistics for individual characteristics and site attributes

Variables	Definition	Mean	Std. Dev	Min	Max
Socioeconomic characteristics (sample size=6375; sample weights applied)					
male	Dummy: 1=yes, 0=no	0.49	0.50	0	1
age	Age	46.53	18.53	17	99
white	Dummy	0.86	0.34	0	1
edu	Years of education	14.38	2.47	10	19
Fulltime	Full time employed, Dummy	0.47	0.49	0	1
Retire	Dummy	0.24	0.42	0	1
under17	Dummy for Children under 17	0.33	0.47	0	1
Income	Thousand in 2011 dollars	73.13	61.07	12.50	300.00.
Site Attributes (sites=451)					
Beach length	Miles	0.76	1.40	0.01	13.11
Temperature	June Temperature	55.50	4.24	48.87	72.57
	July Temperature	67.20	4.385	58.05	81.34
	August Temperature	67.76	4.59	58.49	78.93
	September Temperature	62.28	3.35	55.75	70.40
Closure days	Beach closure days of 2010	1.17	7.56	0	112
Regional dummy	LP northeast	0.20	0.40	0	1
	LP Mideast	0.09	0.29	0	1
	LP southeast	0.04	0.20	0	1
	LP northwest	0.33	0.47	0	1
	LP Midwest	0.06	0.24	0	1
	LP southwest	0.07	0.25	0	1

The trip data as described in section 3.2 consists of the regular matched beach data, grouped beach data and unknown beach data. The resulting structure for the probabilities for this irregular data set cannot be accommodated using standard software packages for the nested logit model. Moreover, the panel data used in this essay contains a time-variant variable, i.e., water temperature, and the choice set consists of 451 alternatives for each observation. The resulting complexity of the panel data increases additional computation burdens. Therefore, the log likelihood function was programmed in matrix language in MATLAB to perform full information maximum likelihood (FIML) procedure. Results from sequential estimation are used as starting values for FIML estimation. Depending on the operating system, estimation usually takes around two hours. Finally, since there are correlations that could arise from repeat observations from the same individual throughout the season, bootstrapping is used to correct for clustering on repeated trips. We bootstrapped 120 draws of the sample to get the bootstrapped standard errors. Given the intensive computation burden, the bootstrap procedures were divided into four smaller computational jobs using remote Compute Serves, which took about 2 to 3 days by using matrix programming of the bootstrap procedures in MATLAB.

4. Results

4.1 Estimation Results

The estimated parameters of the repeated nested logit model for all trip data are presented in Table 1-5. Based on the sign and magnitude of the estimated parameters, the results indicate that travel cost has a negative effect on the probability of choosing a site, which is consistent with our expectation that a higher price leads to lower demand. An increase in beach length increases the probability of choosing a beach; likewise, an increase in water temperature increases the probability of choosing a beach. Thus, an increase in beach length or warmer water temperature will increase demand. The number of closure days in the previous year negatively affects the probability of visiting the beach. Regional dummies reveal that Lake Michigan attracts the most Michiganders, while Lake St. Clair and Lake Erie are less popular, all else equal.

The nesting parameters measure the degree of independence in nests of each level. More intuitively, one minus the nesting parameter is an indicator of the correlation among alternatives within a nest. Therefore, the error terms for beaches are more correlated within each lake than across lakes. When nesting parameters are equal to 1, the nested logit reduces to the conditional logit model. In that sense, the nested logit with nesting parameters, which is significantly different from one, means the nested logit model provides a significant improvement over conditional logit by relaxing the property of independence from irrelevant alternatives (IIA) in logit model.

Regarding the demographic variables, the signs for all the estimated parameters make intuitive sense. In particular, the parameter for having higher education significantly and negatively affects the decision of *not* taking a trip in a choice occasion at a statistical confidence level of 95%. That is to say, Michiganders with higher education level take more trips.

Table 1-5 Full information maximum likelihood (fiml) estimation result

Nested Levels	Variable	Estimates	Bootstrapped Standard Errors	t statistic
Beach Level	Travel Cost	-0.014***	0.001	-10.261
	Log(Length)	0.076***	0.011	7.153
	Temperature	0.025***	0.004	6.443
	Closure Days of 2010	-0.010***	0.002	-4.861
	LP Northeast	-0.096	0.117	-0.827
	LP Mid-East	-0.665***	0.115	-5.770
	LP Southeast	-0.709***	0.123	-5.780
	LP Northwest	0.454***	0.083	5.498
	LP Mid-West	0.354***	0.089	3.984
	LP Southwest	0.045	0.090	0.493
Lake Level	Nesting Parameter	0.347***	0.024	14.228
Trip Level	Nesting Parameter	0.501***	0.040	12.615
No Trip	Male	-0.150	0.096	-1.550
	Age	0.003	0.003	0.911
	White	-0.314	0.196	-1.605
	Education Years	-0.113**	0.017	-6.819
	Full-Time Employed	-0.012	0.115	-0.101
	Retired	0.155	0.153	1.011
	Children under 17	0.095	0.078	1.227
	Constant	7.441***	0.461	16.146

Note: *10% significance level; **5% significance level; *** 1% significance level

4.2 Welfare Results

This section provides the welfare benefits for policies ranging from site closure to improvements in site quality. Specifically, we consider three types of policy scenarios: marginally increasing the length of one beach in one region, closing each beach in a region one at a time, and closing all beaches at a Great Lake. For each policy trip changes and seasonal welfare measures were calculated at individual level. For seasonal value per trip, according to Section 2.3, there are two ways to normalize the seasonal value: one way is dividing the seasonal value by the change of total trips, the other is dividing the seasonal value by the change of trips to the site or sites affected by a policy. At the population level, seasonal consumer surplus value was derived from aggregating the seasonal value per person to all Michiganders living in the Lower Peninsula.

Take the scenario of closing one beach in one region as an example. As described in section 3.3, there are 7 regions *R* in Michigan. For a region *r*, let R_r denote the number of beaches in that region, and let $CS_{j,n}$ denote the seasonal welfare estimates for person *n* to beach *j*. Taking the weighted average across the sample population gives the average per person seasonal value to beach *j*.

$$\overline{CS_j} = \frac{\sum_{n=1}^{N} w_n * CS_{j,n}}{\sum_{n=1}^{N} w_n}$$

This number is then computed for each other beach in the region. Taking the weighted sum of the seasonal value per person to a specific beach across *all* the beaches in the region *r* gives the weighted average seasonal value per person to *any* beach in that region. That is to say, if there is one beach closed in region *r*, the average seasonal value per person is:

$$\overline{CS}^r = \frac{1}{R_r} * \sum_{j=1}^{R_r} \overline{CS_j}$$

where R_r is the number of sites in region *r*.

Now considering the changed trips, taking the weighted sum of the average of the total changed trips per person across *all* the beaches in region *r* gives the total number of changed trips, given one beach is closed in that region.

$$\Delta\bar{Y}_G^r = \frac{1}{R_r} * \sum\nolimits_{j=1}^{R_r} \Delta\bar{Y}_{G,j} = \frac{1}{R_r} * \sum\nolimits_{j=1}^{R_r} \frac{\sum_{n=1}^{N} w_n * \Delta\hat{Y}_{G,j,n}}{\sum_{n=1}^{N} w_n}$$

Then we can get the first season value per trip from normalizing the average seasonal value per person by the change of total trips in the region *r* as:

$$\overline{\overline{CS}}_G^r = \frac{\overline{CS}^r}{\Delta\bar{Y}_G^r}$$

Similarly, if one beach is closed in the region *r*, the change in the number of trips to any beach in the region *r* is:

$$\Delta\bar{Y}_{jlG}^r = \frac{1}{R_r} * \sum\nolimits_{j=1}^{R_r} \Delta\bar{Y}_{jlG} = \frac{1}{R_r} * \sum\nolimits_{j=1}^{R_r} \frac{\sum_{n=1}^{N} w_n * \Delta\hat{Y}_{jlG,n}}{\sum_{n=1}^{N} w_n}$$

Thus, the second season value per trip is found by dividing the average seasonal value per person by the change in the number of trips to any beach in the region *r*.

$$\overline{\overline{CS}}_{jlG}^r = \frac{\overline{CS}^r}{\Delta\bar{Y}_{jlG}^r}$$

Table 1-6 Welfare estimates of changing a beach in 2011 dollars per person

		Seasonal Value	Seasonal Value per Trip	
			Season/Total Trip Change	Season/Site Trip Change
Closure of One Beach in the Region	Huron North	-0.09	74.75	24.76
	Huron South	-0.20	73.07	24.81
	St. Clair	-1.12	72.65	25.77
	Erie	-2.20	72.61	28.07
	Michigan North	-0.11	75.68	24.74
	Michigan Central	-0.89	74.67	24.88
	Michigan South	-0.45	74.20	24.85
Marginal Increase in Length of One Beach in the Region	Huron North	0.06	75.20	25.03
	Huron South	0.09	74.31	25.41
	St. Clair	0.65	72.92	27.34
	Erie	0.65	72.56	31.30
	Michigan North	0.05	72.26	23.55
	Michigan Central	0.27	74.83	25.17
	Michigan South	0.19	74.17	25.06

Table 1-6 displays the regional differences in the welfare measures arising from the changes of a beach. It seems counter-intuitive at first that Lake Erie has the largest seasonal welfare losses if one beach is closed. However, in our choice site, there are only 2 beaches at Lake Erie, 6 beaches at Lake St. Clair, and all other 443 beaches are at Lake Huron and Lake Michigan. Therefore, if one beach is closed in Lake Erie or Lake St. Clair, people's utility decreases dramatically as their substitution is limited. When we consider the seasonal value per trip, based on the first welfare measure (the column titled as "Season/Total Trip Change"), Erie turned out to have the lowest seasonal value per trip, followed by St. Clair as the second lowest. Michigan North has the highest seasonal value per trip, followed by Huron North as the second highest. Regarding the second measure of the seasonal value per trip (the column titled as "Season/Site Trip Change"), the welfare loss for all regions is far less than the first seasonal value per trip. The reason is that total trip changes are smaller than site trip changes; site trip changes mostly are due to substitution from other sites and only partly due to changes in total numbers of trips. Furthermore, Erie and St. Clair's values again become higher in the second measure of seasonal value per trip because their site trip changes are smaller when there comparatively few similar substitute sites in their nest.

If we increase the beach length by one mile on one beach, the seasonal welfare benefits to Michiganders are also larger for Erie and St. Clair. The average length of a public Great Lakes beach in our choice set is 0.76 miles. By contrast, among 8 beaches on the Eire and St. Clair, the maximum length of the beach is just 0.42 miles and the minimum length is only 0.01 miles. Similarly, Huron South has many short beaches as well, with the average length of a beach as 0.46 miles. Moreover, once we take the logarithm of the beach length, the utility of a person is increasing as the beach length is increasing, but the utility increases at a slower rate as the length increases. Therefore, a marginal increase of beach length leads to more utility increase for shorter beaches in Erie and St. Clair than for long beaches in Huron North and Michigan. Correspondingly, the welfare gains accrued to Michiganders are smaller for Huron North and Michigan. Similarly, if we use the "Season/Total Trip Change" as the seasonal value per trip, Michigan North once again has the highest welfare gain, followed by Huron North as the second highest. Huron South turned out to have the lowest season value per trip, followed by Erie as the second lowest.

Table 1-7 Welfare estimates of changing a beach in 2011 dollars (million) at state level

		Seasonal Value (Million)
Closure of One Beach in the Region	Huron North	-0.689
	Huron South	-1.462
	St. Clair	-8.183
	Erie	-16.043
	Michigan North	-0.810
	Michigan Central	-6.490
	Michigan South	-3.305
Marginal Increase in Length of One Beach in the Region	Huron North	0.470
	Huron South	0.672
	St. Clair	4.768
	Erie	4.774
	Michigan North	0.346
	Michigan Central	1.986
	Michigan South	1.385

To calculate the population level welfare, we have to aggregate the weighted average seasonal value at the individual level to the all Michiganders living in the Lower Peninsula. The population number of Michigan adult residents is obtained from the 2010 census as 7,289,085.

Table 1-7 shows the welfare estimates at the population level when a beach is closed or the beach length increases by one mile. Total Michiganders' welfare loss due to a beach closure in a region was estimated to be about $0.69million to $16.04million. Total Michiganders' welfare gain arising from a marginal increase in beach length in a region was estimated to be about $0.35million to $4.77 million. For the same reason that individual seasonal value is larger in Erie and St. Clair, seasonal value at the population level is also larger in Erie and St. Clair.

Table 1-8 Estimated trips and welfare changes of closing all beaches on a great lake in 2011 dollars

Per Person					
		Number of Trips	Seasonal Value	Seasonal Value per Trip	
				Season/Total Trip Change	Season/Site Trip Change
Closure of All Beaches on a Great Lake	Lake Erie	0.15	-5.87	72.52	38.06
	Lake St. Clair	0.25	-9.78	72.52	38.75
	Lake Huron	0.84	-35.35	73.16	41.95
	Lake Michigan	2.55	-146.11	73.52	57.32

State Level (Million)

		Number of Trips (Million)	Seasonal Value (Million)
Closure of All Beaches on a Great Lake	Lake Erie	1.124	-42.767
	Lake St. Clair	1.840	-71.304
	Lake Huron	6.142	-257.650
	Lake Michigan	18.580	-1065.000

Table 1-8 displays the welfare estimates and predicted trips from closing an entire lake. The loss of a Michigander's welfare associated with the elimination of a Great Lake can also be described as the welfare benefits accrued to a Michigander for access to beaches on that lake. Since the majority of the public Great Lakes beaches are located at Lake Michigan and Lake Huron, Lake Michigan generated the highest seasonal welfare measure, with $146.11 in seasonal value obtained from an average Michigan adult resident. When normalized by the site trip change, the season value per trip per person is $57.32. Lake Erie has the lowest seasonal welfare value per person at $5.87, the lowest seasonal welfare value per person per trip at $38.06. On average, a Michigan adult resident takes 2.55 trips to the beaches at Lake Michigan, followed by Lake Huron with 0.84 trips. Lake Erie and Lake St. Clair are much less popular. By contrast, a Michigander only takes 0.15 trips to Lake Erie and 0.25 trips to Lake St. Clair. When aggregated at the population level, seasonal recreational value from Lake Michigan can be realized as $1.06 billion by all Michigan adult residents living in the Lower Peninsula; seasonal value from Lake Erie is lowest at $42.77 million.

5. Conclusions

Beach recreation is an important outdoor activity and is of great value to the beachgoers in Michigan. Despite the importance of beach recreation, not many environmental valuation studies have covered Great Lakes beaches. This essay contributes to this area of study by focusing on applying all trip data from a general population survey to Michigan adults to estimate the economic value of the public beaches on Lake Erie, Lake St. Clair, Lake Huron and Lake Michigan.

The economic estimates and welfare measures of this essay provide policy makers and beach managers with a better understanding of the factors determining Michiganders' site selection and the economic benefits associated with changes in the level of particular beach site attributes. Furthermore, the information on economic benefits is useful for beach restoration and protection programs. Finally, the economic estimates and welfare measures can be applied in subsequent benefit transfer studies.

We found that on average a Michigan adult resident takes 3.8 trips to the Great Lakes beaches in summer. Generally speaking, Michiganders prefer beaches with lower travel cost, longer beaches, beaches with warmer water temperature and all else equal beaches at Lake Michigan. Among all Michigan adults living in the Lower Peninsula, people who are male, with higher education level, not employed full-time, without children under 17 and with more income tend to take more trips.

The seasonal value of access to a public Great Lakes beach ranges from $24.74 to $28.07per person per trip, depending on the region. If we only use single day trip data, the seasonal access value reduced to two-thirds of the value, i.e. $14.25 to $17.24 (Chen, 2013). For all the Michigan adult residents living in the Lower Peninsula, the total Michiganders' welfare loss due to a beach closure in a region was estimated to be about \$0.69 million to \$16.04 million. Although Lake Erie seems to have the highest seasonal value at \$28.07 per person per trip, it only has two alternatives in our choice set. The limited substitution pattern makes the value of Lake Erie higher than those of the other regions and may be an artifact of the nesting structure. However, when we normalized the seasonal value by the total trip changes for a season rather than trip changes to a specific region, the substitution effects counteract each other within the same region. Thus, Michigan North has the highest seasonal value at \$75.68 per person per trip, followed by Huron North as the second highest seasonal value at \$74.75 per person per trip. Lake Erie has the lowest seasonal value per person per total trip change as we expected, followed by the Lake St. Clair. When we compare the seasonal value per trip in benefits transfer studies, the second welfare measure (the column titled as “Season/Site Trip Change” in Table 1-6) is more appropriate when comparing the results to single site models.

The values for access to beaches depend on the range of substitutes that are available. When one beach in a region closes, there are often many other beaches within that same region. However, when access of removed for all beaches at a Great Lake, the substitution patterns are more limited and the values are consequently higher. The seasonal value of access to a lake ranges from \$38.06 to \$57.32 per person per trip across Lake Michigan, Lake Huron, Lake St. Clair and Lake Erie. Lake Michigan is the most popular one among the four lakes in our choice set. On average, a Michigander takes 2.55 trips to the beaches at Lake Michigan during a beach season. In total there are just over 18 million trips taken by all Michigan residents living in the Lower Peninsula to the beaches at Lake Michigan. Lake Huron comes in second place and Lake Erie is the least popular one among the four lakes. At the state level, the seasonal value of access to Lake

Michigan beaches is around $1.06 billion, followed by Lake Huron at $257.65 million, Lake St. Clair at $71.3 million, and Lake Erie at $42.77 million.

In addition to estimating the economic value of the Great Lakes beaches in Michigan, this essay also raises a few empirical issues to discuss. First, we believe that using all trip data can help to derive the complete recreational demand curve, therefore the estimated results and welfare measures would be more accurate. Second, we found that using all partial sites information can make the economic estimates more accurate (see appendix C). It is common that respondents cannot always report the exact beach name in a survey. In such cases, aggregating the unmatched beaches into grouped beaches or even using the trip information without knowing the beach name can contribute to the accuracy of the estimation. Finally, our survey is based on the general population, making the results useful for future benefit transfer studies.

Meanwhile, this essay suffers an obvious caveat. Due to the limited access to the beach quality data, we are not able to include certain important beach attributes that may influence people' site choices. Although the regional fixed effects will capture the regional average differences of any missing attributes, there are likely other factors that also influence Michiganders' site choices that are not accounted for in this model. For example, Pendleton et al. (2012) shows that beach width also matters substantially to beachgoers. Future work may include more site attributes into consideration upon data availability. The alternative way is using a combination of revealed preference data and stated preference data. Essay 2 explores the possibility of incorporating additional water quality attributes by using joint estimation of revealed preference data and stated preference data.

ESSAY 2 Combining Revealed and Stated Preference Methods for Valuing Water Quality Changes to Great Lakes Beaches in Michigan

1. Introduction

Water quality of the Great Lakes is highly valued by policy makers and the public. Many legislative efforts and government regulations, such as Clean Water Act (CWA, 1970, 1972) and Great Lakes Water Quality Agreement (GLWQA, 1972, 1978, 1987, 2012), have been enacted to restore and enhance the water quality of the Great Lakes over the last decades. Public policies toward water quality can benefit from information about the economic benefits of improvement or protection of water quality. Although valuing water quality changes is particularly challenging as compared to other environmental services (Keeler et al. 2012), we can estimate some of the monetary value of water quality improvements by measuring the recreational benefit of water quality improvement, as one of the major benefits from improving water quality accrues to recreational use (Bockstael, Hanemann, & Kling, 1987).

Two primary approaches have been applied to the measurement of recreational benefits: revealed preference (RP) approaches and stated preference (SP) approaches. RP approaches, such as the "travel cost method", rely on observed behaviors to indirectly derive values of environmental services. By contrast, SP approaches, such as "choice experiments" or the "contingent valuation method", ask the individual to make hypothetical choices to directly elicit values.

Both RP and SP approaches have advantages and disadvantages, and each approach faces challenges in valuing water quality changes. For RP approaches, challenges in valuing the water quality changes mainly lie in three aspects. First, unlike air quality, which has a comparatively small number of accepted measures of quality, water quality is scaled by a large number of chemical and biophysical variables. Evaluating overall water quality status from a large number of variables is often difficult (Kannel et al. 2007). Second, understanding the link between the biophysical characteristics and the recreational attributes of water quality has long been, and continues to be a challenge for selecting the appropriate variables to describe water quality (Kneese & Bower, 1968; Keeler et. al, 2012). Third, among the few studies conducted on valuing water quality by using biophysical attributes, they either require a considerably rich dataset (Egan et al. 2009), or they often suffer from problems of multicollinearity (see Bockstael, Hanemann, & Kling, 1987 for a discussion) or missing attribute levels, as suggested by Adamowicz et al. (1997). On the other hand, although SP approaches can readily address subjective measures of water quality changes, SP approaches have been criticized for being hypothetical because their estimates are based on respondents' *ex ante* choices.

Inspired by the fact that the some of the strengths of RP approaches are possible weaknesses of SP approaches, and vice versa (see Whitehead et al. 2008 for a detailed review), a combination of the two methods to jointly estimate RP and SP data has been proposed (Cameron, 1992; Adamowicz, Louviere & Williams, 1994). Based on the underlying theoretical framework, the RP and SP literature in environmental economics can be classified into two strands: those based in random utility theory (RUM), and others. When RP and SP studies are structured as RUM models, the combined approach also follows RUM. A typical example is combining RUM travel cost models with the choice experiments (Adamowicz et al., 1994, 1997; Von Haefen and Phaneuf, 2008). The other strand of literature has different theoretical foundations of RP and SP data, in which at least one model does not follow the RUM theory, such as combinations of contingent valuation and travel cost methods (Cameron, 1992; Loomis, 1997; Huang, Haab, & Whitehead, 1997).

Despite its merits, some argue that combining RP and SP data should be subjected to a consistency test (Morikawa, 1989; Swait and Louviere, 1993; Adamowicz et al., 1994; Von Haefen and Phaneuf, 2008), which is a statistical test of the equality of common parameters in RP and SP models. Empirical evidence about combining RP and SP data in environmental economics, however, is mixed. Some applications have passed the test and concluded that the RP and SP data contain similar preference structure and thus can be combined (Adamowicz et al. 1994, 1997; Carson et al. 1996; Huang et al. 1997; Whitehead et al. 2010). However, many applications have rejected the test (Earnhart, 2001; Haener, Boxall, & Adamowicz, 2001; Azevedo, Herriges & Kling, 2003; Von Haefen & Phaneuf, 2008; Hoyos & Riera, 2013; Jeon, 2014). For instance, even though Adamowicz et al. (1994) found the common parameter equality existed in RP and SP data, Von Haefen and Phaneuf (2008) and Jeon (2014), using the same datasets, rejected consistency between the RP and SP data respectively by using different methods, but still within the RUM framework.

The purpose of this study is to estimate the values of water quality changes for beach recreation in the Great Lakes. By using data from the same web survey of 2,537 Michigan beachgoers, this essay builds on the Essay 1 and an earlier SP study by Weicksel (2012). The web survey consists of two types of data: one is revealed preference data, which is collected by asking about respondents' trips to public beaches at the Great Lakes in Michigan; and the other is stated preference data, which involves asking respondents in a choice experiment to choose from hypothetical choice sets in which the beaches were constructed with different environmental quality attributes. In Essay 1, we employed all trip data to estimate the use value of Great Lakes beaches. Weicksel (2012) used the choice experiment data to estimate preferences for water quality attributes at Great Lakes beaches. However, each data set alone would not be sufficient to value the water quality changes. Therefore, this essay extends the two proceeding studies by combining the two datasets to jointly estimate the values of water quality changes.

In this study, we combine trip data (RP) and choice experiment data (SP) to offer four advantages. First, the combined method makes use of water quality measures from choice experiment data, which avoids potential multicollinearity problems and missing attribute levels from using observed physical measures and reduces the data collection burden. More importantly, the water quality attributes from the SP data are designed to be policy-relevant since they match those that the EPA collects through its occasional beach sanitation surveys (EPA, 2008). Second, the constructed physical indices from choice experiments are easy to understand, match what people can see at beaches, and are likely more relevant to beach recreation than water chemistry and related physical measures. Third, combining data can ground the stated choices from choice experiments within actual trip choices from the travel cost model. Finally, the RP data includes a large number of beach sites (451 alternatives) which enables us to better capture a rich array of substitution effects of trip demand in response to water quality changes.

Furthermore, few environment valuation studies have focused on water quality of the Great Lakes. Huang, Poor and Zhao (2007) combined travel cost method and contingent valuation method to measure the impact of erosion and erosion control programs at eight ocean beaches in New Hampshire and southern Maine. Parsons, Helm, and Bondelid (2003) applied travel cost methods and set up three scenarios for water quality improvements in six northeastern states, and estimated annual benefits in the region due to CWA to be near $100 million per year. Egan et al. (2009) used a mixed logit model and collected extensive physical water quality attributes of 129 lakes in Iowa to value water quality changes. Still, little is known about the value of water quality changes in the Great Lakes. Knowing some of the values of water quality changes, specifically for the Great Lakes, could help fill the gap in the literature and help policy makers better allocate funds and evaluate water quality restoration or improvement programs.

The remainder of the paper proceeds as follows. Section 2 first provides a brief review of the underlying theoretical framework (i.e. Random Utility Model). Within the RUM framework, we further present the revealed preference approach, the stated preference approach, and combined RP and SP approach. Section 3 describes the Great Lakes beaches survey and datasets, which is followed by the empirical specifications of the models in section 4. Estimation results and hypothesis testing are then presented in section 5. Section 6 describes the method to calculate welfare measures and presents the welfare results, and the final section provides conclusion and discussion.

2. Models

2.1 The Random Utility Model (RUM)

The random utility model is widely used in recreation demand studies where an individual chooses among a set of sites to visit. On a single choice occasion, the RUM considers the choice of one site from many mutually exclusive recreational sites to be a function of attributes of the sites. Based on individual's choice, the model implicitly measures the trade-off between site attributes. If we include travel cost into the site attributes, we can get the implicit value of site attributes in dollar terms.

More formally, following Train (2009), we assume a sample of N travelers with the choice set C, and the utility that individual n derives from choosing alternative j from the set C is denoted by

$$U_{jn} = V_{jn} + \varepsilon_{jn}.$$

The systematic component, V_{jn}, is observable to researchers and usually is a function of the attributes of alternative *j* and the individual's socio-demographic characteristics, while the random term ε_{jn} captures all the factors unobservable to researchers. Individuals choose the alternative which generates the highest utility, so the probability that individual *n* chooses alternative *i* rather than alternative *j* is equal to the probability that the utility of choosing *i* is higher than the utility of choosing *j*:

$$P_{in} = P\left(U_{in} > U_{jn}, \forall\, j \in C\right) = P\left(\varepsilon_{in} - \varepsilon_{jn} > V_{jn} - V_{in}, \forall\, j \in C\right)$$

This probability has a cumulative distribution that depends on the density $f(\varepsilon_{jn})$. Different assumptions about the distribution of the unobserved parts of utility (i.e., the random term), will yield different random utility models. When each random term is distributed as generalized extreme value (GEV), it is a nested logit model, which is described further with the application in section 2.2. When the random term is iid with extreme value distribution, it is a conditional logit model, which will be applied in the choice experiment data in section 2.3.

2.2 Repeated Nested Logit Model for Trip Data (RP)

Following Essay 1, a repeated three-level nested logit model is applied to all trip data, which explains the site choice and recreation demand of trips to Great Lakes beaches in a summer season. The model is called "repeated" because the season is divided into choice occasions in which beachgoers decide whether or not to visit a beach. The trips can be a day trip or multiple-days trip.

Generally, in a three-level nested logit model, the alternatives in choice set *C* are grouped in M nests. The decision process can be visualized as choosing among the *M* nests, *M*= {*Trip, No trip*}, among the *L* lakes in the nest *Trip*, *L* = {*Lake Erie, Lake St. Clair, Lake Huron, Lake Michigan}*, and among the *J* beaches at one of the lakes *l*. The decision tree is illustrated in the figure below:

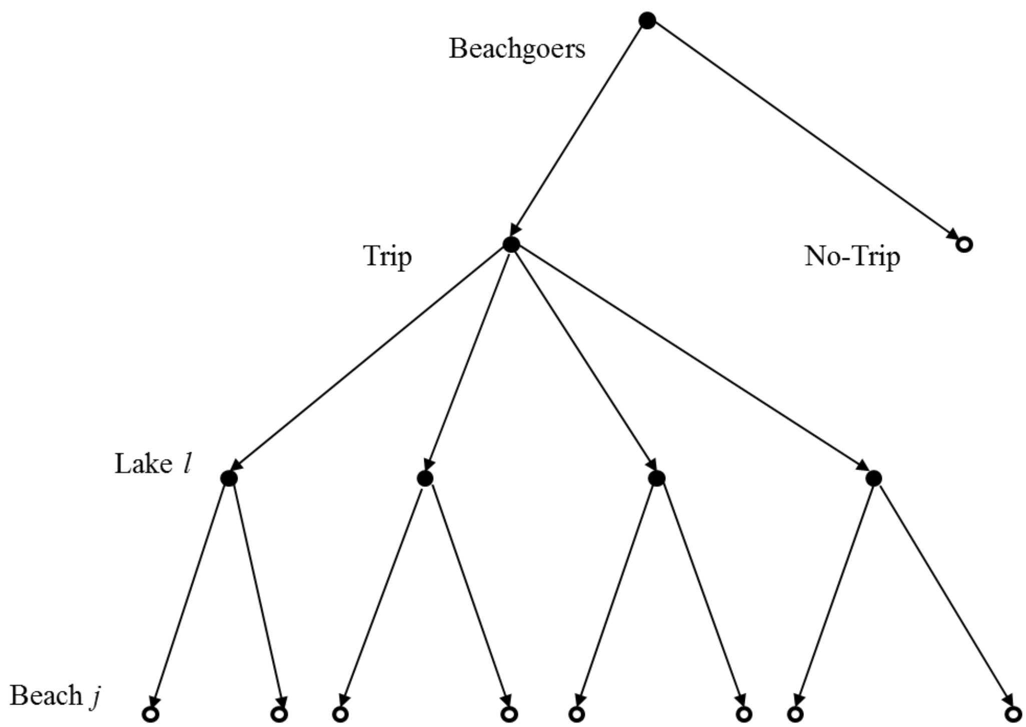


Figure 2-1 Repeated three level decision tree of beach recreation trip

Formally, the utility of a three-level nested logit is given as (individual subscript *n* is omitted to simplify the notation):

$$U_{jlm} = V_{jlm} + \varepsilon_{jlm},\ \ \forall\ (jlm) \in C$$

Assume that the joint density function of the random term is given by the first type of generalized extreme value (GEV) distribution with three nests (McFadden, 1978):

$$F(\varepsilon_{jlm}) = exp\left\{-\sum_{m\in M}\left[\sum_{l\in L_m}\left[\sum_{j\in J_{lm}} exp\left(-\frac{\varepsilon_{jlm}}{\lambda}\right)\right]^{\frac{\lambda}{\rho}}\right]^{\rho}\right\}$$

where

- Beach alternatives *J*= {1, 2, …, 451};
- Lake alternatives *L*= {*Lake. Erie, Lake St. Clair, Lake Huron, Lake Michigan}*;
- Trip alternatives *M*= {*G, No*}; (*G* is short for Trip, *No* is short for No Trip)

The probability of beach *j* being chosen is given by

$$P_{jlG} = P_{(j|lG)} * P_{(l|G)} * P_G$$

where $P_{(j|lG)}$ is the conditional probability of choosing beach *j* given that lake *l* and trip alternative *G* is chosen. $P_{(l|G)}$ is the conditional probability of choosing lake *l* given a trip alternative *G* is made. P_G is the probability of taking a trip. Then, the indirect utility of not taking a trip can be denoted as V_{No}.

The conditional and marginal probabilities are given by:

$$P_G = \frac{\exp(\rho IV_G)}{\exp(\rho IV_G) + \exp(V_{No})}$$

$$P_{(l|G)} = \frac{\exp\left(\frac{\lambda}{\rho} IV_{lG}\right)}{\sum_{k\in L_m}\left[\exp\left(\frac{\lambda}{\rho} IV_{kG}\right)\right]}$$

$$P_{(j|lG)} = \frac{\exp\left(\frac{1}{\lambda}V_{jlG}\right)}{\sum_{i\in J_{lm}}\left[\exp\left(\frac{1}{\lambda}V_{ilG}\right)\right]}$$

The expected utility that each beachgoer receives from the choice of alternatives within each nest is called an inclusive value. IV_G and IV_{lG} are the inclusive values of Trip nest G and Lake nest respectively, where

$$IV_G = \ln\left[\sum_{k\in L_m}\left[\exp\left(\frac{\lambda}{\rho}IV_{kG}\right)\right]\right]$$

$$IV_{lG} = ln\left[\sum_{i\in J_{lm}}\left[\exp\left(\frac{1}{\lambda}V_{ilG}\right)\right]\right]$$

Finally, the unconditional probability of taking a trip to beach *j* is:

$$P_{jlG} = \frac{\exp(\left(\frac{1}{\lambda}V_{ilG}\right) * \left[\sum_{l\in L_m}\left[\sum_{j\in J_{lm}} exp\left(\frac{1}{\lambda}V_{jlG}\right)\right]^{\frac{\lambda}{\rho}}\right]^{\rho-1} * \left[\sum_{j\in J_{lm}} exp\left(\frac{1}{\lambda}V_{jlG}\right)\right]^{\frac{\lambda}{\rho}-1}}{\left[\sum_{k\in L_G}\left[\sum_{i\in J_{kG}} exp\left(\frac{1}{\lambda}V_{ikG}\right)\right]^{\frac{\lambda}{\rho}}\right]^{\rho} + \exp(V_{No})}$$

The unconditional probability of not taking a trip to any beach is:

$$P_{No} = \frac{\exp(V_N)}{\left[\sum_{k\in L_G}\left[\sum_{i\in J_{kG}} exp\left(\frac{1}{\lambda}V_{ikG}\right)\right]^{\frac{\lambda}{\rho}}\right]^{\rho} + \exp(V_{No})}$$

Then, the expected maximum utility for each choice occasion, or the inclusive value of each individual *n*, can be obtained as:

$$IV = ln\left\{\left[\sum_{k\in L_G}\left[\sum_{i\in J_{kG}} exp\left(\frac{1}{\lambda}V_{ikG}\right)\right]^{\frac{\lambda}{\rho}}\right]^{\rho} + \exp(V_{No})\right\}$$

Let *T* denote the total number of choice occasions, called the beach season, and *T*=126. Let $y_{jlG,nt} = 1$, if person *n* visited beach *j* at Lake *l* on occasion t, and $\mathrm{y}_{jlG,\mathrm{nt}} = 0$, otherwise. As long as the beachgoer takes the trip to the beach *j*, $y_{jlG,nt}$ always equals 1, irrespective of the type of trip. To simplify the notation for probability expressions, individual *n* at time *t* will be noted after the comma in the subscript of the probability.

The log-likelihood function for this sample is:

$$LL_{beach}^{RP} = \sum_{n=1}^{N}\sum_{t=1}^{T}\left[\sum_{l\in L_G}\sum_{j\in J_{kG}} w_n * y_{jlG,nt} * \ln(P_{jlG,nt}) + w_n * (1 - y_{jlG,nt}) * \ln(P_{No,nt})\right]$$

where w_n is the weight of person *n.* There are three purposes of the weight (See Appendix A.3). The first is to correct for sampling strata and possible non-representativeness of the sample. The second use is to down-weight number of overnight trips due to the multiple purposes for overnight trips. The third to account for self-reported corrections to trip counts.

As in Essay 1, there is a type of trip data called “grouped beaches”, which has only partial information on the alternatives chosen. The reason is that some people only reported the nearest town or city to the beach, so we don’t know the exact beach name but only an aggregated area. We applied the same approach as Essay 1 to handle trips with partial information. Denoting the grouped area as *a*, the log-likelihood function for this sample of “grouped beaches” is:

$$LL_{group}^{RP} = \sum_{n=1}^{N}\sum_{t=1}^{T}\left[\sum_{l\in L_G}\sum_{j\in a} w_n * y_{jlG,nt} * \ln\left(P_{jlG,nt}\right) + w_n * (1 - y_{jlG,nt}) * \ln(P_{No,nt})\right]$$

That is, the log-likelihood function is the sum of the probabilities of visiting the individual sites within area a.

Finally, we have some reported beaches which were unknown to researchers because the way they were reported did not allow researchers to either locate the exact beach or aggregate the beach into groups. However, we do know that the respondent has taken the trip, so the unconditional probability P_G was applied to the unknown-beach samples yielding

$$LL_{unknown}^{RP} = \sum_{n=1}^{N}\sum_{t=1}^{T}\left[w_n * y_{G,nt} * \ln\left(P_{G,nt}\right) + w_n * (1 - y_{G,nt}) * \ln(P_{No,nt})\right].$$

The resulting log-likelihood function for all the samples in the trip data is:

$$LL^{RP} = LL_{beach}^{RP} + LL_{group}^{RP} + LL_{unknown}^{RP}$$

As we have observations with exact, grouped and unknown sites, conventional syntax in common statistical software can no longer accommodate our needs. Thus, we have to program the log-likelihood function in order to include all the information provided in the data.

2.3 Conditional Logit Model for Choice Experiment Data (SP)

When the correlations of random terms of the utility are zero, the nested logit model reduces to the conditional logit model. As a simple case of nested logit model, the conditional logit model is the easiest and most widely used random utility model (Train, 2009). In the present application, the choice experiment data is estimated using conditional logit model. In the choice experiments, beachgoers were asked to choose between two alternative beaches which vary in their distances and water quality attributes. The conditional logit model gives the probability that individual *n* chooses beach *i* as a function of travel cost and water quality attributes. Based on the individual's choice, the model implicitly captures the trade-off between travel costs and water quality attributes.

More formally, if the random terms of the utility are assumed to be independently and identically distributed with type 1 extreme value distribution, then the choice probability of choosing alternative *i* for individual *n* is:

$$P_{in} = \frac{\mathrm{e}^{V_{in}}}{\sum_{j \in C} e^{V_{jn}}}$$

Correspondingly, the log-likelihood function is:

$$LL^{SP} = \sum_{n=1}^{N} \sum_{i \epsilon C} w_{ns} * y_{in} * \ln(P_{in})$$

where $y_{in} = 1$ if person *n* chooses alternative i, and $y_{in} = 0$, otherwise. w_{ns} is the survey weight of person *n* to correct for sampling strata and possible non-representativeness of the sample.

2.4 Combination of RP and SP Data

Since both the preceding RP and SP approaches are random utility models, it is possible to combine both datasets. When combining different types of data, one needs to account for possible differences in residual variance in each dataset to avoid potential bias. Even under the same random utility framework, data from different data sets could have different variance for the unobserved portion of utility. Morikawa (1989) was one of the first to propose a scaling approach to address this problem by allowing RP and SP data to have different variances within a single model. The idea is to scale the variance of the unobserved factors of the SP data so that RP and SP display identical unobserved effects in a pooled model (see also Ben-Akiva and Morikawa, 1990; Ben-Akiva et al., 1994). Through proper scaling, RP and SP data can be pooled to jointly estimate the parameters of attributes in both datasets. The scaling approach has been applied to value environmental quality changes within the same random utility framework (e.g., Adamowicz et al., 1994; 1997; Earnhart, 2001; Von Haefen et al., 2008).

Formally, the utility functions for individual *n* for site *i* are defined as:

$$U_{in}^{RP} = \beta^{RP} X_{in}^{RP} + \omega Z_{in} + \varepsilon_{in}^{RP}, \forall\ i \in C^{RP}$$

$$U_{in}^{SP} = \beta^{SP} X_{in}^{SP} + \delta W_{in} + \varepsilon_{in}^{SP}, \forall\ i \in C^{SP}$$

where X_{in}^{RP}, X_{in}^{SP} is a vector of observed variables common to both the RP and SP data sets, such as travel cost and beach length. Z_{in} and W_{in} are vectors of observed variables specific to each data set. β^{RP}, β^{SP}, ω, δ are unknown parameters to be estimated. ε_{in}^{RP} and ε_{in}^{SP} are random terms unobserved by researchers.

The perquisite for the joint estimation is that RP and SP data are derived from "the same underlying preference structure" (Hensher and Bradley, 1993; Adamowicz et al., 1994; Louviere et al., 1999). In other words, combing the two data sources involves imposing the restriction that the common attributes have the same parameters in both data sources, i.e. $\beta^{RP} = \beta^{SP} = \beta$. This condition cannot be satisfied when different unobserved error variances are present in each data. However, the scaling approach introduces a scaling parameter θ:

$$\theta^2 = var(\varepsilon_i^{RP})/var(\varepsilon_i^{SP})$$

which enables $\beta^{RP} = \theta\beta^{SP} = \beta$, and thus the joint estimation of two data sets becomes possible. θ can be interpreted as the *relative scale* of SP data with respect to the RP data. (Swait and Louviere, 1993; Bradley and Daly, 1997; Hensher et al., 1998; Louviere, et al., 2000, p.253)

The final parameter vector to be jointly estimated is $\psi = (\beta, \omega, \delta, \theta)$. Assuming the two data sources come from independent samples, the log likelihood of the pooled data is simply the sum of the log likelihoods of the RP and SP data:

$$LL^{joint}(\psi) = LL^{RP}(X_{in}^{RP}, Z_{in} \,|\, \beta, \omega) + LL^{SP}\left(X_{in}^{SP}, W_{in} \,|\, \beta, \delta, \theta\right)$$

If the random terms of the RP and SP data for the same individual are not correlated, maximizing the joint log likelihood function yields consistent and efficient estimates. If the random terms are correlated between RP and SP data, the estimates are consistent but not efficient (Wooldridge, 2010).

3. Survey and Data

3.1 Survey

The data used for this study are drawn from the Great Lakes Beaches Survey[6], which was conducted by Lupi, Kaplowitz, Chen and Weicksel in 2011 and 2012. First, in order to recruit beachgoers, a mail survey on leisure activities was conducted with the general population of Michigan residents. A random sample of 32,230 was drawn from the Michigan driver's license list. To reduce potential self-selection bias that might over-select for those that visit the Great Lakes, the mail survey has numerous questions on a broad range of indoor and outdoor leisure activities, among which there was only one screening question for Great Lakes beach recreation

[6] See Chen (2013), Weicksel (2012) for additional details regarding the survey sampling and implementation.

during two summers in 2010 and 2011. Respondents who answered they had participated in beach recreation were counted as beachgoers and were subsequently invited to take a follow-up web survey.

There are three sections in the follow-up web survey: a travel cost section, which collected trip information about respondents' trips to public Great Lakes beaches in one summer season from Memorial Day weekend to September 30, 2011; a choice experiment section, which gathered respondents' preferred beach in each of three different choice sets with experimentally designed attributes; and finally, a section of demographic questions.

3.2 Data

In the mail survey dataset of 9,591 observations, 5,737 respondents indicated they had visited a Great lakes beach in 2010 or 2011, so they were invited to the web survey. There were 3,196 people who responded to the web survey resulting in a response rate for the web survey of about 59%. Essay 1 made use of all trip data to estimate the value of trips to Great Lakes beaches by applying a nested logit model. Among the 2,573 observations, 1,894 individuals took at least one trip to Great Lakes beaches during the beach season. The trip data consists of self-reported trips to Great Lakes beaches from Memorial Day weekend to September 30, 2011. After matching the reported beaches to the Michigan DEQ beach database, the choice set for each individual is comprised of 451 beaches. There are 643 people who had taken trips to Great Lakes beaches before but didn't take any trip during the indicated season, they are treated as potential users and also included in this study.

Weicksel (2012) utilized choice experiment data from the web-survey to estimate preferences for environmental quality attributes at Great Lakes beaches. The effective samples of respondents was 2,791, which had 254 more individuals than in the trip data because Essay 1 only kept the persons whom the web survey was addressed to, while Weicksel's study included respondents who are "other household member" or "someone else". To maintain the same data implementation procedure, in this combined study, we follow Essay 1 to keep 2,537 effective respondents for data analysis. This approach also ensures that the data weights[7] are consistent and the same for each individual.

In the choice experiment data, each respondent was presented with three choice scenarios, with each choice set including 2 beaches. One attribute of beach alternatives is called "label", which provided the name of the Great Lake where the beach was located (sometimes referred to as a "labeled" or "branded" choice experiment). The web survey had three types of labeling design for the choice experiment: one used "labeled" alternatives with the different Great Lakes; another with "same-labeled" alternatives where each lake in a choice set was for the same Great Lake but the lakes varied across choice sets; the third used "unlabeled" alternatives that did not give names of the Great Lakes.

[7] Detailed procedures of data weights can be found in the appendix C, Chen (2013)

Table 2-1 Sample size for each types of choice experiment data

Data types of choice experiment	Number of respondents	Number of choice sets
All	2494	7300
Labeled	946	2785
Same-labeled	581	1948
Unlabeled	967	3190

In this study, we only use "labeled" data. There are two reasons: first, according to Weicksel (2012), labeling does have a significant effect on people's choice decision; second, we tested for a common preference across the three designs and, like Weicksel, we reject pooling of the three types of labeling data. Therefore, the effective sample size of respondents for SP data is 946 in this study, while for RP data, the effective sample of respondents is 2,537.

4. Econometric Model Specification

4.1 RP Data

For trip data, following Essay 1, in occasion *t*, the indirect utility for individual *n* obtained from visiting beach *j* at Lake *l* is:

$$V_{jlt} = \beta_{tc} * travel\ cost_{jl} + \beta_l * \log\left(beach\ length_{jl}\right) + \omega_t * temperature_{jlt} + \omega_{cd} * closure\ days\ of\ 2010_{jl} + \omega_r * regional\ dummies_{jl}$$

Similarly, the indirect utility for individual *n* who chose not to take a trip is:

$$V_N = \gamma_{male} * male + \gamma_{age} * age + \gamma_{white} * white + \gamma_{edu} * edu + \gamma_{fulltime} * Fulltime + \gamma_{retire} * Retire + \gamma_{under17} * under17 + constant$$

The computation of travel cost also follows Chen (2013):

$$Travel\ cost = round\ trip\ travel\ distance * \$0.2422\ per\ mile + round\ trip\ travel\ time * (annual\ income/2{,}000) * (1/3)$$

The trip data as described in section 2.2 consists of the regular beach data, grouped beach data and unknown beach data. The resulting structure for the probabilities for this irregular data set cannot be accommodated using standard software packages for nested logit model. Therefore, the log likelihood function was programmed in matrix language in MATLAB to perform full information maximum likelihood procedure. Estimation usually takes around one to two hours.

Table 2-2 reports descriptive statistics for both individual characteristics and site attributes in the RP data.

Table 2-2 Descriptive Statistics

Variables	Definition	Mean	Std. Dev	Min	Max
Socioeconomic characteristics (sample size=2537)					
male	Dummy: 1=yes, 0=no	0.40	0.49	0	1
age	age	49.64	15.13	18	94
white	Dummy: 1=yes, 0=no	0.93	0.25	0	1
edu	Years of education	15.09	2.46	10	19
Fulltime	Full time employed, Dummy	0.50	0.50	0	1
Retire	Dummy	0.25	0.44	0	1
under17	Dummy for Children under 17	0.30	0.46	0	1
Income	2011 dollars	83758.22	60368.84	12500	300000
Site Attributes (sites=451)					
Beach length	Miles	0.76	1.40	0.01	13.11
Temperature	June Temperature	55.50	4.24	48.87	72.57
	July Temperature	67.20	4.385	58.05	81.34
	Aug Temperature	67.76	4.59	58.49	78.93
	Sep Temperature	62.28	3.35	55.75	70.40
Closure days	Beach closure days of 2010	1.17	7.56	0	112
Regional dummy	LP northeast	0.20	0.40	0	1
	LP Mideast	0.09	0.29	0	1
	LP southeast	0.04	0.20	0	1
	LP northwest	0.33	0.47	0	1
	LP Midwest	0.06	0.24	0	1
	LP southwest	0.07	0.25	0	1

4.2 SP Data

For the choice experiment data, each respondent has three choice sets, and each choice set consists of two beach alternatives. The indirect utility function for individual *n* to choose beach *i* is:

$$V_{in} = \beta_{tc}{}' * travel\ cost_{in} + \beta_{l}{}' * \log(beach\ length_{in}) + \boldsymbol{\delta W_{in}}$$

where $\boldsymbol{W}$ is the attributes level of water quality (see Table 2-3), and $\boldsymbol{\delta}$ is a vectors of unknown parameters. Travel cost and the logarithm of beach length are variables that are included in both the RP and SP models. Although Weicksel (2012) used one-way distance as an explanatory variable, we transformed the one-way distance to a round-way travel cost following the approach outlined above for the RP data.

Finally, the unit of beach length in the SP data is yard. In order to make the variable compatible with the RP data, we transform yards to miles and take the logarithm of the beach length. Table 2-3 lists the other water quality attributes and attribute levels for the SP model (travel costs and beach length are not show in the table).

Table 2-3 Explanations of attributes and attribute levels ($\boldsymbol{W}$) in sp data

Attributes	Attribute Levels
Label: Great Lakes name	Lake Michigan Lake Huron Lake St. Clair Lake Erie
Algae in the water	None Low (rarely come in contact with algae) Moderate (sometimes come in contact with algae) High (constantly come in contact with algae)
Algae on the shore	None Low (1-20% of the shore has algae) Moderate (21-50% of the shore has algae) High (more than 50% of the shore has algae)
Testing water for bacteria	Never Monthly Weekly Daily

4.3 Pooled Data

When pooling RP and SP data together, according to the scaling approach, we get the indirect utility for joint estimation as:[8]

$$V_{jlt}{}^{joint} = \beta_{tc} * travel\ cost_{jl} + \beta_l * \log\left(beach\ length_{jl}\right) + \omega_t * temperature_{jlt} + \omega_{cd} * closure\ days\ of\ 2010_{jl} + \omega_r * regional\ dummies_{jl} + \theta * \boldsymbol{\delta W_{in}}$$

where θ is the RP/SP scaling parameter, which is imposed on the SP data to allow the β coefficients to be the same for the common variables of both SP and RP data, up to the scale difference. However, since the indirect utility function for the pooled data is no longer linear in all the parameters, the joint log likelihood function is programmed in the MATLAB to perform full information maximum likelihood procedure. Estimation usually takes around three hours with starting values obtained from sequential estimation.

5. Estimation Results

5.1 Conditional Logit Model for Choice Experiment Data (SP)

The results of the conditional logit model for the stated preference data are presented in Table 2-4, and all the estimates have signs consistent with expectations. The results indicate that Michigan beachgoers prefer less algae in the water and less algae on the shore. Furthermore, magnitudes of estimated parameters of algae levels in the water are higher than those of algae levels on the shore, which reveals that beachgoers have a stronger dislike of algae in the water than on the shore. Regarding the frequency of testing water for bacteria, beachgoers prefer water tested daily to water tested weekly or never tested at all. All else equal, beachgoers favor Lake Michigan

[8] If the observation was from the SP data, then there would be a $\theta\beta_{tc}$ and $\theta\beta_l$ instead of just the $\beta's$.

the most, followed by Lake Huron. All the above results are similar to those found in Weicksel (2012).

For SP data only, the difference between this study and Weicksel (2012) lies in three aspects. One, this study applies conditional logit model, while Weicksel used random-effects logit model. Although random-effects logit model circumvents the restrictive assumption of homogeneity for the conditional logit model, random-effects model itself incurs high computational cost when it deals with large data sets and many alternatives. Since each choice set in SP data only has two alternatives, random-effects model works well with SP data alone. However, given that the RP data has 451 alternatives, once we combine the data sets, the computation burden of random-effects model would impede estimation (Wooldridge, 2010). Therefore, conditional logit model is applied in this study, and as mentioned above it yields results very similar to those of Weicksel (2012). Second, the number of observations differs from Weicksel's sample size of 1,041, with 3,062 choice sets. To be compatible with RP data, we only kept respondents to whom the web survey was addressed in the datasets (rather than other household members that may have completed the survey), which leaves us with 946 observations with 2,785 choice sets. In addition, the survey weight for each respondent is applied to correct for sampling strata and possible non-representativeness of the sample, which was not available for Wiscksel's study. Third, the variable definitions of two beach attributes differ. Weicksel used the one-way distance, and treated beach length as a categorical variable. This study transforms the one-way distance to round-trip travel cost, and treats beach length as a continuous variable and then takes the logarithm of beach length. As a result, we have two common variables in both RP and SP models which, along with common weights and sample definition, enables us to perform joint estimation on the pooled data.

Table 2-4 SP estimation result

Variables	Attribute levels	Estimates	Robust Standard Errors	t statistic
Travel Cost		-0.007***	0.001	-10.320
Log(length of beach)		0.164***	0.026	6.440
Algae in the water	None	1.554***	0.143	10.850
(base:high)	Low	1.382***	0.136	10.180
	Moderate	1.127***	0.131	8.590
Algae on the shore	None	1.326***	0.124	10.730
(base:high)	Low	1.048***	0.120	8.700
	Moderate	0.658***	0.112	5.890
Testing water for bacteria	Never	-1.449***	0.121	-12.020
(base:Daily)	Monthly	-0.226**	0.107	-2.110
	Weekly	-0.344***	0.109	-3.140
Label of Great Lakes	Lake Michigan	1.127***	0.127	8.850
(base: Lake Erie)	Lake Huron	0.490***	0.108	4.550
	Lake St. Clair	-0.013	0.102	-0.120

Note: *10% significance level; **5% significance level; *** 1% significance level

5.2 Repeated Nested Logit Model for Trip data (RP)

The results of the repeated nested logit model for the revealed preference data are presented in Table 2-5. Since there are correlations that could arise from repeat observations from the same individual across the season, bootstrapping was used to correct for clustering on repeated trips. We bootstrapped 120 draws of the sample to get the bootstrapped standard errors in MATLAB.

Based on the sign and magnitude of the estimated parameters, the results indicate that travel cost has a negative effect on the probability of choosing a site, which is consistent with our expectation that higher price leads to lower demand. An increase in beach length increases the probability of choosing the beach as does an increase in water temperature. That is to say, an increase in beach length and water temperature will increase demand. The number of closure days in the previous year negatively affects the probability of visiting the beach. Regional dummies reveal that Lake Michigan attracts the most beachgoers, while Lake St. Clair and Lake Erie are less popular, all else equal.

The nesting parameters measure the degree of independence in nests of each level. More intuitively, one minus the nesting parameter is an indicator of the correlation among alternatives within a nest. Therefore, the error terms for beaches are more correlated within each lake than across lakes. When nesting parameters are equal to 1, the nested logit reduces to the conditional logit model. In that sense, nesting parameters are significantly different from 1 which means that in the RP data the nested logit model provides a significant improvement over conditional logit by relaxing the property of independence from irrelevant alternatives (IIA) in logit model.

Regarding the demographic variables, the parameters for being male significantly and negatively affect the decision of *not* taking a trip in a choice occasion at a statistical significance level of 95%. That is to say, male beachgoers take more trips.

Table 2-5 RP estimation result

Nested Levels	Variable	Estimates	Bootstrapped Standard Errors	t statistic
Beach Level	Travel Cost	-0.0115***	0.0011	-10.8485
	Log(Length)	0.0643***	0.0089	7.2600
	Temperature	0.0216***	0.0036	6.0716
	Closure Days of 2010	-0.0083***	0.0021	-3.9628
	LP Northeast	-0.0457	0.0997	-0.4587
	LP Mid-East	-0.5189***	0.0956	-5.4288
	LP Southeast	-0.5545***	0.1103	-5.0279
	LP Northwest	0.3880***	0.0714	5.4306
	LP Mid-West	0.2920***	0.0780	3.7433
	LP Southwest	0.0239	0.0723	0.3301
Lake Level	Nesting Parameter	0.2959***	0.0230	12.8708
Trip Level	Nesting Parameter	0.4527***	0.0418	10.8342
No Trip	Male	-0.1860**	0.0901	-2.0638
	Age	-0.0040	0.0031	-1.2779
	White	0.1532	0.2003	0.7652
	Education Years	-0.0278	0.0179	-1.5507
	Full-Time Employed	0.1195	0.0950	1.2585
	Retired	0.1470	0.1487	0.9886
	Children under 17	0.1225	0.0810	1.5129
	Constant	5.2328***	0.4412	11.8606

Note: *10% significance level; **5% significance level; *** 1% significance level

5.3 Joint Estimation of RP and SP Data

The results of the FIML joint estimation of RP and SP data are presented in Table 2-6. Similar to the situation with the RP method, bootstrapping was used to account for clustering on repeated trips in RP data and repeated choices in SP data. The procedures for bootstrapping the standard errors for 120 draws were coded using matrix language in MATLAB. Since each model estimation takes about 3 hours and hence a total bootstrapping time of about 15 days, the task was divided into smaller jobs to simultaneously implement on multiple remote servers.

The scaling parameter represents the relative scale of SP model to RP model. When the scale is between 0 and 1, the SP model contains more variation in the errors than the RP model (Ben-Akiva & Morikawa, 1990). The estimated scaling parameter is 0.622, which indicates the variance of the random term in SP model is 2.58 times of that in RP model. Other studies have also found SP model contains more variation (Ben-Akiva & Morikawa, 1990; Von Haefen & Phaneuf, 2008)

Compared to the RP-only model results, most of the variables from the RP model maintain the same sign and have only a slight change in magnitude in the joint estimation results. For instance, travel cost, closure days of 2010, and nesting parameters almost remain the same in joint estimation[9]. All other parameters of statistically significant variables change within a relatively small magnitude of 3% or less.

Compared to the SP-only results, travel cost in the joint model was forced to increase by about 1.6 times, while the logarithm of the beach length decreased from 0.164 to 0.064. Most of

[9] The RP and SP data we weighted so that each RP and SP *choice* was given equal weight (Von Haefen & Phaneuf, 2008 pp.29 footnote 10). We also followed Adamowicz et al. (1997) to give each RP and SP *individual* equal weight. The result is robust to alternative weighting schemes for the RP versus SP data within the likelihood ratio test.

the water quality variables from SP-only model increased by roughly 1.6 times, the same amount that travel cost increased because the pooled results will maintain the underlying marginal rates of substitution implicit in the choice experiment data. The signs of the SP variables never change, mainly because almost all water quality attributes are statistically significant in SP-only model.

If one compares the estimated coefficient of travel cost in the above RP-only and SP-only models, the parameter of travel cost in SP-only method (-0.007) is only around two-thirds of the value in RP method (-0.0115). Meanwhile, the coefficient of the logarithm of beach length in the SP-only method (0.165) is 2.6 times higher than the value in RP-only model (0.0643). Given that there are only two common variables, the opposite direction of changes in each coefficient between these two methods suggests the pooled model may face difficulties with the hypothesis of common parameters. We can further use a likelihood ratio test to formally test the hypothesis.

Table 2-6 FIML joint estimation result

Model Levels	Nest Levels/ Variables	Variable/ Attribute Levels	Estimates	Bootstrapped s.e.	t statistic
RP	Beach Level	Travel Cost	-0.0115***	0.0010	-11.3729
		Log(Length)	0.0660***	0.0088	7.5099
		Temperature	0.0215***	0.0038	5.7158
		Closure Days of 2010	-0.0083***	0.0020	-4.1165
		LP Northeast	-0.0494	0.0942	-0.5243
		LP Mid-East	-0.5239***	0.0915	-5.7291
		LP Southeast	-0.5581***	0.1059	-5.2685
		LP Northwest	0.3827***	0.0672	5.6948
		LP Mid-West	0.2863***	0.0735	3.8961
		LP Southwest	0.0191	0.0696	0.2749
	Lake Level	Nesting Parameter	0.2957***	0.0219	13.4937
	Trip Level	Nesting Parameter	0.4522***	0.0396	11.4307
	No Trip	Male	-0.1858***	0.0902	-2.0593
		Age	-0.0040	0.0031	-1.2813
		White	0.1537	0.2041	0.7532
		Education Years	-0.0277	0.0178	-1.5575
		Full-Time Employed	0.1195	0.0900	1.3269
		Retired	0.1471	0.1429	1.0292
		Children under 17	0.1225	0.0799	1.5338
		Constant	5.2207***	0.4608	11.3301
Scale		Scaling Parameter	0.6223***	0.0822	7.5680
SP	Algae in the water (base:high)	None	2.4362***	0.2257	10.7925
		Low	2.1953***	0.2007	10.9399
		Moderate	1.8232***	0.1774	10.2802
	Algae on the shore (base:high)	None	2.1071***	0.2324	9.0667
		Low	1.6102***	0.2210	7.2847
		Moderate	0.9439***	0.1731	5.4526
	Testing water for bacteria (base:Daily)	Never	-2.2813***	0.2832	-8.0560
		Monthly	-0.3788**	0.1715	-2.2082
		Weekly	-0.5331***	0.1508	-3.5348
	Great Lake (base: Lake Erie)	Lake Michigan	1.8342***	0.2089	8.7820
		Lake Huron	0.7274***	0.1469	4.9534
		Lake St. Clair	-0.0329	0.1427	-0.2304

Note: *10% significance level; **5% significance level; *** 1% significance level

More formally, according to Swait and Louviere (1993), to accept the hypothesis of common parameter equality between RP and SP method, we have to pass the following likelihood ratio test:

$$-2(LL^{joint} - (LL^{RP} + LL^{SP}) \sim \chi(k-1)$$

where k is the number of common variables.

In its present form, our pooled model rejects the test of common preference parameters (see Table 2-7, Model 1). Given only 1 degrees of freedom, this test significantly rejects the hypothesis of equal parameters with scaling. This finding indicates that the variances from the error term in one preference method are different from those in the other one, and the scaling approach does not eliminate preference parameter differences in the current model specification. To increase the number of common variables that can explain the difference of the variances in the two data sets, we further decompose the beach length into 6 categorized variables in the RP model and 5 categorized variables in the SP model, with 4 categories being the same for both RP and SP data. Thus, including the travel cost variable, we have 5 common variables in Model 2. Still, Model 2 strongly rejects the common parameter test. In Model 3, we incorporate lake dummies into the RP model, and change the 7 regional dummies into North and South dummies. In this way, we have the 3 lake dummies, the logarithm of beach length, and the travel cost in both RP and SP data, which also give us 5 common variables. This test similarly significantly rejects the hypothesis of equal parameters.

Following Earnhart (2001), we examine whether certain subsets of parameters might be compatible in two data sets, although not all common parameters are compatible. Therefore, we separate travel cost of RP data and SP data in Models 5 to 7. However, all models strongly reject the test that the RP and SP data contain equal scaled common parameters.

Table 2-7 Different model specifications for combining RP and SP data

Model	Common variables	Number of common variables	likelihood ratio test
1	Travel Cost Log(beach length)	2	-2*(-117773.3-(-115617.1-2126.0))=60.3, Reject
2	Travel Cost Beach length dummies	5	-2*(-105128.2-(-102968.3-2112.6))=94.5, Reject
3	Travel Cost Log(beach length), Lake dummies	5	-2*(-106340.3-(-104106.5-2126.0))= 215.7, Reject
4	Travel Cost Beach length dummies Lake dummies	8	-2*(-105432.2-(-103196.6-2112.6))= 245.9, Reject
5	Beach length dummies	4	-2*(-105121.1-(-102968.3-2112.6))= 80.3, Reject
6	Log(beach length) Lake dummies	4	-2*(-106273.3-(-104106.5-2126.0))= 81.7, Reject
7	Beach length dummies Lake dummies	7	-2*(-105435.0-(-103196.6-2112.6))= 251.6, Reject

Current model specifications have rejected the scaling approach outlined above for combining the RP and SP data. An alternative strategy for combining RP and SP data is the calibration of SP to RP approach (Von Haefen and Phaneuf, 2008). This approach mainly relies on RP data, and uses the SP data to fill in the parameter estimates of interest that are missing in RP data, which in our case are the water quality attributes. Some reasons to use the calibration of SP to RP approach are that the RP data has much less variance than SP data and the SP data might suffer hypothetical bias.

In the approach of Von Haefen and Phaneuf (2008), in response to a rejection of the common parameter test, the scaling parameter was not estimated from the joint log likelihood function, but instead was calibrated as the ratio of parameters in the separate RP and SP models. In our case, the scaling parameter is calibrated as the ratio of beach length parameters in the RP and SP models.

$$\theta^c = \beta_l^{RP}/\beta_l^{SP}$$

In our study, the ratio is 0.064 divided by 0.164, which means the scaling parameter is 0.39. Using the calibrated scaling parameter to rescale the SP estimates of water quality attributes provides the parameters of the calibrated joint model.

6. Welfare Measures

6.1 Welfare Calculation Method

Once we get the calibrated scaling parameters from the calibration approach, we can use the calibrated joint model to measure the change in consumer surplus in response to a particular policy. Specifically, the indirect utility for calibrated joint model is:

$$
\begin{aligned}
V_{jlt}{}^{c} = {} & \beta_{tc}{}^{RP} \cdot travel\ cost_{jl} + \beta_{l}{}^{RP} \cdot \log\left(beach\ length_{jl}\right) + \omega_{t}{}^{RP} \cdot temperature_{jlt} + \omega_{cd}{}^{RP} \\
& \cdot closure\ days\ of\ 2010_{jl} + \omega_{r}{}^{RP} \cdot regional\ dummies_{jl} + \theta^{c}(\delta_{aw}{}^{SP} \\
& \cdot algae\ water\ dummies_{jt} + \delta_{as}{}^{SP} \cdot algae\ shore\ dummies_{jt} + \delta_{bt}{}^{SP} \\
& \cdot bacteria\ testing\ dummies_{jt})
\end{aligned}
$$

for beach alternative $j \in \{1, 2, \dots, 451\}$, choice occasion $t \in \{1, 2, \dots, 126\}$. To simplify the notation for welfare calculation, we use the abbreviations for dummy variables listed in Table 2-8.

Table 2-8 Abbreviations for dummy variables

Variable name	Abbreviation	Variable Definition	Attribute Levels
regional dummies	RD	The region of the beach located (base: Upper Peninsula)	LP Northeast LP Mid-East LP Southeast LP Northwest LP Mid-West LP Southwest
algae water dummies	AW	Algae in the water (base: high)	None Low Moderate
algae shore dummies	AS	Algae on the shore (base: high)	None Low Moderate
bacteria testing dummies	BT	Testing water for bacteria (base: Daily)	Never Monthly Weekly

To construct the status quo of the water quality for the Great Lakes beaches, we rely on the RP data. Under the status quo situation, assume the indirect utility for individual *n* who takes a trip to beach *j* at Lake *l* at the choice occasion *t* is:

$$\hat{V}_{jl,nt}{}^{0} = \hat{\beta}_{tc}{}^{RP} \cdot travel\ cost_{jl,n}{}^{0} + \hat{\beta}_{l}{}^{RP} \cdot \log\left(beach\ length_{jl,n}{}^{0}\right) + \widehat{\omega}_{t}{}^{RP} \cdot temperature_{jl,nt}{}^{0} + \widehat{\omega}_{cd}{}^{RP} \cdot closure\ days\ of\ 2010_{jl,n}{}^{0} + \widehat{\omega}_{r}{}^{RP} \cdot RD_{jl,n}{}^{0}$$

Specifically, the regional dummies *RD* are the regional average effects that account for all unidentified factors, which include water quality attributes. To separate the regional dummies, we further define the indirect utility as

$$\hat{V}_{jl,nt}{}^{0} = \tilde{V}_{jl,nt}{}^{0} + \widehat{\omega}_{r}{}^{RP} \cdot RD_{jl,n}{}^{0} \qquad (1)$$

where

$$\tilde{V}_{jl,nt}{}^{0} = \hat{\beta}_{tc}{}^{RP} \cdot travel\ cost_{jl,n}{}^{0} + \hat{\beta}_{l}{}^{RP} \cdot \log\left(beach\ length_{jl,n}{}^{0}\right) + \widehat{\omega}_{t}{}^{RP} \cdot temperature_{jl,nt}{}^{0} + \widehat{\omega}_{cd}{}^{RP} \cdot closure\ days\ of\ 2010_{jl,n}{}^{0}.$$

When we take the water quality attributes into the calibrated indirect utility, the baseline effects of the water quality attributes from SP data need to be netted out of the regional dummies. More formally, at region *r*, the original regional average effects are the sum of the regional water quality effects and the other regional effects:

$$\underbrace{\widehat{\omega}_{r}{}^{RP} \cdot RD_{jl,n}{}^{0}}_{\text{regional average effects}} = \underbrace{\omega_{r}{}^{remain} \cdot RD_{jl,n}{}^{0}}_{\text{the remainder}} + \underbrace{\hat{\theta}^{c}\ \left(\hat{\delta}_{aw}{}^{SP} \cdot AW_{r,n}{}^{0} + \hat{\delta}_{as}{}^{SP} \cdot AS_{r,n}{}^{0} + \hat{\delta}_{bt}{}^{SP} \cdot BT_{r,n}{}^{0}\right)}_{\text{regional water quality effects}} \qquad (2)$$

By inserting equation (2) into equation (1), we get the indirect utility with water quality attributes at the status quo point as

$$\hat{V}_{jl,nt}{}^{0} = \tilde{V}_{jl,nt}{}^{0} + \widehat{\omega}_{r}{}^{RP} \cdot RD_{jl,n}{}^{0}$$

$$= \tilde{V}_{jl,nt}{}^{0} + \omega_{r}{}^{remain} \cdot RD_{jl,n}{}^{0} + \hat{\theta}^{c}\left(\hat{\delta}_{aw}{}^{SP} \cdot AW_{r,n}{}^{0} + \hat{\delta}_{as}{}^{SP} \cdot AS_{r,n}{}^{0} + \hat{\delta}_{bt}{}^{SP} \cdot BT_{r,n}{}^{0}\right) \qquad (3)$$

The indirect utility for an individual who does *not* take a trip is:

$$\hat{V}_{No} = \hat{\gamma}_{male} \cdot male + \hat{\gamma}_{age} \cdot age + \hat{\gamma}_{white} \cdot white + \hat{\gamma}_{edu} \cdot edu + \hat{\gamma}_{fulltime} \cdot Fulltime + \hat{\gamma}_{retire} \cdot Retire + \hat{\gamma}_{under17} \cdot under17 + constant$$

Then, the expected maximum utility for each choice occasion *t*, or the inclusive value each individual *n* can obtain, is:

$$\widehat{IV}_{jl,nt}{}^{0}(status\ quo) = \ ln\left\{\left[\sum_{k\in L_G}\left[\sum_{i\in J_{kG}} exp\left(\frac{1}{\lambda}\hat{V}_{jl,nt}{}^{0}\right)\right]^{\frac{\lambda}{\rho}}\right]^{\rho} + \exp(\hat{V}_{No})\right\}$$

Now consider a change of water quality at one or more regions, for instance, change the algae level in the water. Assume that $AW_{r,n}{}^{0}$ represents the algae level in the water at region *r* for person *n* without an improvement and assume that $AW_{r,n}{}^{*}$ represents algae level in the water with an improvement. All other site characteristics remain the same, only the algae level in the water at region *r* has changed between the two states of the world. With the change in the water quality, the indirect utility for individual *n* for a trip to beach *j* at Lake *l* at choice occasion *t* is:

$$\hat{V}_{jl,nt}{}^{*} = \tilde{V}_{jl,nt}{}^{0} + \omega_{r}{}^{remain} \cdot RD_{jl,n}{}^{0} + \hat{\theta}^{c}\left(\hat{\delta}_{aw}{}^{SP} \cdot AW_{r,n}{}^{*} + \hat{\delta}_{as}{}^{SP} \cdot AS_{r,n}{}^{0} + \hat{\delta}_{bt}{}^{SP} \cdot BT_{r,n}{}^{0}\right)$$

$$= \tilde{V}_{jl,nt}{}^{0} + \omega_r{}^{remain} \cdot RD_{jl,n}{}^{0}$$

$$+ \hat{\theta}^{c} \left(\hat{\delta}_{aw}{}^{SP} \cdot AW_{r,n}{}^{0} + \hat{\delta}_{as}{}^{SP} \cdot AS_{r,n}{}^{0} + \hat{\delta}_{bt}{}^{SP} \cdot BT_{r,n}{}^{0} - \hat{\delta}_{aw}{}^{SP} \cdot AW_{r,n}{}^{0} \right.$$

$$\left. + \hat{\delta}_{aw}{}^{SP} \cdot AW_{r,n}{}^{*} \right)$$

$$= \tilde{V}_{jl,nt}{}^{0} + \widehat{\omega}_r{}^{RP} \cdot RD_{jl,n}{}^{0} + \hat{\theta}^{c} \left(\hat{\delta}_{aw}{}^{SP} \cdot \left(AW_{r,n}{}^{*} - AW_{r,n}{}^{0} \right) \right)$$

$$= \hat{V}_{jl,nt}{}^{0} + \hat{\theta}^{c} \left(\hat{\delta}_{aw}{}^{SP} \cdot \left(AW_{r,n}{}^{*} - AW_{r,n}{}^{0} \right) \right)$$

$$= \hat{V}_{jl,nt}{}^{0} + \hat{\theta}^{c} \left(\hat{\delta}_{aw}{}^{SP} \cdot \Delta AW_{r,n} \right)$$

With the change in the water quality, the expected maximum utility for each choice occasion *t* for each individual *n* is:

$$\widehat{IV}_{jl,nt}{}^{*}(scenario) = \ ln \left\{ \left[\sum_{k \in L_G} \left[\sum_{i \in J_{kG}} exp\left(\frac{1}{\lambda} \hat{V}_{jl,nt}{}^{*} \right) \right]^{\frac{\lambda}{\rho}} \right]^{\rho} + \exp(\widehat{V_{No}}) \right\}$$

As in Essay 1, the welfare change can be calculated as the change of expected maximum utility, i.e. the change of inclusive value, divided by the marginal utility of income.

$$cs_{nt} = \frac{\widehat{IV}_{jl,nt}{}^{*}(scenario) - \widehat{IV}_{jl,nt}{}^{0}(status\ quo)}{-\hat{\beta}_{tc}}$$

For individual *n,* the seasonal welfare change will be the sum of all consumer surplus changes in each choice occasion *t:*

$$CS_n = \sum_{t=1}^{T} cs_{nt}$$

The weighted average seasonal value per person is:

$$\overline{CS} = \frac{\sum_{n=1}^{N} w_n * CS_n}{\sum_{n=1}^{N} w_n}$$

For individual *n* at choice occasion *t*, the predicted total number of trips is:

$$\hat{Y}_{G,n} = \sum_{t=1}^{T} \hat{P}_{G,nt}{}^{0}$$

For individual *n* at choice occasion *t*, the predicted total number of taking trips to beach *j* at lake *l* is:

$$\hat{Y}_{jlG,n} = \sum_{t=1}^{T} \hat{P}_{jlG,nt}{}^{0}$$

If the water quality attributes changed, the change in predicted total number of trips is:

$$\Delta\hat{Y}_{G,n} = \sum_{t=1}^{T} \hat{P}_{G,nt}{}^{*}(scenario) - \sum_{t=1}^{T} \hat{P}_{G,nt}{}^{0}(status\ quo)$$

Similarly, the change in predicted total number of trips to beach *j* at Lake *l* is:

$$\Delta\hat{Y}_{jlG,n} = \sum_{t=1}^{T} \hat{P}_{jlG,nt}{}^{*}\ (scenario) - \sum_{t=1}^{T} \hat{P}_{jlG,nt}{}^{0}(status\ quo)$$

It is sometimes convenient to compare the seasonal value to other literature by normalizing the value to the change in trips. There are two ways to normalize the weighted average seasonal value per person to per trip units. One is to divide the value by the weighted average total trip change

$$\overline{\overline{CS}}_G = \frac{\overline{CS}}{\Delta\overline{Y}_G} = \frac{(\sum_{n=1}^{N} w_n * CS_n)}{\sum_{n=1}^{N} w_n * \Delta\hat{Y}_{G,n}}$$

and another is to divide the value by the weighted average trip change to beach *j* on lake *l*.

$$\overline{\overline{CS}}_{jlG} = \frac{\overline{CS}}{\Delta\bar{Y}_{jlG}} = \frac{\sum_{n=1}^{N} w_n * CS_n}{\sum_{n=1}^{N} w_n * \Delta\hat{Y}_{jlG,n}}$$

6.2 Welfare Results

As described above, for welfare measurement the status quo water quality level is partly captured by the regional effects from the RP part of our model and these status quo effects should be accounted for in any policy scenario. The status quo information for the water quality in each region was obtained from the 2011 Great Lakes Beach Sanitary Survey (EPA, 2011), which provided incomplete water quality information for 191 Great Lakes beaches. The surveyors went to sites and categorized the algae level in the water and on the shore to three levels: low, medium and high. There are 1,955 observations from Great Lakes Beach Sanitary Survey for 128 beaches in our choice set, of which 74 beaches have the information for algae levels in the water and 66 beaches have the information for algae levels on the shore. When we aggregated the water quality information at the regional level, information for the Northeast region is missing, so we assume the water quality in the Northeast is same as the Northwest. In the sanitary survey data testing for bacteria rarely happened since it is reported elsewhere. Therefore, the attribute of testing for bacteria is no longer included in water quality scenarios we examine here. Water quality is thus defined by algae level in the water and algae level on the shore as *low, medium,* or *high.* In our policy scenarios, when we refer to water quality change, we mean the algae level in the water and the algae level on the shore are simultaneously changed in the same direction.

Table 2-8 and Table 2-9 provide the baseline distribution of water quality across regions. The tables show that water quality in the LP Mid-East region and LP Southeast region is much lower than the water quality of the other regions based on the amounts of algae present. It reinforces our impression that, because of the algae problems, water quality of the Saginaw Bay, Lake Erie, and Lake St. Clair is worse than Lake Michigan.

Table 2-9 The baseline distribution of algae level in the water across region in 2011

	Low	Medium	High
LP Northeast	81.18%	18.04%	0.78%
LP Mid-East	52.43%	20.39%	27.18%
LP Southeast	57.79%	18.85%	23.36%
LP Northwest	81.18%	18.04%	0.78%
LP Mid-West	95.65%	2.17%	2.17%
LP Southwest	100.00%	0.00%	0.00%
Upper Peninsula	91.30%	6.52%	2.17%

Table 2-10 The baseline distribution of algae level on the shore across region in 2011

	Low	Medium	High
LP Northeast	86.99%	12.20%	0.81%
LP Mid-East	59.48%	20.69%	19.83%
LP Southeast	75.33%	22.91%	23.79%
LP Northwest	86.99%	12.20%	0.81%
LP Mid-West	100.00%	0.00%	0.00%
LP Southwest	100.00%	0.00%	0.00%
Upper Peninsula	94.05%	4.76%	1.19%

We consider two types of welfare scenarios using our calibrated joint model. The first scenario assumes that water quality at half of the sites in a region is improved *up* by one level. Simply put, half of Great Lakes beaches in a region with the high algae level are improved to the medium level and half of beaches in a region with the medium algae level are improved to the low level. Take Northeast region as an example, under the first scenario, high algae level in the water/on the shore becomes half of the baseline value of the low level, which means that 0.39% of Great Lakes beaches in the Northeast maintain a high algae level in the water and 0.4% of beaches maintain a high algae level on the shore. Medium algae level in the water/on the shore turns out to be half of the sum of baseline values of the low level and the medium level, which means 9.41% of beaches in the Northeast attain a medium algae level in the water and 6.51% of beaches attain a medium algae level on the shore. Finally, 90.2% of Great Lakes beaches in the Northeast attain a low algae level in the water and 93.09% of beaches attain a low algae level on the shore. The same procedures are applied to the water quality of the other five regions under the first scenario.

The second scenario assumes that water quality is deteriorated by shifting half of the sites' water quality in a region *down* by one level. This is a significant change in water quality, because half of beaches with the low algae level are degraded to the medium level and half of beaches with the medium algae level are degraded to the high level. The distribution of algae levels moves in the opposite direction to the algae levels in the first scenario. In both types of scenarios the algae changes are made only within one region at a time, resulting in twelve total welfare scenarios (an improvement and decrement to quality in each of six regions).

Table 2-11 displays the predicted trips and welfare estimates from the first scenario of water quality improvement. If we improve half of Great Lakes beaches' water quality in a region *up* by one level, compared to the trips taken at status quo, the trips increases by 33.62% for Middle-East region (Huron South) and 20.49% for Southeast region (St. Clair and Erie).[10] Trips increase slightly for Huron North and Lake Michigan. The intuition behind this is that the baseline algae levels in Huron South, St. Clair, and Erie are higher than those in Huron North and Lake Michigan. Once we increase the water quality, the utility of a person is increasing as the algae level decreases. Therefore, improving water quality leads to more utility increase for beaches with initially higher algae level in Huron South, St. Clair, and Erie than beaches with initially lower algae level in Huron North and Lake Michigan. In particular, trips to Southwest region never change, because the baseline water quality in the Southwest region was already at the highest level.

Under the water quality improvement scenario, the seasonal welfare benefits to beachgoers are larger for Huron South, St. Clair, and Erie as well. St. Clair and Erie generate the largest seasonal welfare gains, with $9.92 in seasonal value obtained for an average Michigan beachgoer. When normalized by the site trip change, the seasonal value per person per trip is $50.73. Although Huron South has the second highest seasonal value per person at $4.9, it has a relatively small

[10] Again, bear in mind that the 12 policy scenarios were run separately, so here we are comparing separate scenarios and are not referring to site substitution patterns within a scenario.

number of trips, so the seasonal value per person per trip turns out to be the second lowest at $33.36 when normalizing by the site trip change. South Lake Michigan has zero seasonal value since the water quality improvement does not affect this region at all.

To calculate the population level welfare, we follow the approach of in Essay 1 to aggregate the weighted average seasonal value at the individual level to the entire population of beachgoers living in the Lower Peninsula. The population number of beachgoers is derived from the participation rate of beach recreation, which is 58.01%, multiplied by 7,289,085 Michigan adults living in the Lower Peninsula. When aggregated at the population level, 0.83 million more trips were taken to Lake Erie and Lake St. Clair due to improving half of Great Lakes beaches' water quality in a region *up* by one level. Improvements at Lake St. Clair and Lake Erie result in $41.94 million in welfare gains by all Michigan beachgoers living in the Lower Peninsula. Again, welfare gains from South Michigan were zero because it had the highest water quality at status quo.

By contrast, if we degrade half of Great Lakes beaches' water quality in a region *down* one level, trips decrease dramatically and welfare loss turns out to be significant. Table 2-12 displays the predicted trips and welfare estimates from the second scenario of the water quality deterioration. Compared to the trips taken at status quo, all regions lose trips and the magnitude of decreased trips ranges from 24.09% to 32.66% across the six regions. When aggregated at the state level, 1.76 million trips are lost in the Northwest region due to degrading half of Great Lakes beaches' water quality *down* by one level. Mid-west region loses 1.75 million trips, followed by Southwest region losing 1.04 million trips. Mid-East region loses 0.6 million trips, which is the least trip loss among the six regions. The range of trip loss indicates that the water quality degradation impacts Lake Michigan most and Huron south least.

Under the water quality deterioration scenario, Michigan North has the largest seasonal welfare losses to beachgoers, with welfare losses from the Northwest region at $18.86 per person and from the Middle-west region at $16.81 per person. When normalized by the site trip change, St. Clair and Erie incur the highest seasonal welfare losses, with the seasonal value per person per trip at $48.41, followed by Lake Michigan ranging from $37.58 to $45.23 per person per trip. When aggregated at the state level, North Michigan loses $79.77 million by all Michigan beachgoers living in the Lower Peninsula from the water quality degradation. South Huron incurs the least welfare losses at $18.96 million. Finally, Lake St. Clair and Lake Erie incur $48.02 million welfare losses.

Table 2-11 Estimated trips and welfare measures of shifting half of sites' water quality up by one level in a region in 2011 dollars

		Number of Trips	Number of Site Trips Change	% Changes in Trips	Seasonal Value	Season/Total Trip Change	Season/Site Trip Change
Take Half of Sites' Algae in the Water & Algae on the Shore *up* by one Level	LP Northeast	0.68	0.03	4.96%	1.21	92.34	37.77
	LP Mid-East	0.58	0.15	33.62%	4.90	90.79	33.36
	LP Southeast	1.15	0.20	20.49%	9.92	89.98	50.73
	LP Northwest	1.62	0.06	4.05%	2.91	94.54	46.07
	LP Mid-West	1.74	0.02	1.21%	0.88	92.74	42.40
	LP Southwest	0.97	0.00	0.00%	0.00	0.00	0.00

State level

		Number of Trips (Million)	Number of Site Trips Change (Million)	% Changes in Trips (Million)	Seasonal Value (Million)
Take Half of Sites' Algae in the Water & Algae on the Shore up by one Level	LP Northeast	2.872	0.136	4.96%	5.122
	LP Mid-East	2.468	0.621	33.62%	20.717
	LP Southeast	4.862	0.827	20.49%	41.937
	LP Northwest	6.857	0.267	4.05%	12.283
	LP Mid-West	7.357	0.088	1.21%	3.719
	LP Southwest	4.111	0.000	0.00%	0.000

Note: The table rows are for the 12 regional scenarios each run separately. Only changes within a region are shown and site substitution patterns for each scenario are omitted for brevity.

Table 2-12 Estimated trips and welfare measures of shifting half of sites' water quality *down* by one level in a region in 2011 dollars

Per Person

		Number of Trips	Number of Site Trips Change	% Changes in Trips	Seasonal Value	Season/Total Trip Change	Season/Site Trip Change
Take Half of Sites' Algae in the Water & Algae on the Shore *down* by one Level	LP Northeast	0.44	-0.21	-32.14%	-7.57	92.25	36.37
	LP Mid-East	0.29	-0.14	-32.66%	-4.49	90.68	31.44
	LP Southeast	0.72	-0.24	-24.58%	-11.36	89.74	48.41
	LP Northwest	1.14	-0.42	-26.74%	-18.86	94.26	45.26
	LP Mid-West	1.31	-0.41	-24.09%	-16.81	92.56	40.58
	LP Southwest	0.73	-0.25	-25.28%	-9.24	92.02	37.58

State level

		Number of Trips (Million)	Number of Site Trips Change (Million)	% Changes in Trips (Million)	Seasonal Value (Million)
Take Half of Sites' Algae in the Water & Algae on the Shore *down* by one Level	LP Northeast	1.857	-0.880	-32.14%	-31.986
	LP Mid-East	1.244	-0.603	-32.66%	-18.963
	LP Southeast	3.044	-0.992	-24.58%	-48.015
	LP Northwest	4.828	-1.763	-26.74%	-79.766
	LP Mid-West	5.518	-1.751	-24.09%	-71.076
	LP Southwest	3.071	-1.039	-25.28%	-39.050

Note: The table rows are for the 12 regional scenarios each run separately. Only changes within a region are shown and site substitution patterns for each scenario are omitted for brevity

7. Conclusion and Discussion

This paper investigated combining revealed and stated preference data to jointly estimate the monetary value of water quality attributes and their economic benefits to recreational beachgoers. To combine the trip data and choice experiment data from a 2011 Great Lakes Beach Survey, we first applied a scaling approach to jointly estimate the parameters of attributes in both RP and SP datasets under a unified RUM framework. Different model specifications for common preferences across the data types were tested. Common preference tests between the RP and SP data were consistently rejected. Our results provide empirical evidence that passing the hypothesis of equal common parameters is difficult when combining both RP and SP.

With some caveats, we then applied the calibration of SP to RP approach to measure the change in consumer surplus in response to two types of water quality scenarios. If we improve half of Great Lakes beaches' water quality in a region *up* by one level, compared to the trips taken at status quo, trips increase by 33.62% for Middle-East region (Huron South) and 20.49% for Northeast region (St. Clair and Erie). Trips increase slightly for Huron North and Lake Michigan. At the state level, we found 0.83 million more trips were taken to Lake Erie and Lake St. Clair. Improvements at Lake St. Clair and Lake Erie result in $41.94 million in welfare gains by all Michigan beachgoers living in the Lower Peninsula. By contrast, trip changes and welfare gains from South Michigan were zero because it had the highest water quality at status quo.

If we degrade half of Great Lakes beaches' water quality in a region *down* one level, compared to the trips taken at status quo, each region loses trips so dramatically that the magnitude of decreased trips ranging from 24.09% to 32.66% across the six regions. Northwest region lost most trips at 1.76 million. It also resulted in the lowest seasonal welfare losses at $79.77 million to all Michigan beachgoers living in the Lower Peninsula. The South Huron scenario incurs the largest welfare losses at $518.96 million. Distributions of trip losses and welfare losses across the six regions indicate that the water quality degradation impacts Lake Michigan most, Huron south least.

We note that even if one rejects the consistency test and thus the data sets cannot be jointly estimated, a simple calibration approach still provides a way to combine the data sets. However, the estimated changes in consumer surplus could still be biased, even if they intuitively make sense. Finally, this essay provided the empirical evidence that the scaling approach is not sufficient to account for differences in the amount of unexplained variance when using RP and SP data together in some applications. Therefore, more empirical strategies should be proposed and implemented in the future.

ESSAY 3 Estimating Spending for Trips to Great Lakes Beaches in Michigan

1. Introduction

1.1 Beach Recreation is Important to the Michigan Economy

With 3,126 miles of Great Lakes shoreline[11], Michigan has over 500 beaches on the shoreline of the Great Lakes. Each year millions of visitors from all over the state visit Great Lakes beaches. During their visits, they may spend money on transportation, food, beverages, and lodging. This spending will contribute to local economic development because the recreation demand induces consumption at local gas stations, grocery stores, restaurants, and hotels.

Despite their popularity among Michigan residents' recreational activities, Great Lakes beaches face threats from a combination of factors that include bacterial contaminants, invasive species, algal growth, harmful algae blooms, shoreline development and land uses, and climate change. For instance, Dorfman and Haren (2009) indicated that water quality samples from the Great Lakes region had the highest percentage of *E. Coli* exceeding the EPA's standards in the Nation. Shorebirds have been killed by a toxin-producing bacterium (*Botulism*) found in food,[12] leading to dead shorebirds on some Great Lakes beaches. Decaying algae and some invasive species, such as zebra and quagga mussels, have accumulated and fouled some Great Lakes beaches. Shoreline development and land use can decrease the opportunities of outdoor recreation and degrade the water quality by increasing the non-point source pollution and point source pollution (USEPA, 2009). Finally, climate change will have uncertain effects on Great Lakes water levels, which affects beach width and potentially erosion. For example, Hartmann (1990) raised

[11] Source: "Michigan's State Facts", http://www.michigan.gov/kids/0,4600,7-247-49069-67959--,00.html
[12] http://www.environmentalhealthnews.org/ehs/news/2014/aug/mass-murder-by-botulism-scientists-exploring-surge-in-great-lakes2019-bird-deaths

concerns that rising lake levels caused by climate change could increase erosion threats to shorelines in Michigan. All of these threats pose challenges for beach recreation.

1.2 Spending Analysis and its Significance

Quantifying the contribution of beaches to the local economy can inform policy makers of the some of the importance of preserving and restoring beaches. Because there are limited funds for competing uses of many natural resources, policy makers need to evaluate preservation and restoration programs to justify funding decisions. Policy makers evaluating beach programs not only need to consider the costs and benefits but also the distributional implications of the program. Understanding the regional distribution of the recreational activity, however, requires measurement of the locations and economic impacts.

Visitor spending is an essential component of economic impact analysis (Wilton and Nickerson, 2006). An economic impact analysis focused on beach recreation can help policy makers, more specifically, park and recreation administrators and planners, as well as the local community, evaluate potential beach development or protection programs.

1.3 Research Gaps in Studying Spending of Beach Recreation

Despite the importance of spending analyses, there is very limited information from prior studies on the spending of trips to freshwater beaches. In contrast, trip spending to ocean beaches has been investigated by many researchers. Dwight et al. (2012) surveyed 2,455 visitors at 14 southern California beaches and computed expenditures per person to be $72.31per trip in 2014 dollars. Nelsen et al (2007) collected 973 samples from a web survey and estimated the average expenditure per-person per-visit for surfers visiting Trestles Beach was $47.05 in 2014 dollars. King (2002) interviewed 283 groups at San Clemente's beaches in the summer of 2001 and found that expenditures specifically for beach recreation were $103.04 per person in 2014 dollars. King (1999) used spending data from a telephone survey in 1995 regarding 641 California residents' trips to beaches and found that the average expenditure per person on day trips was $21.4 and on overnight trips was $130.67 in 2014 dollars. Hanemann et al. (2004) found that per-person per-trip expenditures were $31.89 in 2014 dollars in a survey of beach-goers who took at least one trip in the summer of 2000 in southern California. Bell and Leeworthy (1986) found the annual average household expenditures on visiting saltwater beaches in Florida to be $1065 in 2014 dollars

Nevertheless, very few spending studies have specifically focused on Great Lakes, which have different characteristics from saltwater beaches. Murray, et al (2000) surveyed 1,587 visitors at 15 Lake Erie beaches and suggested that single day visitors to beaches on Lake Eire in Ohio spent $20 per trip in 1998. However, their study was applied only to beaches on Lake Erie and may not be representative of other areas of the Great Lakes. In Michigan, the National Park Service has provided visitor spending and economic impact reports[13]. However, these reports only include one national park with a beach: Sleeping Bear Dunes National Lakeshore. For instance, in 2009, Cook (2009) found direct spending of visitors to Sleeping Bear Dunes National Lakeshore to be $103.5 million in the region within a one-hour drive of the park. This revenue was estimated to support 1,279 jobs in the region, while the indirect labor income and induced value added were estimated to be $90.6 million. Still, there are hundreds more Great Lakes beaches in Michigan in need of economic impact studies to demonstrate their importance.

Trip visitation data is an essential part of economic impact studies and is useful information for recreation planning and management. However, visitation data are not always available and

[13] Source: National Park Service: http://www.nps.gov/state/mi/index.htm

can be difficult to collect. National parks and national forests have the capability to maintain large and diverse trip datasets, while some states or regions do not. For instance, national parks and national forests in some areas track visitation by using automated traffic counters with calibration (Watson et al., 2000), but this approach is not typically used for Great Lakes beach recreation, in part due to the costs. In California, the Lifeguard system provides beach visitation data based on head counts (Dwight et al, 2007). In Florida, a statewide survey was implemented every 12 years since 1995 to visitors of the Florida Keys (English et al, 1996; Leeworthy et al, 2010).

By contrast, in Michigan, an official and publicly accessible record of beach visitation does not exist, so there is almost no accurate data on the number of trips taken to Great Lakes public beaches in Michigan every year. By using data on beachgoers' from a web survey, Chen (2013) was the first to apply a recreation demand model to value Great Lakes beaches. Essay 1 extended Chen's study by using both day trips and overnight trips data to value the Great Lakes beaches. In this study, we use the demand system based on Essay 1 to predict the regional variation of trips to Great Lake beaches.

When collecting survey data such as recreation spending, non-respondents always exist regardless of the survey methods. Therefore, the possibility of nonresponse biases deserves attention. The most common way of addressing nonresponse biases in economic impact studies is to compare the characteristics of respondents to those of the general population, but there is often no reason that the population that engages in recreation has the same characteristics as the general population. Another approach which can be a convenient and inexpensive, compares variables such as age and income from different surveys of the desired target population (Armstrong and Overton, 1977), often using *t*-tests or *chi-square* tests of differences in means between the respondents and non-respondents (Lee, 2001). However, the differences might arise from measurement errors associated with the different survey instruments rather than nonresponse bias. The third way is to apply econometric methods which have the advantage of addressing the self-selection bias, such as the Heckman model (sometimes referred to as a type II Tobit model). English (1997) found that an approach which does not correct for nonresponse bias overestimates visitor's spending per trip to Cumberland Island National Seashore by 15% and economic impacts to industry output by 10%. Leeworthy et al. (2001) used a type II Tobit model to correct the nonresponse bias for tourism impacts in the Florida Keys. Still, these methods are less popular in economic impact analysis mainly because of the complexity inherent in the model and the extra requirement of having non-respondents' information that was obtained either through screener surveys or follow-up surveys (Whitehead, 1991). More numerous applications can be found in environmental economics literature. For instance, Messonnier et al (2000) estimated the willingness to pay for aquatic plant management in Lake Guntersville, Alabama. By using the Heckman Model, they found the amount of non-fishers' willingness-to-pay was underestimated without correcting for nonresponse bias. Cameron et al (1999) demonstrated how to correct nonresponse bias by using sample members zip code information which alleviates the burden of collecting non-respondents' information.

1.4 Objectives of This Study

To address the above research gaps, there are two objectives of this study:

- *To estimate an expenditure function for trips to Great Lakes beaches by using the Heckman Model to control for possible nonresponse bias*. Unlike most economic impact literature, a Heckman model is used to address potential nonresponse bias problem in the spending data, to more accurately estimate visitors' spending.

- *To estimate regional variation of spending per trip per person to Great Lakes beaches in Michigan during a beach season.* This paper focuses specifically on Great Lakes beaches in Michigan, thus contributing to the spending studies by reporting the spending of beach recreation along the shoreline of Great Lakes in Michigan.

2. Methods

2.1 Spending Estimation: Heckman Model

To estimate spending, a spending survey is essential. Following Greene (2003, Chapter 19), assume the spending survey collects a random sample of *N* visitors or visitor groups on site. The intercepted samples are asked to provide some information *z*, such as zip code. In addition, they are invited to fill out a survey which asks questions about their party's expenditures (m) for this trip.

Let q_i^* be the latent propensity for intercepted visitor *i* to fill out a completed response to the survey:

$$q_i^* = z_i'\gamma + v_i, \qquad v_i \sim N(0,1)$$

The vector z_i includes individual attributes, such as the zip code of the primary residency. However, researchers cannot observe the visitors' propensity to respond, they can only observe the response or nonresponse outcome. Let $q_i = 1$ if the survey was completed and $q_i = 0$ if the visitor never logged in to fill out the survey, or if the survey was insufficient to be included in the analysis:

$$q_i = \begin{cases} 1, & if\ q_i^* > 0 \\ 0, & otherwise \end{cases}$$

Assume that visitor *i* reported an expenditure m_i^*, which is a linear function of the variables in x_i :

$$m_i^* = x_i'\beta + \varepsilon_i, \qquad \varepsilon_i \sim N(0, \sigma^2)$$

The vector x_i includes individual attributes intended to explain systematic variations in spending, including the distance from primary residency to the site, trip types, and demographic information.

However, the spending function can be constructed only for those who responded. Let the actual spending for visitor i as m_i:

$$m_i = \begin{cases} m_i^*, & if\ q_i = 1 \\ unobserved, & otherwise \end{cases}$$

Assume

$$(v_i, \varepsilon_i) \sim bivariate\ normal\ with\ correlation\ \rho$$

Then the conditional mean for m_i given the reported spending (see, Greene 2003 p. 854) is:

$$E(m_i | q_i = 1) = x_i'\beta + (\rho\sigma)\frac{\phi(z_i'\gamma)}{\Phi(z_i'\gamma)}$$

The above clarifies that the regression using *m* needs to control for the second term, which is a non-linear function of the response selection propensities. The Heckman estimation procedure appropriately controls for these terms.

According to English (1997) and Stynes (1997), spending surveys should elicit information of the party/group's spending and the size of the party/group. Therefore the estimate of per person per trip spending is:

$$\bar{m} = \frac{E(m | q = 1)}{E(party\ size)}$$

2.2 Trip Prediction

Following the same notation in essay 1, for individual *n*, in the given beach season, the predicted number of trips taken by person *n* to beach *j* on lake *l* is the sum over days of the probability of going to beach *j* on any day:

$$\hat{Y}_{jlG,n} = \sum_{t=1}^{T} \hat{P}_{jlG,nt}$$

where $\hat{P}_{jlG,nt}$ is the unconditional probability of individual *n* taking a trip to beach *j* at choice occasion *t*. If we divide the beach season into 4 months, we can compare the trip visitation by each month. For each month *m*, the predicted total number of trips taken is:

$$\hat{Y}_{jlG,nm} = \sum_{t=1}^{T_m} \hat{P}_{jlG,nt}$$

Adding each month's visitation per person gives the total number of trips to beach *j* for the whole beach season:

$$\hat{Y}_{jlG,n} = \sum_{m=1}^{4} \sum_{t=1}^{T_m} \hat{P}_{jlG,nt}$$

Taking the weighted average across every individual, the average number of trips taken to beach *j* during the beach season per person is:

$$\overline{Y_{JlG}} = \frac{\sum_{n=1}^{N} w_n * \hat{Y}_{jlG,n}}{\sum_{n=1}^{N} w_n}$$

As in Essay1, we divided the Lower Peninsula of Michigan into 7 regions, *R={Huron North, Huron South, St. Clair, Erie, Michigan North, Michigan Center, Michigan South }*. For region *r* in *R*, taking the weighted sum of trips to beaches in the region gives the predicted total number of trips per person in that region.

$$\overline{Y_{G,r}} = \frac{\sum_{n=1}^{N} w_n \ * (\sum_{j=1}^{R_r} \hat{Y}_{jlG,n})}{\sum_{n=1}^{N} w_n}$$

where R_r is the total number of beaches in the region r.

Because the demand model only considers the users and potential users of beaches, and does not consider the non-participants of beach recreation, we have to take the participation rate into consideration. Assume the participation rate of beach recreation is *bpr*, then, for region *r*, the total visitation of trips by Michigan residents during the beach season is:

$$Y_r = bpr * Total\ population * \overline{Y_{G,r}}$$

For the state, the total visitation of trips by Michigan residents living in the Lower Peninsula during the beach season is:

$$Y = bpr * Total\ population * \sum_{r \in R} \overline{Y_{G,r}} = bpr * Total\ population * \overline{Y_G}$$

2.3 Estimation Procedures

As a summary of section 2, Figure 2-1 provides an overview of the approach to predict trips and estimate an average beachgoer's spending of visiting Great Lakes public beaches. Trip estimation and spending estimation are the top two parts in the Figure 2-1. Following the flows in each component leads to spending for beach visitation to 451 beaches.

To calculate the regional variation of spending on beach visitation, let the predicted spending for person *n* to beach *j* be $\widehat{m}_{j,n}$. Then the total spending for person *n* to beach *j* in a beach season is

$$\widehat{M}_{j,n} = \widehat{Y}_{jlG,n} * \widehat{m}_{j,n}$$

Taking the weighted sum of the total spending per person to a specific beach across *all* the beaches in region *r* gives the weighted average total spending per person to beaches in that region. That is to say, the predicted total spending per person per season to beaches in a region is:

$$\bar{M}^r = \frac{\sum_{n=1}^{N} w_n \ * (\sum_{j=1}^{R_r} \widehat{M}_{j,n})}{\sum_{n=1}^{N} w_n}$$

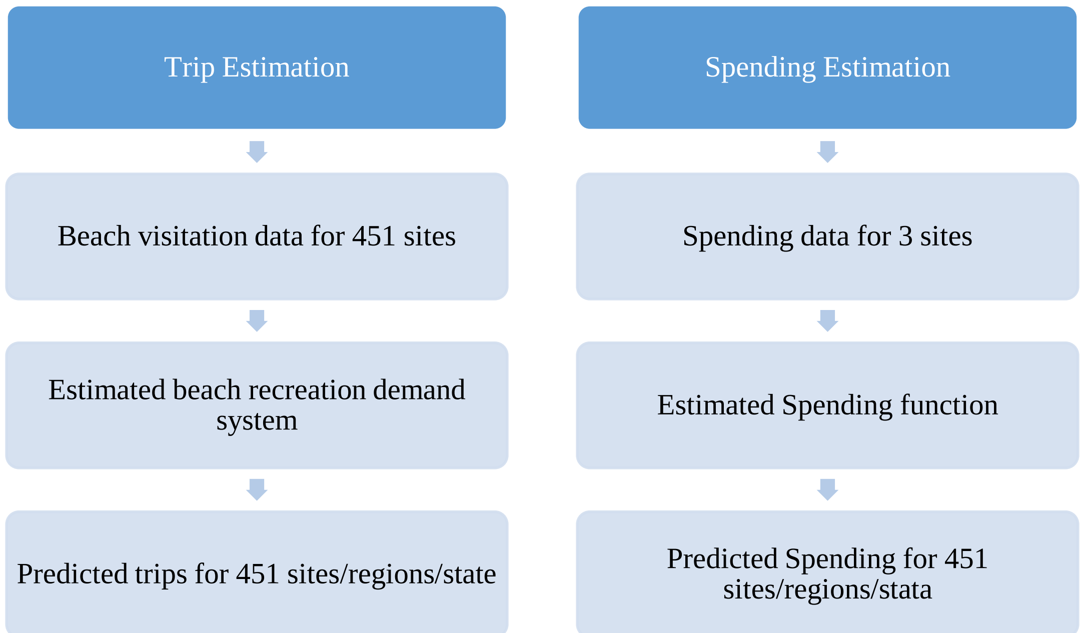


Figure 3-1 Detailed approach to estimate spending for visits to Great Lakes public beaches

3. Survey and Data

3.1 Surveys

Two surveys are applied in this essay. The first one is the Great Lakes Beaches Survey, which is used for the purpose of trip prediction. The second survey is the Beach visitor spending survey, which is used for the spending estimation.

Great Lakes Beaches Survey is a two-stage survey developed by Lupi, Kaplowitz, Chen, and Weicksel in 2011 as described in detail in Essay 1 (see also Chen 2013 and Weicksel 2012). In this essay, we use the trip prediction based on the demand systems in Essay 1 and the demographic information from the Great Lakes beaches survey.

The Beach Visitor Spending Survey first involves on-site recruitment of subjects by intercepting beachgoers and distributing an invitation letter with a unique web address to access a web-based survey. The recruitment of subjects took place in three public beaches in Michigan in the summer of 2014, specifically, the Bay City Recreation State Park (Lake Huron), the Grand Haven State Park (Lake Michigan), and the Metropolitan Beach Metro Park (Lake St. Clair). The interviewer would ask for the individual's zip code, and contact information including, if possible, an email address and a mailing address. If the person refused to give the email address, they were asked to provide a mailing address; if the person still did not want to provide any contact addresses, then only the invitation letter was given.

If the intercepted beachgoers did not have access to the Internet, a mail survey was sent to their residency. To reduce recall bias, expenditure surveys should be conducted as soon after the recreational event as possible (Champ & Bishop, 1996). Therefore, three waves of email reminders were subsequently sent within two weeks after the date of each on-site sampling. The fourth wave of email reminder was sent after one month. Because some intercepted beachgoers left a resident address instead of an email address, we sent three waves of mail reminders in one month to those who gave residential address information. Detailed survey implementation is described in the Appendix H, along with copies of all correspondence letters and e-mails.

The survey has two parts: one asked people's itemized expenditures, and another collected demographic information. Survey instruments are listed in Appendix I.

3.2 Data

During the 2014 summer period, 336 groups were intercepted at three beaches on the Great Lakes and invitation letters were successfully handed to 334 groups. By the end of survey period, we received 150 fully complete responses out of 170 overall responses. After replacing missing demographic information with mean values of the samples, we obtained 7 more useable responses, which leads to 157 effective observations, with a response rate of 47%. Detailed statistics of response rate are listed in Appendix J.

Following Stynes (1997) and English (1997), a beachgoer's spending from the visitor survey is measured for the party. Party is defined in the survey as the persons arriving in the same vehicle. Therefore, party size is very important when transforming the spending per party to spending per person.

From Table 3-1, Michigan beachgoers spend much more on the overnight trips than on day trips to beaches. The average spending per party for day trips was \$32.38 in 2014 dollars. Beachgoers for overnight trips and stopover visits spent \$718.5 and \$236 respectively on average per party. However, the variances are also large, which means the spending varies substantially across parties.

Table 3-1 The average spending per Party for Michigan beachgoers

Trip type	Mean ($)	Standard Deviation	Frequency
Day trip	32.38	28.49	104
Overnight trip	718.50	475.38	30
Stop over	236.00	366.13	6
Total	188.13	362.67	140

Dividing the spending per party by party size gives the average spending per person. Compared to other spending studies such as Murray et al (2000) with a range of $18 to $24 in 1998 dollars, our day trip spending per person is comparatively lower with $15.57. One reason is that we differentiate the beachgoers within the state and outside of the state, and in these tables we only include the beachgoers who are Michigan residents.

Table 3-2 The average spending per person for Michigan beachgoers

Trip type	Mean ($)	Standard Deviation	Frequency
Day trip	15.57	17.96	104
Overnight trip	269.65	228.78	30
Stop over	94.92	113.20	6
Total	73.41	149.85	140

Table 3-3 presents the average spending per person across each beach site. People tended to spend more at Grand Haven, less at Saginaw Bay, and the least at St. Clair Metro Park.

Table 3-3 The average spending per person for each site for Michigan beachgoers

Site	Mean ($)	Standard Deviation	Frequency
Grand Haven	102.33	171.39	90
Saginaw Bay	39.62	128.01	18
St. Clair Metro Park	11.11	16.35	32
Total	73.41	149.85	140

Essay 1 estimated of the value of both day trips and overnight trips to Great Lakes beaches by using a repeated three-level nested logit model. Among the 2,573 observations, 1,894 individuals took at least one trip to Great Lakes beaches from Memorial Day weekend to September 30, 2011. The trip data consists of self-reported trips to Great Lakes beaches. After matching the reported beaches to the Michigan DEQ beach database, the choice set for each individual is comprised of 451 beaches. There are 643 people who had taken trips to Great Lakes beaches before but didn't take any trip during the indicated season, they are also include in this study. Based on the demand systems of Essay 1, predicted number of trips were obtained for each site and then aggregated to each region in the beach season following the methods for trip predictions outlined above.

4. Results

4.1 Spending Estimation results

The Heckman model described above was applied to both day trips and overnight trips data to estimate a spending equation. In the selection equation (response/nonresponse) stage of the Heckman model, the explanatory variables rely on our knowledge of the intercepted parties' zip codes and the beach they were intercepted at. The zip codes are used to derive census variables for the neighborhoods of the parties, The 2010 Census data at ZCTA level (ZIP Code Tabulation Areas), which is census data aggregated to the level of zip codes, provides measures of social-demographic characteristics of non-respondents. The zip codes are also used to derive travel distances. The data for the explanatory variables for the spending equation are presented in Table 3-6, which shows the variable values for the 314 cases from the full intercepted sample and for the 157 respondents with day trips and overnight trips.

Table 3-4 Statistic summary of the explanatory variables from census data at ZCTA level for the entire sample (N=314) and for the 157 respondents[14] which are used in the selection equation.

Variable	Definition	Obs	Mean	Std.	Min	Max
distance	One way driving distance	314	72.48	162.92	0	1928
		157	97.32	220.99	0	1928
population	Population in the zip code (million)	314	0.03	0.01	0.0007	0.0531
		157	0.03	0.01	0.0016	0.0531
pwhite	Percentage of white in zip code	314	86.87	11.50	7.80	97.80
		157	86.61	10.89	43.00	97.80
medhhincome	Medium Household income (Thousand dollar)	314	50.80	13.21	5.00	106.83
		157	51.66	12.96	5.00	97.44
graduate	Percentage of people with graduate degrees in zip code	314	8.86	6.57	0.50	40.60
		157	9.30	6.53	0.50	40.60
GrandHaven	Was interviewed at Grand Haven	314	0.61	0.49	0	1
		157	0.67	0.47	0	1

In the second stage of the Heckman model, the total spending for overnight trips only include the expenditures spent with 35 miles of the destination, for the purpose of measuring the impacts specifically for that local region (Stynes, 1997). By intuition, we expect that people would spend more if they drive longer distance with larger party size. In terms of demographic variables, we expect that people with higher income tend to spend more. The total party expenditure is

[14] For each variable, the first row is for the whole sample which include respondents and non-respondents. The second row is for respondents with day trips and overnight trips.

regressed on travel distance, party size, and demographic characteristics. The data for the explanatory variables for the spending equation are presented in Table 3-7, which shows the descriptive statistics for the independent variables in the expenditure equation.

Table 3-5 Statistic summary of the explanatory variables in spending equation

Variable	Definition	Obs	Mean	Std.Dev.	Min	Max
spending	Total spending per party (dollar)	157	200.69	379.78	0.00	1985
distance	One way driving distance	157	97.32	220.99	0.00	1928
size	Party size	157	2.76	1.52	1	8
male	Male=1, Female=0	157	0.23	0.42	0	1
age	Age	157	45.05	13.82	19	81
white	Dummy=1, if white	157	0.95	0.22	0	1
eduyear	Years of education	157	15.30	2.36	12	19
income	Income (Thousand dollar)	157	96.07	65.58	12.50	300
fulltime	Dummy=1, if employed full-time	157	0.53	0.50	0	1
retire	Dummy=1, if retired	157	0.69	0.46	0	1
couple	Dummy=1, if couple only	157	0.34	0.48	0	1
child17	Dummy=1, if with children under 17	157	0.13	0.34	0	1

The estimation results from the full Heckman model are presented in Table 3-8. The estimation procedure controls for possible selection bias due to response/nonresponse. The results show that the correlation of the two equations (rho) is different from zero at 10% statistical significance level, so there was evidence of a sample selection problem at the 10% but not 5%

level of significance. The estimated parameter on travel distance has a negative sign and is statistically significant at 1% level, which means people have more propensity to respond to the spending survey if they travel from further away. This result makes sense since their willingness to travel signals their interests in beach recreation, which in turn might make them more likely to respond the beach spending survey.

Table 3-6 Heckman model estimation results

	Variables	Estimates	Standard Errors	t-statistics
Spending	Log(distance)	67.64***	19.93	3.39
	size	12.96	19.15	0.68
	male	38.68	67.48	0.57
	age	5.02***	2.53	1.98
	white	93.33	119.59	0.78
	eduyear	-22.05**	11.67	-1.89
	income	1.19***	0.44	2.68
	fulltime	-24.69	63.58	-0.39
	retire	-84.44	110.32	-0.77
	couple	99.38	68.94	1.44
	child17	-12.83	61.66	-0.21
	constant	-76.50	247.65	-0.31
Respond	distance	0.003***	0.001	2.65
	population	-4.55	6.43	-0.71
	pwhite	-0.01	0.01	-1.41
	medhhincome	0.00	0.01	0.49
	graduate	0.00	0.01	0.31
	GrandHaven	0.15	0.16	0.96
	constant	0.45	0.62	0.74
rho		-0.518	0.169	
LR test of indep. eqns. (rho = 0): chi2(1) = 3.52 Prob > chi2 = 0.0607				

Note: *10% significance level; **5% significance level; *** 1% significance level

We tried 4 model specifications for the spending equation with 4 different combinations of distance and logarithm of distance in the spending equation and selection equation. The model in Table 3-8 had the best fit in terms of the value of log likelihood, AIC and BIC. Moreover, using logarithm of distance allows for a declining effect of distance which is advantageous when predicting for more distant sites in the 2011 Great Lakes Beaches Survey. In the spending equation, we found the logarithm of driving distance was statistically significant, and had a positive effect on spending. The driving distance matters a lot for spending, and its importance decreases as distance increases. All else equal, age, income, and education years have statistically significant effects on beach spending. Specifically, people with higher education tend to spend less on trips, while people who are older and with higher income tend to spend more on trips. Robustness checks for the model specification is provided in the Appendix F.

4.2 Spending Prediction

The estimated spending equation was applied to 2,537 beachgoers from the Great Lakes Beaches Survey. Because each beachgoer has 451 beach alternatives in the choice set, the sample for prediction has 1,144,187 observations, which derived from 2,537 people times 451 beaches. Since the demographic information collected in the Beach Visitor Spending Survey was by design exactly the same as that in the Great Lakes Beaches Survey, we had the demographic variables needed for spending predictions. The round-trip travel distance in the trip data was transformed to the one-way distance by dividing by two. For party size information, Great Lakes Beaches Survey also asked respondents the number of children and adults that they traveled with for a random selection of the total trips. For prediction, we used the weighted average party size across all trips in the data—2.94 for every observation's party size. Table 3-9 shows the statistical summary of the explanatory variables for spending prediction to the site in people's choice sets. All explanatory variables follow the same definition as those in Table 3-7. In the final spending prediction 6.14% of the cases are less than zero, which are then truncated at zero.

In Table 3-9, the average predicted spending per party across the 1,144,187 observations is $439.53, which seems much higher than the spending in studies in a similar regions (Stynes, 1998; 2004). However, the predicted spending reported in the table is for sites in the choice set and does not account yet for the probability that those sites are visited. Thus, the data in the table summarizes the spending predictions if the sites were to be visited and cannot be directly comparable to the spending that actually happened, because the predicted spending is conditional on the trips taken by a beachgoer. For instance, if a beachgoer has a near-zero probability of going to a specific beach, the predicted spending to that beach may be high but when weighted by his trips to the beach it will be near zero.

Table 3-7 Statistical summary of the explanatory variables using 2011 Great Lakes Beaches Survey and Predicted spending if a visit were to be taken to each of the 451 sites in the recreation demand model choice set

Variable	Obs	Mean	Std. Dev.	Min	Max
predicted spending[15] (per party)	1,144,187	439.53	134.70	0.00	940.08
distance	1,144,187	220.82	92.06	0.00	557.20
size	1,144,187	2.94	0.00	2.94	2.94
male	1,144,187	0.48	0.50	0.00	1.00
age	1,144,187	44.40	17.19	18	94
white	1,144,187	0.91	0.29	0	1
eduyear	1,144,187	14.82	2.38	10	19
income (1000s)	1,144,187	81.91	61.34	12.50	300.00
fulltime	1,144,187	0.52	0.50	0	1
retire	1,144,187	0.19	0.39	0	1
couple	1,144,187	0.34	0.47	0	1
child17	1,144,187	0.35	0.48	0	1

4.3 Trip Prediction

According to the Beach Activity Survey (see, Chen 2013), the share of people classified as potential beachgoers was 58.01%; that is, 58.01% of the Michigan adult population fall within the share of the population that was included in the trip demand model. Based on 2010 census data, the total number of adults living in the Lower Peninsula of Michigan is 7,289,085. The

[15] The reported value for the predicted spending is the estimated spending that would be made if a trip was taken to the sites (i.e., the figure is not yet weighted by the probabilities of visiting the sites).

unconditional probability of taking trips to 451 sites was obtained from Essay 1. After additional calculations listed in section 2.1, the predicted trips per person per season for the region and for the entire Lower Peninsula of the state can be obtained.

4.4 Total Spending by Region

This section provides the regional total spending to Great Lakes beaches. Table 3-10 displays the regional differences in the total spending of beach visitation per person per season. If we assume the trips taken by an average Michigan beachgoer during the beach season in 2011 maintains the same as in 2014, the total spending of an average beachgoer to Great Lakes beaches in one region ranges from $35.92 to $248.80 in 2014 dollars. Specifically, during a beach season, an average Michigan beachgoer spent the least on trips to Lake Erie at $35.92, followed by Lake St. Clair at $54.57. The average beachgoer spent the most on trips to Michigan Central at $248.80 per person per season, followed by Michigan North at $229.92 per person per season.

Table 3-8 Economic impacts of beach visitation in 2014 dollars per person per season

	Number of Trips (per person per season)	Total Spending by Region (per person per season)
Huron North	0.68	99.51
Huron South	0.69	96.55
St. Clair	0.42	54.57
Erie	0.27	35.92
Michigan North	1.59	229.92
Michigan Central	1.72	248.80
Michigan South	0.97	140.95

To calculate the state level economic spending, we aggregated the weighted average regional spending per person to all beachgoers living in the Lower Peninsula. Table 3-11 displays the differences in the total regional spending for beach visitation at the state level. Beachgoers spent $151.90 million in the Lake Erie region, which is the lowest among the 7 regions. Lake St. Clair generated $230.74 million in total expenditures, which is the second lowest. By contrast, Michigan Central received the largest amount of total spending at $1.05 billion, followed by Michigan North at $972.19 million and the Michigan South at $596.01 million.

Table 3-9 Economic impacts of total spending by region in 2014 dollars at state level

State level	Number of Trips (millions)	Total Spending by Region (millions)
Huron North	2.86	420.78
Huron South	2.93	408.26
St. Clair	1.79	230.74
Erie	1.16	151.90
Michigan North	6.73	972.19
Michigan Central	7.27	1052.00
Michigan South	4.11	596.01

5. Conclusions and Discussion

Spending analysis is an essential component of economic impact analysis. By using a visitor spending survey, this essay aims to estimate trip spending to Great Lakes beaches in order to provide the spending information to enable the quantification of the contribution of beach recreation to local economies. An on-site recruitment of beachgoers was conducted in three public beaches in Michigan in 2014. Intercepted beachgoers were asked to take a web survey about their beach activities and their spending of the visits. Unlike most literature, a sample selection model is used to address potential nonresponse bias problem in the spending data, so that the estimation of visitors' spending would be more accurate.

We further used the estimated spending equation to extrapolate an average beachgoer's spending per trip by using the 2011 Great Lakes Beaches Survey. Based on the demand system from Essay1, we were able to obtain the regional variation of spending from recreation trips to Great Lake beaches. We found the regional spending of an average beachgoer to Great Lakes beaches ranges from $35.92 to $248.80 in 2014 dollars. During the beach season, an average Michigan beachgoer spends $35.92 at Lake Erie. The average beachgoer spent the least at Lake Erie, followed by Lake St. Clair at $54.57. Beachgoers spent the most at $248.80 per person per season at Michigan Central, followed by Michigan North at $229.92 per person per season. The average beachgoer spent $99.51 in the North Huron region and $96.55 in the South Huron region.

This essay provides the necessary spending information for economic impact analysis in Essay 4. However, this essay also suffers some caveats. Primarily, the Beach Visitor Spending Survey is not a general population survey, although we used the selection model to correct the response bias, the estimated spending may not be representative of the whole state

ESSAY 4 Estimating the Economic Impacts of Changes in Water Quality by Linking a Recreational Demand System with Spending Data

1. Introduction

1.1 Motivations

By providing ample water-related recreational opportunities, the Great Lakes play an important role in Michigan residents' leisure activities. However, Great Lakes' water quality issues have long been a public concern. Bacterial contaminants, invasive species, algal growth, and harmful algae blooms are some common issues. Environmental agencies and local governments have spent public funding to improve the water quality of the Great Lakes. For instance, from 2010 to 2014, the Great Lakes Restoration Initiative has invested about $1.657 billion into over 2,500 projects over the Great Lakes (Great Lakes Restoration Initiative report, 2015), many of which are aimed at improving water quality. Thus, policy makers need information to understand the significance of the water quality improvements and to understand the impacts of their funding decisions.

Quantifying the contributions of water quality improvements to local economies can inform policy makers about some of the importance of improving water quality, as well as illuminating some of the distributional implications of programs. Understanding the regional distribution of the economic impacts water quality improvements, however, requires measurement of these economic impacts. Specifically, the core question is: What do water quality improvements of the Great Lakes contribute to local economies?

Economic impact analysis is a tool to address the proceeding question. Following Stynes (1997), economic impact analysis for recreation traces the flow of spending associated with visitation in a given region in order to determine the effects of recreation on the sales, income, and employment of that region's residents. Quantifying the economic impacts of water quality improvement to the local economy can demonstrate some of the importance of improving water quality and help policy makers evaluate water quality restoration and improvement programs.

However, measuring the regional economic impacts from water quality improvements is very challenging. Because water quality is a public good, water quality improvements can benefit a range of different activities for different people at different levels. Therefore, one challenge lies in the complexity of identifying the group of beneficiaries from water quality improvements (Keeler et al, 2012). As Bockstael, Hanemann and Kling (1987) indicated, significant benefits from surface water quality improvements accrue to recreational use. Thus we consider recreational beach use as the one of the beneficiaries and the medium to link water quality improvements of the Great Lakes and the local economic impacts.

This essay builds on Essay 2 and Essay 3 to quantity the economic impacts from water quality changes. Specifically, there are two steps involved: the first step of is to measure the economic impacts of beaches to the local economy; the second step is to set up the linkages between water quality and beach recreation to estimate the economic impacts of water quality improvements. By integrating the recreation demand system from Essay 2 and economic impact analysis from Essay 3, this essay establishes the critical linkages between water quality and beach recreation to estimate the regional economic impacts of access to beaches and the regional economic impacts of changes in water quality.

1.2 Research Gaps

Studies of the economic impacts of recreation have addressed a wide variety of activities and sites. The most common examples are national parks (Stynes and Sun, 2003; 2004; National

Park Service Visitor Spending Effects Reports, 2012; 2013; 2014), state parks (Bergstrom et al. 1990a; 1990b; Stynes 1998), recreational fishing (Lovell, Steinback, Hilger, 2013), lake recreation (Bergstrom et al, 1996), forest recreation (Starbuck et al, 2006), and recreational boating (Stynes 1983, Lee 2001).

Nevertheless, few economic impact studies have focused on freshwater beach recreation and water quality changes, and little is known about these impacts in the Great Lakes. The lack of such information can lead policy makers and the public to neglect the economic contributions of Great Lakes beaches and water quality improvement programs. Knowing some of the economic impacts from water quality changes, specifically for the Great Lakes, could fill the gap in the literature and help policy makers better allocate funds and evaluate the water quality restoration or improvement programs. In addition, an economic impact analysis focused on beach recreation can demonstrate the contribution of beach recreation itself and help policy makers, more specifically, park and recreation administrators and planners, as well as the local community, evaluate potential beach development or protection programs.

Furthermore, there are only a few studies in the existing literature on economic impacts where researcher have attempted to incorporate a trip demand model[16] (Bergstrom, et al., 1996; Hamel, et al, 2002; Starbuck, et al., 2006; Hutt, et al., 2013; Deisenroth, et al., 2013). Based on the methods applied in a recreation demand model, we can categorize the literature into three strands: the revealed preference method (RP), the stated preference method (SP) and combinations of RP and SP method.

[16] As Stynes (Economic Impacts of Tourism, 1997 pp.8) put it, "This step is usually the weakest link in most tourism impact studies, as few regions have accurate counts of tourists, let alone good models for predicting changes in tourism activity or separating local visitors from visitors from outside the region."

Deisenroth et al. (2013) applied a repeated three level nested logit model to predict the number of fishing visits among 48 sites in Mono County, CA. When fish stocking changes, they used IMPLAN input-output software to simulate the changes in economic impacts. Pendleton et al. (2011) employed a three-level nested logit model to predict trip changes due to beach width changes for 51 public beaches in southern California. Through the relationship of sea level changes and beach erosion, they estimated the potential impact of climate change on the economy due to beach use. However, in these two studies, only one environmental quality attribute (i.e., fish stocking, beach width) was examined. This reveals a limitation of some RP approaches; in RP analyses it can be difficult to collect attributes of environmental quality, especially water quality attributes, either because the data does not exist for all sites or because it lacks variation across sites.

For the SP method, Bergstrom et al. (1996) used a Tobit model for a single site recreation demand equation that incorporated contingent behavior questions. They estimated economic impacts of recreation spending at Lake Guntersville, Alabama under 5 scenarios of aquatic plant management. Hutt et al. (2013) estimated the economic value of recreational fishing at Mississippi reservoirs using the contingent valuation method, and then estimated economic impacts of these fisheries. Hamel (2002) linked a simple participation-rate model incorporating contingent behavior questions with an input-output model to estimate the economic impact from saltwater sport fisheries in Alaska. Although these studies had more environmental quality attributes to use, none of the results can be directly applied to the Great Lakes.

Starbuck, et al. (2006) used a truncated Poisson model to pool the RP data and contingent behavior (SP) data to simulate linkages from fire and fuels management activities to changes in forest recreation demand, and ultimately to regional economic impacts. However, Starbuck et al.

used a single site of demand model. A drawback of the single site framework is the difficulty of effectively modeling potential substitute sites. This has led to the development of random utility maximization (RUM) models to analyze the discrete choice among several recreation sites to visit (Phaneuf and Smith, 2005).

In this study, we adapt the demand system from the essay2, which used the calibration approach of combining RP and SP methods to incorporate the water quality attributes into the recreational demand system. In addition, the RP data was estimated by a repeated three-level nested logit model under the RUM framework, which can better capture the substitution effects among recreational sites. As Deisenroth, Loomis and Bond (2013) pointed out, failure to account for substitution effects in recreational demand from water quality changes results in overestimation of economic impacts. Moreover, most economic impact studies only provide a "snapshot" of an economic activity's contribution at a given point in time. However, when environmental quality attributes change, recreational demand will change. For instance, if water quality decreases on one beach site, the probability of it being chosen decreases. Beachgoers would go to other beaches or would forego visiting at all. Therefore, when quantifying economic impacts from water quality changes, researchers should not ignore the substitution effects.

1.3 Objectives

To address the above research gaps, there are two objectives of this study:

- *To estimate the economic impacts of beach recreation at regional levels*. This paper focuses specifically on Great Lakes beaches in Michigan, thus contributing to the economic impacts studies in the region by reporting the economic impacts of beach recreation along the shoreline of the Great Lakes in Michigan.

- *To establish the critical linkages between water quality and beach recreation to estimate the economic impacts of water quality changes by region*. By integrating the recreation demand system from Essay 2 and spending analysis from Essay 3, this essay is able to establish the critical linkages between water quality, beach recreation and spending to estimate the economic impacts of water quality improvements. Moreover, by using the repeated three-level nested logit model of RP data under the RUM framework, substitution effects are accounted for in the results, and the availability of multiple water quality attributes enables us to enlarge the scope of scenario analysis and policy implications.

2. Method

Economic impact analysis is originally from the Input-Output Model developed in macroeconomics to study the interdependence of industry sectors. According to Bergstrom et al. (1990), when non-resident[17] beachgoers take a trip to a region, the region basically "exports" the recreation services associated with the beach. The revenue generated from beachgoers stimulates the local economy by direct, indirect, and induced effects.

For example, assume beachgoers dine in restaurants near the beach. In order to provide food to beachgoers, restaurants need to purchase food, which ultimately comes from farmers. This first-round purchase is a direct effect of spending. Farmers need to increase their production by purchasing more inputs, such as fertilizer, which leads fertilizer producers to increase purchases of their inputs to produce more of their product. These "chain effects" of additional purchases are considered indirect effects. Both the direct effect and the indirect effects of the beachgoer's

[17] The current approach in this essay does not differentiate residents and non-residents of a local area.

spending stimulate the overall increase of production, along with the increased employment and income in the region. This increased income leads to more consumer demand, considered the induced effects.

Economic impact analysis measures the direct, indirect, and induced effects from an economic activity. In the area of recreation and tourism, methods for economic impact analysis are well defined (see, Stynes, 1997; Bergstrom et al, 1990a; Frechtling, 2000) and can be summarized in three steps. The first step is to measure the total number of trips in the studied areas. The second step is to measure the average spending per person per trip, which can be obtained by conducting a visitor spending survey. The final step is to apply multipliers from an Input-Output model to calculate the indirect and induced effects. The basic equation for estimating economic impacts of tourism can be simplified as:

$$Economic\ Impact = Number\ of\ Trips * Average\ spending\ per\ trip * Multiplier$$

This section consists of four parts: the recreational demand system (from Essay 2), the spending analysis (from Essay 3), the multiplier and finally, the economic impacts on local economies from beach recreation and changes in water quality.

2.1 Recreational Demand System

In Essay 1, we applied a repeated three-level nested logit model to the Great Lakes Beaches trip data, which explains the site choice and recreation demand for trips to Great Lakes beaches in a summer season. Trips are distinguished by Great Lake and beach location. During the Memorial Day weekend to September 30, 2011, we assume every single-day, a beachgoer simultaneously decides whether or not to go and if so where to go to a beach. The summer season consists of a fixed number of 126 choice occasions (T).

Due to a lack of beach quality data, we were not able to include all water quality attributes that may influence beachgoers' site choice by using RP data only. Essay 2 explored the strategies to incorporate water quality attributes by first using joint estimation of revealed preference data and stated preference data. Because of failing to pass the consistency test, we applied the calibration of SP to RP approach instead. By using the calibrated joint model, we are now able to predict the trip change in response to a particular water quality policy that alters the amount of algae in the water and on the shore of beaches.

Following the same notation and variables' definition of Essay 2, the indirect utility for the calibrated joint model is:

$$
\begin{aligned}
V_{jlt}{}^{c} = {} & \beta_{tc}{}^{RP} \cdot travel\ cost_{jl} + \beta_{l}{}^{RP} \cdot \log\left(beach\ length_{jl}\right) + \omega_{t}{}^{RP} \cdot temperature_{jlt} + \omega_{cd}{}^{RP} \\
& \cdot closure\ days\ of\ 2010_{jl} + \omega_{r}{}^{RP} \cdot regional\ dummies_{jl} + \theta^{c}(\delta_{aw}{}^{SP} \\
& \cdot algae\ water\ dummies_{jt} + \delta_{as}{}^{SP} \cdot algae\ shore\ dummies_{jt} + \delta_{bt}{}^{SP} \\
& \cdot bacteria\ testing\ dummies_{jt})
\end{aligned}
$$

where beach alternative $j \in \{1, 2, \dots, 451\}$ and choice occasion $t \in \{1, 2, \dots, 126\}$. To construct the status quo of the water quality for the Great Lakes beaches, we rely on the RP data. Under the status quo situation, assume the indirect utility for individual *n* who takes a trip to beach *j* at Lake *l* at the choice occasion *t* is:

$$
\begin{aligned}
\hat{V}_{jl,nt}{}^{0} = {} & \hat{\beta}_{tc}{}^{RP} \cdot travel\ cost_{jl,n}{}^{0} + \hat{\beta}_{l}{}^{RP} \cdot \log\left(beach\ length_{jl,n}{}^{0}\right) + \widehat{\omega}_{t}{}^{RP} \cdot temperature_{jl,nt}{}^{0} \\
& + \widehat{\omega}_{cd}{}^{RP} \cdot closure\ days\ of\ 2010_{jl,n}{}^{0} + \widehat{\omega}_{r}{}^{RP} \cdot RD_{jl,n}{}^{0}
\end{aligned}
$$

Now consider a change of water quality at one or more regions, for instance, change the algae level in the water. Assume that $AW_{r,n}{}^{0}$ represents the algae level in the water at region *r* for person *n* without an improvement and assume that $AW_{r,n}{}^{*}$ represents algae level in the water with an improvement. All other site characteristics remain the same, only the algae level in the water at region *r* has changed between the two states of the world. With the change in the water quality, the indirect utility for individual *n* for a trip to beach *j* at Lake *l* at choice occasion *t* is[18]:

$$\hat{V}_{jl,nt}{}^{*} = \hat{V}_{jl,nt}{}^{0} + \hat{\theta}^{c}\left(\hat{\delta}_{aw}{}^{SP} \cdot \left(AW_{r,n}{}^{*} - AW_{r,n}{}^{0}\right)\right) = \hat{V}_{jl,nt}{}^{0} + \hat{\theta}^{c}\left(\hat{\delta}_{aw}{}^{SP} \cdot \Delta AW_{r,n}\right)$$

For individual *n* at choice occasion *t*, the predicted total number of taking trips to beach *j* at lake *l* is:

$$\hat{Y}_{jlG,n} = \sum_{t=1}^{T} \hat{P}_{jlG,nt}{}^{0}$$

Taking the weighted sum of the number of trips to a specific beach across *all* the beaches in the region *r* gives the weighted average number of trips per person to beaches in that region. If the water quality attributes changed, the change in predicted total number of trips to beach *j* at Lake *l* is:

$$\Delta\hat{Y}_{jlG,n} = \sum_{t=1}^{T} \hat{P}_{jlG,nt}{}^{*}\,(scenario) - \sum_{t=1}^{T} \hat{P}_{jlG,nt}{}^{0}(status\ quo)$$

2.2 Spending of Trips to Great Lakes Beaches

Essay 2 estimated trip spending to Great Lakes beaches by using the 2014 Beach Visitor Spending Survey. A Heckman model was used to address potential nonresponse bias problem in the spending data to more accurately estimate visitors' spending. The estimated spending equation

[18] Derivation of the indirect utility under the scenario can be found in Essay 2.

was further applied to the 2011 Great Lakes Beaches Survey to extrapolate an average beachgoer's spending per trip for all possible sites they could visit.

Following the same notation as Essay 2, let the predicted spending for person *n* to beach *j* be $\widehat{m}_{j,n}$. Then the total spending for person *n* to beach *j* in a beach season is:

$$\widehat{M}_{j,n} = \widehat{Y}_{jlG,n} * \widehat{m}_{j,n}$$

Taking the weighted sum of the total spending per person to a specific beach across *all* the beaches in region *r* gives the weighted average total spending per person to beaches in that region. That is to say, the predicted total spending per person per season to beaches in a region is:

$$\overline{M}^r = \frac{\sum_{n=1}^{N} w_n * (\sum_{j=1}^{R_r} \widehat{M}_{j,n})}{\sum_{n=1}^{N} w_n}$$

If the water quality changed, the change in the total spending for person *n* to beach *j* in a beach season is:

$$\Delta\widehat{M}_{j,n} = \Delta\widehat{Y}_{jlG,n} * \widehat{m}_{j,n}$$

The change in the predicted total spending per person per season to beaches in a region is:

$$\Delta\overline{M}^r = \frac{\sum_{n=1}^{N} w_n * (\sum_{j=1}^{R_r} \Delta\widehat{M}_{j,n})}{\sum_{n=1}^{N} w_n}$$

2.3 Multipliers

Multipliers can be used to estimate the indirect and induced economic effects of an economic activity. Multipliers are derived from Input-Output models to measure the interdependencies between sectors within a particular region's economy. Because different regions have different economic sectors, multipliers may be different for each region. Generally, larger regions or regions with more diversified economies have higher multipliers, while smaller regions or regions with limited economic development have lower multipliers.

Multipliers can be borrowed from published studies or other sources. The National Park Service has provided visitor spending and economic impact analysis for Sleeping Bear Dunes National Lakeshore (Cook, 2009). According to Cook's study, the multiplier of the tourism related sales for a three-county region is 1.64, which is derived from input-output models estimated with the IMPLAN software using 2008 economic databases. Because our study is also applied to beach recreation at Great Lakes, we adopt the multiplier of 1.64 to use in our application. This multiplier is specifically applied to the direct sales, which means that every dollar of direct sales made by beachgoers within the region generates $1.64 of total sales in the region through indirect and induced effects.

2.4 Economic Impact Analysis

As a summary of section 2, Figure 4-1 provides an overview of the approach to estimate economic impacts of visiting Great Lakes public beaches. Economic impact analysis consists of three components: trip estimation, spending estimation and economic multipliers, which are the top three parts in the Figure 4-1. Following the flows in each component leads to the average economic impacts of beach visitation per person, which are calculated as the total amount of trips in the region times the average spending per person times the multipliers in the region.

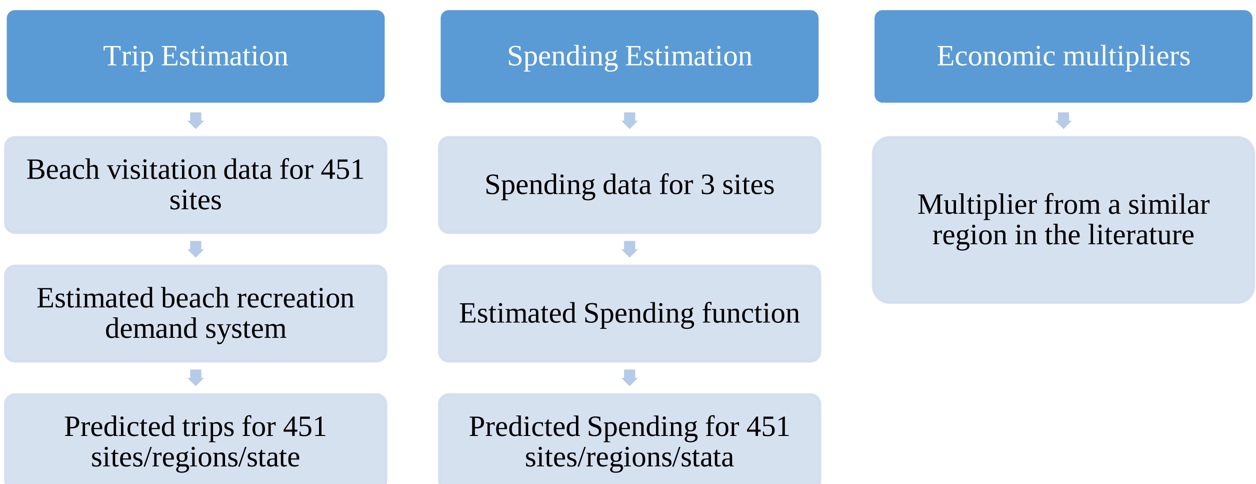


Figure 4-1 Detailed approach to estimate economic impacts of visiting Great Lakes public beaches

To calculate the regional variation of economic impacts of beach visitation, let the economic impact in region *r* to be EI^r. In region *r*, the economic impact of beach visitation per person per season is:

$$EI^r = \frac{\sum_{n=1}^{N} w_n * \left(\sum_{j=1}^{R_r} \hat{Y}_{jlG,n} * \hat{m}_{j,n}\right) * \varphi^r}{\sum_{n=1}^{N} w_n} = \bar{M}^r * \varphi^r$$

where φ_r are the economic multipliers for region *r*.

<u>*Changes in Economic Impacts in Response to Water Quality Changes*</u>

This section aims to answer the following question: to what extent will economic impacts from beach visitation change if the sites' water quality changes? From section 2.1, if the water quality attributes changed, the change in predicted total number of trips to beach *j* at Lake *l* is:

$$\Delta\hat{Y}_{jlG,n} = \sum_{t=1}^{T} \hat{P}_{jlG,nt}{}^{*}\,(scenario) - \sum_{t=1}^{T} \hat{P}_{jlG,nt}{}^{0}(status\ quo)$$

Plugging the change of trips into the economic impact equation, we can obtain the change of economic impacts as:

$$\Delta EI^{r} = \Delta\bar{M}^{r} * \varphi^{r} = \frac{\sum_{n=1}^{N} w_n \ * \left(\sum_{j=1}^{R_r} \Delta\hat{Y}_{jlG,n} * \hat{m}_{j,n}\right) * \varphi^{r}}{\sum_{n=1}^{N} w_n}$$

Based on the calculation, we can know the extent to which these economic impacts change in response to a change in some observed water quality attributes, such as algae levels. The change in economic impacts shows that the responsiveness of economic impacts to water quality changes (ΔEI^{r}) comes from the trips change to water quality changes ($\Delta\hat{Y}_{jlG,n}$). The above procedures for linking the recreational demand system with spending data are illustrated in Figure 4-2.

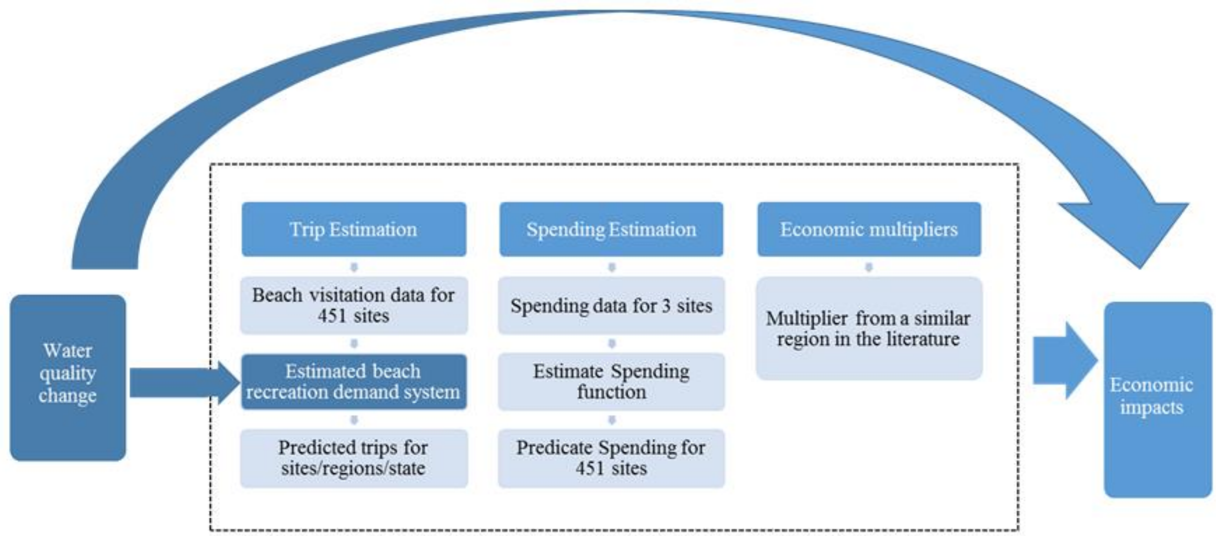


Figure 4-2 the linkage between water quality change and economic impacts

3. Data

Two surveys are applied in this essay. The first one is the Great Lakes Beaches Survey, which was used in the recreation demand system in Essay 2. The second survey is the Beach Visitor Spending Survey, which was used for the spending estimation in Essay 3. Details of the survey can be referred to Essay 2 and Essay 3.

In the Great Lakes Beaches Survey, we used the trip data and choice experiment data. The trip data was collected by asking respondents' trips to public Great Lakes beaches in one summer season from Memorial Day weekend to September 30, 2011. The choice experiment data was gathered by asking respondents' preferred beach in each of three different choice sets with experimentally designed attributes. The trip data has 2,573 observations, 1,894 individuals took at least one trip to Great Lakes beaches during the beach season. The choice set for each individual consists of 451 beaches. The sample size of respondents for choice experiment data is 946, with 2,785 choice sets. Each choice set has two alternatives.

The Beach Visitor Spending Survey has 157 observations used for spending estimation, 336 observations were used to correct for response/nonresponse bias. The estimated spending equation was applied to 2,537 beachgoers from the Great Lakes Beaches Survey. Because each beachgoer has 451 beach alternatives in the choice set, the sample for prediction has 1,144,187 observations.

4. Results

4.1 Economic Impact of Beach Visitation by Region

This section provides the economic impacts of Great Lakes beaches visitors' spending on the local economy. Table 4-1 displays the regional differences in the economic impacts of beach visitation per person per season. The direct sales of an average beachgoer to Great Lakes beaches in one region ranges from $61.41 to $248.62 per season in 2014 dollars. If the sales multiplier for every region is 1.64 (Cook, 2009), the spending by an average Michigan beachgoer had a total economic impact of direct sales on one region that ranges from $100.72 to $407.74 per season. Specifically, during a beach season, an average Michigan beachgoer to Mid-East region generates the lowest total sales at $100.72, followed by Northeast region at $155.65. Beachgoers to Mid-West region have the highest total sales at $407.74 per person per season, followed by Northwest region at $368.94 per person per season.

Table 4-1 Economic Impacts of access to great lakes beaches by region in 2014 dollars

Per Person Per Season			
		Direct Sales	Total Sales
Access to Beaches	LP Northeast	94.91	155.65
	LP Mid-East	61.41	100.72
	LP Southeast	125.64	206.04
	LP Northwest	224.96	368.94
	LP Mid-West	248.62	407.74
	LP Southwest	140.92	231.11
State level			
		Direct Sales (Million)	Total Sales (Million)
Access to Beaches	LP Northeast	401.30	658.13
	LP Mid-East	259.68	425.87
	LP Southeast	531.24	871.23
	LP Northwest	951.23	1560.00
	LP Mid-West	1051.30	1724.10
	LP Southwest	595.87	977.23

To calculate the state level economic impacts for access to beaches in each region, we aggregated the weighted average economic impacts per person to all beachgoers living in the Lower Peninsula. The population number of beachgoers is derived from the participation rate of beach recreation, which is 58.01%, multiplied by 7,289,085 Michigan adults living in the Lower Peninsula. Table 4-1 displays the regional differences in the economic impacts of beach visitation at the state level. Multiplied with the sales multiplier—1.64, the $259.68 million spent by beachgoers to Mid-East region had a total economic impact on the region of $425.87 million in direct sales, which is the lowest among the 6 regions. Visitors to the beaches in the Northeast region supported $658.13 million of total direct sales, which is the second lowest. By contrast, Michigan Central received the largest amount of total direct sales at 1.72 billion, followed by Michigan North at $1.56 billion and the Michigan South at $977.23 million. Figure 4-1 shows regional variation of the total sales at state level from beach visitation.

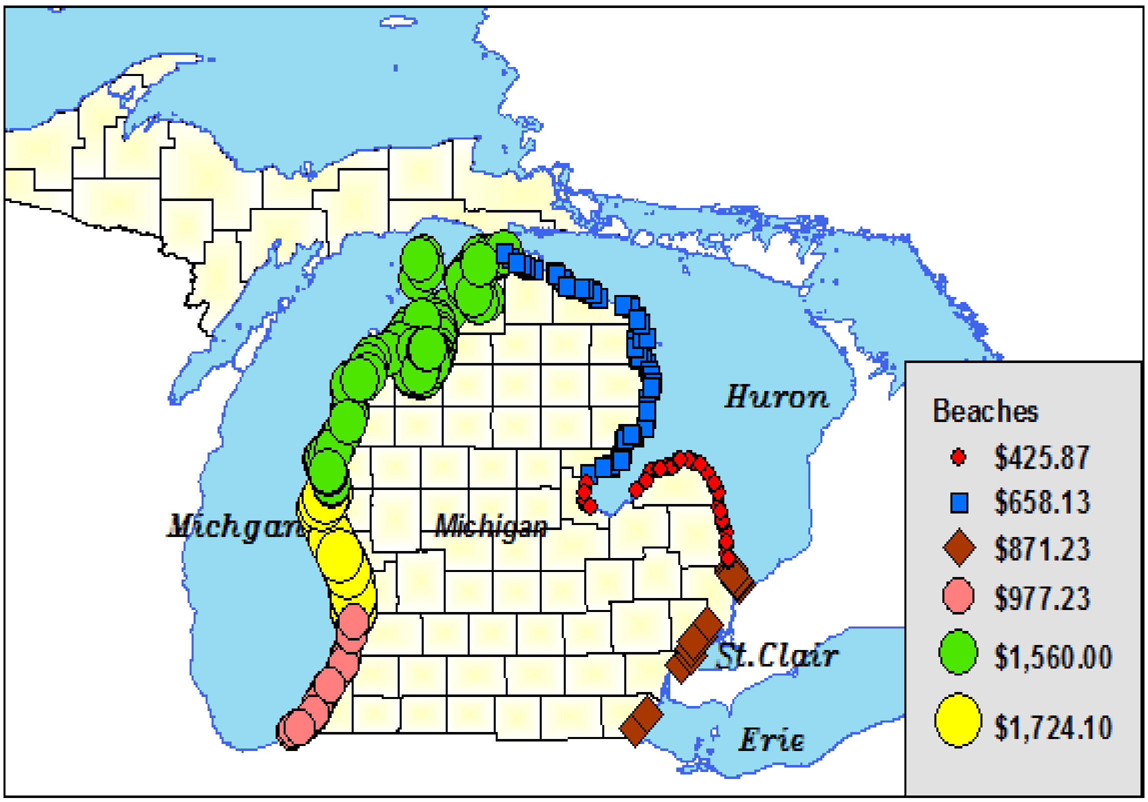


Figure 4-3 Total sales from beach visitation by region in 2014 dollars (millions)

4.2 Economic Impacts in Response to Water Quality Changes

As in Essay 2, we consider two types of welfare scenarios using our calibrated joint model. The first scenario assumes that water quality at half of the sites in a region is improved *up* by one level. Simply put, half of Great Lakes beaches in a region with the high algae level are improved to the medium level and half of beaches in a region with the medium algae level are improved to the low level. The second scenario assumes that water quality is deteriorated by shifting half of the sites' water quality in a region *down* by one level. This is a significant change in water quality, because half of beaches with the low algae level are degraded to the medium level and half of beaches with the medium algae level are degraded to the high level. In both types of scenarios the algae changes are made only within one region at a time, resulting in twelve total welfare scenarios (an improvement and decrement to quality in each of six regions).

Table 4-2 displays the economic impact and the changes in the economic impact from the first scenario of water quality improvement. If we improve half of Great Lakes beaches' water quality in a region *up* by one level, compared to the direct sales at status quo, the direct sales increases by 33.52% for Middle-East region (Huron South) and 20.63% for Southeast region (St. Clair and Erie). Direct sales increase slightly for Huron North and Lake Michigan. The intuition behind this is that the baseline algae levels in Huron South, St. Clair, and Erie are higher than those in Huron North and Lake Michigan. Once we increase the water quality, the utility of a person is increasing as the algae level decreases. Therefore, improving water quality leads to more utility increase for beaches with initially higher algae levels in Huron South, St. Clair, and Erie than for beaches with initially lower algae levels in Huron North and Lake Michigan. In particular, direct sales from Southwest region never change, because the baseline water quality in the Southwest region was already at the highest level.

Under the water quality improvement scenario, the change of total sales of an average beachgoer to Great Lakes beaches in one region ranges from $0 to $42.50 per season in 2014 dollars. When aggregated at the state level, improvements of water quality in Southeast region (Lake St. Clair and Lake Erie) results in $179.7 million more total sales by all Michigan beachgoers living in the Lower Peninsula, which is the highest change of total sales in 6 regions, followed by Mid-East region with $142.76 million more total sales. Again, change of total sales from South Michigan were zero because it had the highest water quality at status quo. Figure 4-1 shows the changed total sales from water quality improvement in a region in 2014 Dollars at the state level.

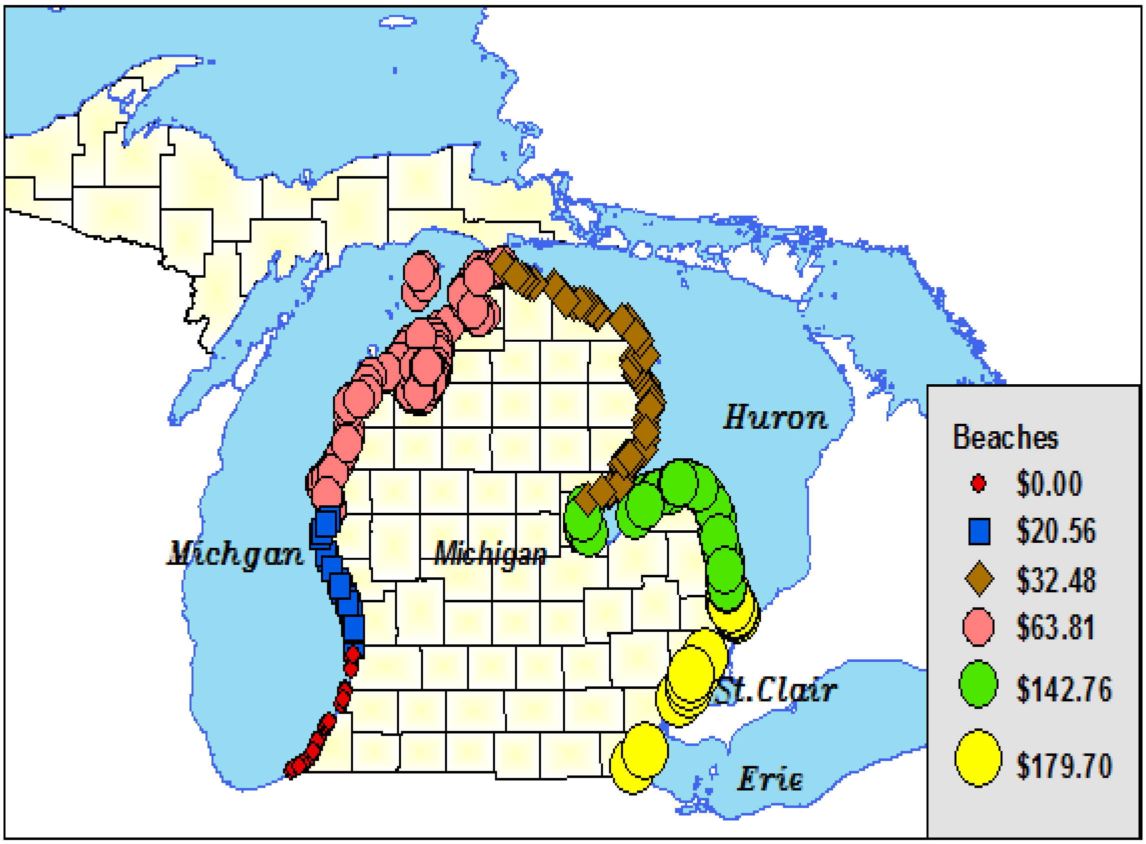


Figure 4-4 Changed total sales from improving water quality by one level at half of the sites in a region in 2014 dollars (millions)

By contrast, if we degrade half of Great Lakes beaches' water quality in a region *down* one level, direct sales decrease dramatically and loss of total sales turns out to be significant. Table 4-3 displays the economic impact and changes in economic impacts from the second scenario of the water quality deterioration. Compared to the direct sales at status quo, all regions lose sales and the magnitude of decreased direct sales ranges from 23.87% to 32.58% across the six regions. When aggregated at the state level, 421.12 million total sales are lost in the Northwest region due to degrading half of Great Lakes beaches' water quality in that region *down* by one level. Mid-west region loses $411.61 million total sales, followed by Southwest region losing $246.12 million total sales. Mid-East region loses $138.76 million total sales, which is the least sales loss among the six regions. The range of total sales loss indicates that the water quality degradation impacts Lake Michigan most and Huron south least. Figure 4-2 shows the changed total sales from water quality degradation in a region in 2014 Dollars at the state level.

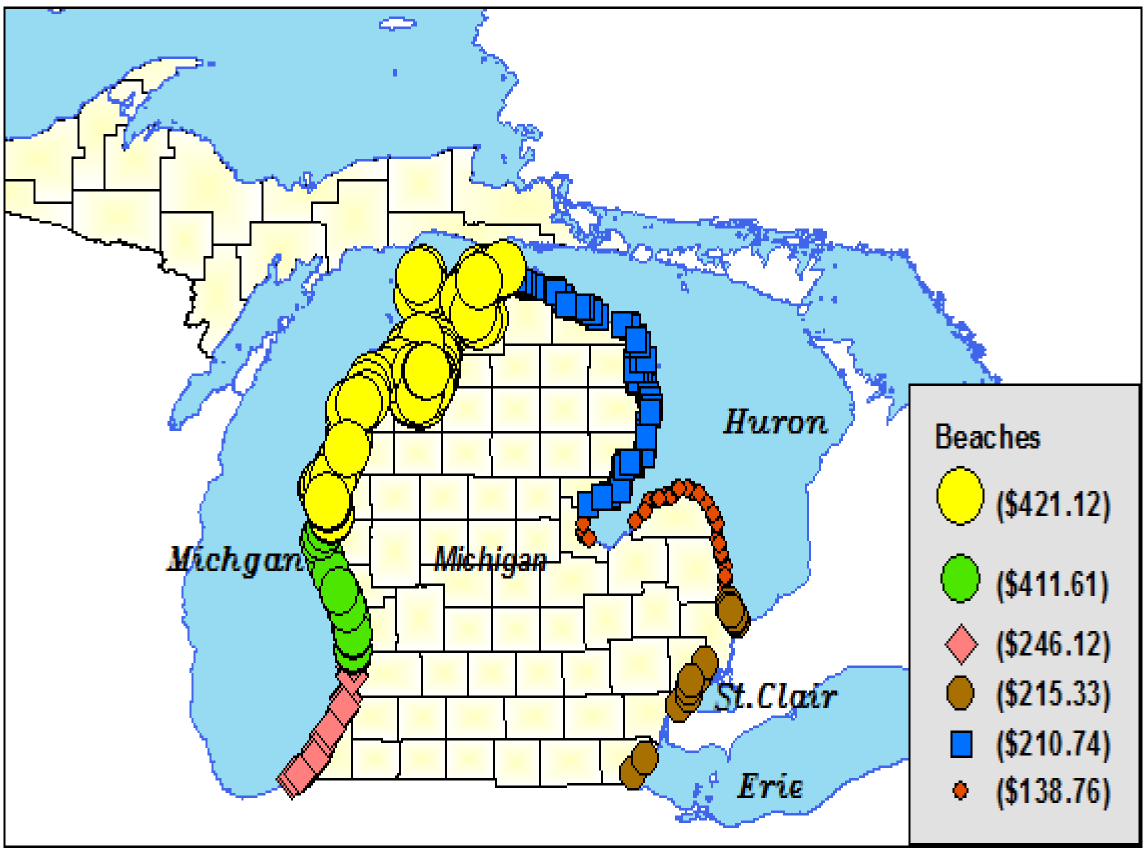


Figure 4-5 Changed total sales from decreasing water quality by one level at half of the sites in a region in 2014 dollars (millions)

Table 4-2 Changes in economic impacts from improving water quality by one level at half of sites in a region in 2014 dollars

Per Person Per Season

		Direct Sales	Total Sales	Change of Direct Sales	% Change in Direct Sales	Change of Total Sales
Take Half of Sites' Algae in the Water & Algae on the Shore *up* by one Level	LP Northeast	99.59	163.33	4.68	4.94%	7.68
	LP Mid-East	82.00	134.48	20.59	33.52%	33.76
	LP Southeast	151.55	248.54	25.91	20.63%	42.50
	LP Northwest	234.16	384.03	9.20	4.09%	15.09
	LP Mid-West	251.59	412.60	2.96	1.19%	4.86
	LP Southwest	140.92	231.11	0.00	0.00%	0.00

State level

		Direct Sales (Million)	Total Sales (Million)	Change of Direct Sales (Million)	% Change in Direct Sales	Change of Total Sales (Million)
Take Half of Sites' Algae in the Water & Algae on the Shore *up* by one Level	LP Northeast	421.10	690.61	19.80	4.94%	32.48
	LP Mid-East	346.73	568.64	87.05	33.52%	142.76
	LP Southeast	640.82	1050.90	109.58	20.63%	179.70
	LP Northwest	990.14	1623.80	38.91	4.09%	63.81
	LP Mid-West	1063.80	1744.70	12.54	1.19%	20.56
	LP Southwest	595.87	977.23	0.00	0.00%	0.00

Table 4-3 Changes in economic impacts from decreasing water quality by one level at half of the sites in a region in 2014 dollars

Per Person Per Season

		Direct Sales	Total Sales	Change of Direct Sales	% Change in Direct Sales	Change of Total Sales
Take Half of Sites' Algae in the Water & Algae on the Shore *down* by one Level	LP Northeast	64.52	105.81	-30.39	-32.02%	-49.84
	LP Mid-East	41.40	67.90	-20.01	-32.58%	-32.82
	LP Southeast	94.59	155.12	-31.05	-24.71%	-50.92
	LP Northwest	164.23	269.34	-60.73	-27.00%	-99.59
	LP Mid-West	189.27	310.40	-59.36	-23.87%	-97.34
	LP Southwest	105.43	172.90	-35.49	-25.19%	-58.21

State level

		Direct Sales (Million)	Total Sales (Million)	Change of Direct Sales (Million)	% Change in Direct Sales	Change of Total Sales (Million)
Take Half of Sites' Algae in the Water & Algae on the Shore *down* by one Level	LP Northeast	272.80	447.39	-128.50	-32.02%	-210.74
	LP Mid-East	175.07	287.11	-84.61	-32.58%	-138.76
	LP Southeast	399.94	655.91	-131.30	-24.71%	-215.33
	LP Northwest	694.45	1138.90	-256.78	-27.00%	-421.12
	LP Mid-West	800.30	1312.50	-250.98	-23.87%	-411.61
	LP Southwest	445.80	731.11	-150.07	-25.19%	-246.12

5. Conclusions

Essay 4 estimated regional variation in economic impacts from trips to Great Lakes beaches in Michigan. By integrating the recreation demand system from Essay 2 and spending analysis from Essay 3, this essay established the critical linkages between water quality and beach recreation to estimate the economic impacts of water quality improvements. By constructing two types of water quality scenarios, this essay further estimated the changes in economic impacts to the local region when water quality changes.

In considering the impacts of a loss of access to beaches within a region, we found the spending by all Michigan beachgoers living in the Lower Peninsula had a total economic impact of direct sales within a region that ranged from $425.87 million to $1,724.1 million per season in 2014 dollars. Michigan Central received the largest amount of total direct sales at 1.72 billion, in contrast to Huron South region with the lowest total sales at $425.87million. At the state level, under the water quality improvement scenario, the gains of total sales of beachgoers to Great Lakes beaches in a region ranged from $0 to $179.70million per season in 2014 dollars. Under the water quality degradation scenario, the loss of total sales of Michigan beachgoers to Great Lakes beaches in one region ranged from $246.12 million to $421.12 million per season in 2014 dollars.

The results of Essay 4 can demonstrate the contribution of beach recreation, some of the importance of improving water quality, and help policy makers to evaluate water quality restoration and improvement programs. However, this essay is not without caveats. Due to the small sample size of the Beach Visitor Spending Survey, we did not differentiate local residents and non-residents. In addition, according to Stynes (1997) an economic impact analysis should not include any local residents who live in the same county as the trip destination. Since the latter has not yet been done, the economic impacts in Essay 4 might be overestimated. Finally, the multipliers were transferred from the best available study—the Sleeping Bear Dunes National Lakeshore. Therefore, we might overestimate the economic impacts of beaches located in a rural region and underestimate the economic impacts of beaches located in a metropolitan region. but future research could consider running the spending profiles through an input-output model for each region.

APPENDICES

Appendix A

Trips Trimming Strategy and Weighting Method in Essay 1

A.1 Trips Trimming Strategy

If the total trips, day trips plus overnight trips, are greater than the total number of days in that month, the number of trips is trimmed into the total number of days in that month, i.e., Jun. 34, Jul. 31, Aug. 31, Sept. 30.

Table A-1 Trips trimming strategy and number of observations trimmed

Month	Trips Trimmed at	Number of Individuals with Trimmed Trips	Percentage of all Trips
Jun.	34	2	0.0788%
Jul.	31	6	0.2365%
Aug.	31	6	0.2365%
Sept.	30	7	0.2759%

A.2 Average Number of Days Staying on the Beaches

In the web survey, if respondents reported overnight trips, we randomly drew a trip, then asked them to report the number of days staying on the beaches. The average number of days of staying on the beaches for overnight trips are reported in Table A-2.

Table A-2 Average number of days staying on the beaches for overnight trips

Type of Trips	Observations	Mean	Std. Dev	Min	Max
Short Overnight Trips	1,211	2.0553	0.8801	1	4
Long Overnight Trips	632	4.0427	5.6932	0.5	126

A.3 Weighting Method

The final weight applied to each person is the product of 3 components. The first one is the sample weight for each person, which corrects for sampling strata and possible non-representativeness of the sample (see Chen, 2013, Appendix C). The second one is the downward weight to correct for multiple purposes for overnight trips, which is 91.08% for short overnight trips and 92.42% for long overnight trips, respectively. The final one is the weight used for correcting for adjusted trip counts. In our web survey, after respondents finished their trip log section, we summarized the number of each type of trip they reported into a table, then verified whether the numbers in the table sound correct to them or not. Only 3.59% of the total samples reported “No” in this trip verification question. For the person who reported “No”, they were given a new table and asked to correct the number of trips they took. For each type of trip, less than 1% of sample changed their number of trips. We used the ratio of the first reported number of trips to the changed number as the weight to correct for the trip adjustments. For instance, if a person first reported 20 for the total number of day trips, then changed to 10 after the verification question, we apply 10/20=0.5 to weight the monthly trip number of day trips. Similarly, we used the same method to correct for the downward adjustments.

Table A-3 Final weights applied to the three types of trips

	Individual Sample Weight	Main purpose Adjustment	Individual Trip Count Adjustment
Day trips	Applied to all	1	Applied if reported
Short overnight trips	Applied to all	0.9108	Applied if reported
Long overnight trips	Applied to all	0.9242	Applied if reported

Appendix B

Missing Income Imputation for 2011 Great Lakes Beaches Survey

B.1. Income imputation for the web survey

Some web survey respondents did not report their income. This appendix discussed the missing ncome imputation procedure.

Step1: Incorporate additional income categories from follow-up questions, and transfer categorized income into continuous income.

In the demographic section of the web survey, participants have the option to skip the income question, or choose a range of income from a choice of 8 categories. If respondents chose to not disclose the income, we gave them two follow-up questions with broader income categories: "Was your total household income in 2011 less than $XX,XXX?" In this way, the two follow up questions generated 4 more broad categories.

Table B-1 Income categories, continuous income that was assigned to the category, and their frequency in the web survey

Income category	Continuous Income	Frequency
Less than $24,999	$12,500	224
Less than $50,000	$30,404	16
$25,000 to $34,999	$30,000	188
$25,000 to $49,999	$37,500	39
$35,000 to $49,999	$42,500	315
$50,000 or more	$76,957	51
$50,000 to $74,999	$62,500	510
$50,000 to $99,999	$75,000	61
$75,000 to $99,999	$87,500	421
$100,000 or more	$164,210	35
$100,000 to $149,999	$125,000	385
$150,000 to $199,999	$175,000	128
$200,000 or more	$300,000	101
In total		2474

The method to transfer categorized income into continuous income is listed below:

For the bounded income categories, we take the midpoints.

For the unbounded income categories, we take $12,500 as the “less than $24,999” and $300,000 as “$200,000 or more”.

For the other unbounded categories, the broad category of $100,000 or more can be divided into $100,000 to $149,999; $150,000 to $ 199,999; and $200,000 or more. As for all 2,544 web respondents, 385 chose $100,000 to $149,999; 128 chose $150,000 to $199,999; and 101 chose $200,000 or more. Thus, the weighted average for the income of people who indicated $100,000 or more is: (385*$125,000+128*$175,000+101*$300,000)/(385+128+101)=$164,210.

Similarly, the weighted average for the income of people who chose $50,000 or more is (62500*510+75000*61+87500*421+164210*35)/(510+61+421+35)= $76,957. The weighted average as the income of people who chose $50,000 or less is (12500*224+30000*188+42500*315+37500*39)/(224+188+315+39)=30,404.

Step2: Filled the missing data with screener survey.

The total number of the web survey sample is 2,544, and the number of respondents who reported income is 2,413. Missing rate for income in the web survey is 131/2544=5.15%. However, if the respondent missed reporting income in the web survey but happened to report it in the screener survey, we will use the income in screener survey as replacement, considering the respondents are from the same household. In this way, the number of respondents who have reported income is 2,474, and the missing rate is (2544-2474)/2544=2.75%.

Step3: Regression Income imputation.

Our purpose is to impute the missing income by setting up a linear regression model (Little and Robin, 2002). We use income as the response variable, and it is treated as a continuous variables. It is hypothesized that gender will be a significant factor in determining individual income after controlling other variables, and income is increased by years of education, age (as a proxy for job experience), and employment. It also indicates that individual income is expected to increase with the number of children and household size (Resetar, 1978). Metropolitan and micropolitan areas are selected based on geographic categories of the metropolitan statistical area/ micropolitan statistical area of Michigan in 2010 by the United States Census Bureau. By their definition, a metropolitan area contains a core urban area of 50,000 or more residents; a micropolitan area contains an urban core of at least 10,000, but less than 50,000 residents. We are expecting metropolitan and micropolitan areas will have an increased effect on income.

Table B-2 Shows the variable choices.

Table B-2 Variable choices and description

Variable name		Description
wincome (dependent)		weighted income
gender		Gender, Male=1, Female=2
metro		Metropolitan Statistical Area
age		Age
race		White, African, Hispanic, Asian, Indian, Other
eduyear		Education Years
employment	Full Time	Full Time
	Part Time	Part Time
	Unemployment	Unemployment
	Stay Home Parent	Stay Home Parent
	Retire	Retire
Household size	couple	Couple
	child5	Single with Children under 5
	child17	Single with children 6 to 17
	imm	Single with Immediate Family
	ext	Single with Extended Family
	withchild	Single with Children
	cc5	Couple with Children under 5
	cc17	Couple with Children 6 to 17
	cc	Couple with Children
	wac5	Single with Adult and Children under 5
	wac17	Single with Adult and Children 6 to 17
	wa	Single with Adult

Table B-3 displays the OLS regression result:

Table B-3 Income estimates of OLS model for the web survey missing income imputation

Variable	OLS model
metro	15491.7***
	(3049.1)
gender	-7140.4*
	(3029.8)
age	281.3**
	(106.0)
white	8249.1
	(4618.4)
Asian	17878.8
	(15095.7)
eduyear	6523.6***
	(578.8)
fulltime	9709.8*
	(3869.7)
parttime	-4546.4
	(5643.8)
unemployment	-14002.3*
	(6673.4)
couple	34001.6***
	(2899.1)
child5	21059.2
	(13622.0)
child17	33904.2***
	(7595.9)
imm	52493.8***
	(7380.2)
ext	25592.9
	(15593.8)
withchild	22728.1*
	(11170.3)
cc5	53323.4***
	(9105.9)
cc17	53806.3***
	(4045.2)
cc	37454.4***
	(6739.1)
wac5	45405.9***
	(10205.3)
wac17	46329.8***
	(7520.6)
wac	63463.7***
	(17811.8)

Table B-3: (cont'd)

Variable	OLS model
wa	35496.9***
	(4574.7)
_cons	-77072.3***
	(13084.7)
N	2474
R-sq	0.170
adj. R-sq	0.163
rmse	56602.4

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001 .

Step4: Applied Different normalized weights.

According to the weights analysis, we have 4 weight choices. From the comparison between the income estimations by using different weights, there are no big differences in magnitude of the coefficients and the significant levels. To make full use of all the information, we finally applied the weights without truncation as the final weight.

Table B-4 Comparison between income estimations of OLS model for web survey by using different weights

	Income finalweight	Income Weight 0.3-3	Income Weight0.4-2.3	Income Weight 0.37-2.45
metro	15491.7***	15391.1***	15300.4***	15330.8***
	(-3049.1)	(-3023.8)	(-2972.3)	(-2929.2)
gender	-7140.4*	-6905.5*	-6611.0*	-6536.8*
	(-3029.8)	(-2935.6)	(-2828.3)	(-2787.30
age	281.3**	280.4**	288.3**	292.4**
	(-106)	(-102.5)	(-99.92)	(-98.72)
white	8249.1	7952.4	7775.1	7723.4
	(-4618.4)	(-4598.9)	(-4512.6)	(-4507.3)
Asian	17878.8	20331.7	21092.3	21366.9
	(-15095.7)	(-14441)	(-14197.7)	(-14155.8)
eduyear	6523.6***	6530.9***	6559.2***	6574.7***
	(-578.8)	(-568.6)	(-556.7)	(-550.5)
fulltime	9709.8*	9966.4**	10607.4**	10992.5**
	(-3869.7)	(-3755.6)	(-3643.6)	(-3577.4)
parttime	-4546.4	-4680.5	-4563.6	-4356.8
	(-5643.8)	(-5382.4)	(-5111.5)	(-5003.9)
unemployment	-14002.3*	-13472.5*	-12627.4*	-12357.3
	(-6673.4)	(-6544.2)	(-6435.3)	(-6358.3)
couple	34001.6***	34547.2***	35033.7***	35246.5***
	(-2899.1)	(-2807.6)	(-2734.6)	(-2711)
child5	21059.2	21049.1	21822.4	22481.7
	(-13622)	(-13590)	(-13807.5)	(-13975.4)
child17	33904.2***	33124.7***	32381.4***	32058.7***
	(-7595.9)	(-7320.6)	(-7092.9)	(-7006.5)
imm	52493.8***	51634.4***	50684.7***	50377.0***
	(-7380.2)	(-7051.5)	(-6753.4)	(-6656.3)
ext	25592.9	23758.1	22659.4	22633.7
	(-15593.8)	(-13997.8)	(-12412.1)	(-12015.8)
withchild	22728.1*	22506.3*	21291.5*	21104.9*
	(-11170.3)	(-11022.5)	(-10518.3)	(-10392.2)
cc5	53323.4***	53540.5***	53353.0***	52950.1***
	(-9105.9)	(-9079.5)	(-8960.2)	(-8816.9)
cc17	53806.3***	53922.4***	54185.0***	54276.1***
	(-4045.2)	(-3991.3)	(-3949)	(-3937.9)
cc	37454.4***	37584.6***	37846.0***	37943.8***
	(-6739.1)	(-6720.5)	(-6698.5)	(-6691.6)
wac5	45405.9***	45403.2***	45603.8***	45562.3***
	(-10205.3)	(-10208.4)	(-10211.7)	(-10217.2)
wac17	46329.8***	46762.7***	47305.6***	47318.7***

Table B-4: (cont'd)

	Income finalweight	Income Weight 0.3-3	Income Weight0.4-2.3	Income Weight 0.37-2.45
	(-7520.6)	(-7482.6)	(-7425.2)	(-7307.6)
wa	35496.9***	36347.8***	36779.9***	36937.7***
	(-4574.7)	(-4304.7)	(-4197.1)	(-4165.2)
wac	63463.7***	63310.0***	64152.3***	64540.0***
	(-17811.8)	(-17800.6)	(-17686.3)	(-17596)
_cons	-77072.3***	-77512.4***	-79126.8***	-79957.0***
	(-13084.7)	(-12797.3)	(-12486.8)	(-12345.9)
N	2474	2474	2474	2474
R-sq	0.17	0.173	0.176	0.178
adj. R-sq	0.163	0.163	0.166	0.169
rmse	56602.4	56309.5	55973.5	55855.4

B.2. Income Imputation for the Screener Survey

The total sample size of the screener survey is 9,591, and the number of respondents who did not report their income is 488. Missing rate for income in the screener survey is 1031/9591=10.75%.

In the screener survey, there are not as many income categories as in the web survey. The income categories are assigned as follows: 1 "Less than $25,000"; 2 "$25,000 to $49,999"; 3 "$50,000 to $99,999"; and 4 "$100,000 and higher". The frequency distribution of income is listed in Table B-5 below:

Table B-5 Frequency distribution of income in screener survey

Income	Frequency	Percent	Cum.
Less than $25,000	1,931	22.56	22.56
$25,000 to $49,999	2,366	27.64	50.20
$50,000 to $99,999	2,749	32.11	82.31
$100,000 and higher	1,514	17.69	100.00
Total	8,560	100.00	

We applied a multinomial logit model and chose from the most likely category to compute the missing income values. The base category is “Less than $25,000”. The result is listed in Table D-6 below:

Table B-6 Income estimates of multinomial logit model for screener survey missing income imputation (base category "less than $25,000")

	$25,000 to $49,999	$50,000 to $99,999	$100,000 and higher
metro	0.112	0.501***	1.171***
	(-0.0891)	(0.0988)	(0.129)
gender	-0.155	-0.249**	-0.483***
	(0.0791)	(0.0847)	(0.104)
age	0.00607	0.00946**	0.993***
	(0.00320)	(0.00344)	(0.154)
white	0.343**	0.793***	0.972***
	(0.107)	(0.123)	(0.150)
eduyear	0.154***	0.346***	0.566***
	(0.0191)	(0.0200)	(0.0237)
fulltime	1.395***	1.361***	0.525
	(0.189)	(0.232)	(0.298)
parttime	0.163	-0.0721	-0.899**
	(0.181)	(0.225)	(0.308)
unemployment	-0.367	-0.936***	-2.134***
	(0.206)	(0.267)	(0.358)
home	0.190	0.186	-0.621
	(0.233)	(0.263)	(0.329)
couple	1.183***	2.054***	3.025***
	(0.102)	(0.121)	(0.195)
child5	-0.662*	-1.237*	0.743
	(0.490)	(0.796)	(0.619)
child17	0.450	1.116***	1.397***
	(0.233)	(0.248)	(0.383)
imm	0.398***	0.925***	2.114***
	(0.148)	(0.178)	(0.268)
ext	0.284	-0.269**	1.114***
	(0.234)	(0.336)	(0.411)
withchild	-0.107	-0.604*	-5.720*
	(0.591)	(0.681)	(2.878)
cc5	1.338***	2.186***	2.361***
	(0.255)	(0.247)	(0.300)
cc17	1.340***	2.533***	3.644***
	(0.175)	(0.178)	(0.235)
cc	1.115***	1.965***	2.922***
	(0.221)	(0.223)	(0.282)
wac5	0.406	0.803*	1.336**
	(0.304)	(0.404)	(0.516)
wac17	0.716***	1.997***	3.051***
	(0.203)	(0.203)	(0.287)
wac	0.822*	0.764	1.897***
	(0.332)	(0.393)	(0.495)

Table B-6: (cont'd)

	\$25,000 to \$49,999	\$50,000 to \$99,999	\$100,000 and higher
wa	1.337***	2.325***	3.404***
	(0.167)	(0.174)	(0.240)
_cons	-3.708***	-7.443***	-12.27***
	(0.353)	(0.430)	(0.596)
Log pseudolikelihood = -9799.3632		Number of obs = 8560	
		Wald chi2 = 1761.64	
		Prob > chi2 = 0.0000	
		Pseudo R2 = 0.1733	

Standard errors in parentheses
* p<0.05, ** p<0.01, *** p<0.001 .

Table B-7 Frequency distribution of imputed income in screener survey

Income	Frequency	Percent	Cum.
Less than \$25,000	255	24.73	24.73
\$25,000 to \$49,999	354	34.34	59.07
\$50,000 to \$99,999	347	33.66	92.73
\$100,000 and higher	75	7.27	100.00
Total	1,031	100.00	

Appendix C

The Importance of Partial Sites

Table C-1 Full information maximum likelihood (FIML) estimation results

		Model 1		Model 2		Model 3	
		With partial sites		Drop trips to partial sites		Drop person to partial sites	
Model Levels	Variables	Estimates	t Statistics	Estimates	t Statistics	Estimates	t Statistics
Beach Level	Travel Cost	-0.012***	-10.849	-0.022***	-7.924	-0.025***	-5.361
	Log(Length)	0.064***	7.260	0.068***	5.241	0.065***	2.953
	Temperature	0.022***	6.072	0.036***	8.525	0.040***	5.723
	Closure Days of 2010	-0.008***	-3.963	-0.014***	-3.858	-0.016***	-3.080
	LP Northeast	-0.046	-0.459	-0.332*	-1.943	-0.226	-0.766
	LP Mid-East	-0.519***	-5.429	-0.663***	-3.443	-0.924***	-2.968
	LP Southeast	-0.555***	-5.028	-0.593***	-2.484	-0.726**	-2.078
	LP Northwest	0.388***	5.431	0.519***	2.831	0.572**	2.168
	LP Mid-West	0.292***	3.743	0.471**	2.257	0.548**	1.956
	LP Southwest	0.024	0.330	0.049	0.269	0.199	0.716
Lake Level	Nesting Parameter	0.296***	12.871	0.405***	9.845	0.356***	5.643
Trip/No Trip Level	Nesting Parameter	0.453***	10.834	0.666***	8.204	0.718***	5.539
No Trip	Male	-0.186**	-2.064	-0.200**	-2.010	-0.309**	-2.151
	Age	-0.004	-1.278	0.000	0.086	0.005	0.813
	White	0.153	0.765	0.310	1.102	0.654	1.758
	Education Years	-0.028	-1.551	-0.053**	-2.287	0.010	0.284
	Full-Time Employed	0.120	1.258	-0.029	-0.233	0.174	0.948
	Retired	0.147	0.989	0.147	0.711	0.387	1.477
	Children under 17	0.122	1.513	0.080	0.654	0.306	1.371
	Constant	5.233***	11.861	7.047***	10.147	6.217***	6.206

Note: *10% significance level; **5% significance level; *** 1% significance level.

The standard errors of Model 1, Model 2 and Model 3 were bootstrapped 120 draws.

Table C-2 Welfare estimates of changing a beach in 2011 dollars per person

		Per Season			Per Season Per Trip					
					Season/Total Trip Change			Season/Site Trip Change		
		Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Closure of One Beach in the Region	Huron North	-0.160	-0.027	-0.012	93.474	46.883	40.479	25.643	18.203	14.134
	Huron South	-0.337	-0.123	-0.036	90.412	45.711	39.869	25.702	18.269	14.194
	St. Clair	-1.970	-0.887	-0.379	89.545	45.477	39.746	26.797	19.144	15.136
	Erie	-4.122	-1.785	-0.834	89.472	45.433	39.737	29.488	21.357	17.49
	Michigan North	-0.183	-0.052	-0.017	94.768	47.717	40.769	25.613	18.172	14.096
	Michigan Central	-1.585	-0.587	-0.230	92.562	46.663	40.367	25.768	18.294	14.21
	Michigan South	-0.835	-0.282	-0.129	91.902	46.421	40.403	25.739	18.269	14.192
Marginal Increase in Length of One Beach in the Region	Huron North	0.109	0.013	0.007	98.68	42.12	41.388	26.926	16.466	14.535
	Huron South	0.147	0.039	0.012	87.687	46.266	36.121	25.117	18.701	13.074
	St. Clair	1.123	0.348	0.145	89.414	46.113	39.484	28.468	20.622	16.297
	Erie	1.199	0.375	0.173	89.712	45.365	39.024	33.524	24.312	20.471
	Michigan North	0.079	0.017	0.006	100.08	35.302	152.01	26.503	13.564	46.58
	Michigan Central	0.481	0.133	0.058	93.125	46.229	40.985	26.178	18.332	14.638
	Michigan South	0.347	0.086	0.043	92.462	45.439	39.698	26.132	18.069	14.138

Table C-3 Welfare estimates of changing a beach in 2011 dollars (million) at state level

		Season (Millions)		
		Model 1	Model 2	Model 3
Closure of One Beach in the Region	Huron North	-0.675	-0.115	-0.052
	Huron South	-1.426	-0.519	-0.151
	St. Clair	-8.332	-3.751	-1.602
	Erie	-17.430	-7.547	-3.526
	Michigan North	-0.774	-0.220	-0.074
	Michigan Central	-6.702	-2.483	-0.974
	Michigan South	-3.529	-1.193	-0.547
Marginal Increase in Length of One Beach in the Region	Huron North	0.460	0.057	0.029
	Huron South	0.622	0.165	0.050
	St. Clair	4.747	1.472	0.612
	Erie	5.072	1.585	0.731
	Michigan North	0.334	0.070	0.026
	Michigan Central	2.032	0.563	0.247
	Michigan South	1.469	0.363	0.181

Table C-4 Estimated trips and welfare changes of closing all beaches on a great lake in 2011 dollars

Per Person	Number of Trips			Season			Season/Total Trip Change			Season/Lake Trip Change		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Erie	0.274	0.161	0.089	-11.493	-5.105	-2.663	89.363	45.38	39.738	41.991	31.74	30.001
St. Clair	0.424	0.263	0.137	-18.154	-8.484	-4.154	89.333	45.38	39.717	42.825	32.27	30.385
Huron	1.371	0.517	0.226	-64.002	-17.542	-7.138	90.551	45.66	39.871	46.683	33.93	31.547
Michigan	4.284	1.998	1.002	-284.25	-83.967	-37.623	90.680	45.93	39.980	66.360	42.03	37.561

State Level

	Number of Trips (Million)			Season (Million)		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Erie	1.157	0.68	0.375	-48.599	-21.59	-21.585
St. Clair	1.793	1.11	0.578	-76.764	-35.87	-35.874
Huron	5.797	2.19	0.957	-270.620	-74.17	-74.173
Michigan	18.112	8.45	4.236	-1201.900	-355.04	-355.040

Appendix D

Robustness Checks for Essay 1

Table D-1 Full Information maximum likelihood (FIML) estimation results for three model specifications

		Model 1		Model 2		Model 3	
Model Levels	Variables	Estimates	t Statistics	Estimates	t Statistics	Estimates	t Statistics
Beach Level	Travel Cost	-0.012***	-10.85	-0.017***	-10.24	-0.015***	-11.29
	Log(Length)	0.064***	7.26	0.093***	6.79	0.084***	6.94
	Temperature	0.022***	6.07	0.033***	8.07	0.029***	7.49
	Closure Days of 2010	-0.008***	-3.96	-0.013***	-3.73	-0.011***	-3.90
	LP Northeast	-0.046	-0.46	0.040	0.28	-0.019	-0.15
	LP Mid-East	-0.519***	-5.43	-0.588***	-4.48	-0.608***	-5.07
	LP Southeast	-0.555***	-5.03	-0.594***	-3.75	-0.628***	-4.49
	LP Northwest	0.388***	5.43	0.583***	5.36	0.502***	5.21
	LP Mid-West	0.292***	3.74	0.499***	4.04	0.399***	3.81
	LP Southwest	0.024	0.33	0.120	1.08	0.055	0.56
Lake Level	Nesting Parameter	0.296***	12.87	0.443***	10.78	0.392***	12.53
Trip/No Trip Level	Nesting Parameter	0.453***	10.83	0.691***	9.06	0.607***	10.49
No Trip	Male	-0.186**	-2.06	-0.165*	-1.85	-0.184**	-2.04
	Age	-0.004	-1.28	-0.002	-0.74	-0.002	-0.55
	White	0.153	0.77	0.196	1.01	0.217	1.10
	Education Years	-0.028	-1.55	-0.004	-0.21	-0.005	-0.30
	Full-Time Employed	0.120	1.26	0.174*	1.85	0.212**	2.11
	Retired	0.147	0.99	0.108	0.74	0.102	0.69
	Children under 17	0.122	1.51	0.178**	2.23	0.186**	2.24
	Constant	5.233***	11.86	6.337***	13.39	5.783***	12.59

Note: *10% significance level; **5% significance level; *** 1% significance level.

The standard errors of all three models were bootstrapped 120 draws.

Table D-1 (cont'd)

		Model 1		Model 2		Model 3	
Model Levels	Variables	Estimates	t Statistics	Estimates	t Statistics	Estimates	t Statistics
No Trip	Income (thousand)			-0.0065***	-7.60		
No Trip	Income 2nd Quartile					-0.326***	-2.32
	Income 3rd Quartile					-0.300	-1.35
	Income 4th Quartile					-0.758***	-5.23
Value of Log likelihood at convergence		-115617.139		-115023.4174		-115181.4551	

Note: *10% significance level; **5% significance level; *** 1% significance level

The standard errors of all three models were bootstrapped 120 draws.

Appendix E

Robustness Checks for Essay 2

Table E-1 Full Information maximum likelihood (FIML) estimation results for three additional model specifications for essay 2

Joint Estimation				Model 1		Model 2		Model 3	
Model Levels	Variables			Estimates	t Stats	Estimates	t Stats	Estimates	t Stats
RP	Beach Level	Common Variables	Travel Cost	-0.012***	-10.85	-0.012***	-11.3	-0.016***	-14.0
			Log(Length)	0.066***	7.26	0.068***	7.6	0.075***	8.0
			Lake Michigan					0.316***	2.0
			Lake Huron					0.227*	1.7
			Lake St. Clair					0.923***	6.0
			Temperature	0.022***	6.07	0.022***	6.1	0.019***	5.7
			Closure Days of 2010	-0.008***	-3.96	-0.008***	-4.0	-0.010***	-4.4
			LP Northeast	-0.049	-0.46	-0.040	-0.4		
			LP Mid-East	-0.524***	-5.43	-0.516***	-5.5		
			LP Southeast	-0.558***	-5.03	-0.548***	-5.0		
			LP Northwest	0.383***	5.43	0.391***	5.6		
			LP Mid-West	0.286***	3.74	0.295***	3.8		
			LP Southwest	0.019	0.33	0.028	0.4		
			North Region					0.255***	7.4
	Lake Level		Nesting Parameter	0.296***	12.87	0.296***	13.5	0.320***	13.1
	Trip/No Trip		Nesting Parameter	0.452***	10.83	0.452***	11.2	0.484***	10.7
	No Trip		Male	-0.186**	-2.06	-0.186**	-2.1	-0.191**	-2.1
			Age	-0.004	-1.28	-0.004	-1.3	-0.005	-1.4
			White	0.154	0.77	0.154	0.8	0.210	1.1
			Education Years	-0.028	-1.55	-0.028	-1.6	-0.024	-1.3
			Full-Time Employed	0.119	1.26	0.119	1.3	0.156	1.7
			Retired	0.147	0.99	0.147	1.0	0.161	1.1
			Children under 17	0.122	1.51	0.122	1.5	0.108	1.3
			Constant	5.221***	11.86	5.234***	11.8	5.589***	11.9

Note: *10% significance level; **5% significance level; *** 1% significance level.

The standard errors of Model 1-3 were corrected for clustering by bootstrapping with 120 draws.

Table E-1 (cont'd)

Model Levels		Variables	Model 1 Estimates	Model 1 t Stats	Model 2 Estimates	Model 2 t Stats	Model 3 Estimates	Model 3 t Stats
Scale		Scale Parameter	0.622***	7.57	0.933***	9.0	0.451***	5.4
SP	Algae in the water (base:high)	None	2.436***	10.79	2.052***	8.9	3.114***	9.1
		Low	2.195***	10.94	1.924***	8.3	2.872***	9.3
		Moderate	1.823***	10.28	1.336***	7.6	2.342***	9.0
	Algae on the shore (base:high)	None	2.107***	9.07	1.515***	7.6	2.753***	8.1
		Low	1.610***	7.28	1.534***	7.5	2.001***	6.9
		Moderate	0.944***	5.45	0.829***	5.5	1.222***	5.4
	Testing water for bacteria (base:Daily)	Never	-2.281***	-8.06	-1.548***	-7.3	-2.966***	-8.0
		Monthly	-0.379**	-2.21	-0.583***	-3.7	-0.545***	-2.5
		Weekly	-0.533**	-3.53	-0.123	-1.0	-0.771***	-3.6
	Great Lake (base: Lake Erie)	Lake Michigan	1.834***	8.78				
		Lake Huron	0.727***	4.95				
		Lake St. Clair	-0.033	-0.23				

Likelihood Ratio Test

	Value of Log likelihood at convergence		
Likelihood Ratio Test	Model 1	Model 2	Model 3
Pooled Data (RP&SP)	-115617.14	-115617.14	-104106.51
RP data	-2125.97	-1974.87	-2125.97
SP data	-117773.26	-117642.16	-106340.33
LL ratio test	60.31	100.30	215.71
Chi squared at 5% significance level	3.84	3.84	9.49

Note: *10% significance level; **5% significance level; *** 1% significance level

The standard errors of Model 1-3 were corrected for clustering by bootstrapping with 120 draws

Appendix F

Robustness Checks for Essay 3

Table F-1 Heckman model estimation results

	Variables	Model 1	Model 2	Model 3	Model 4
Spending	Log(distance)	67.64*** (19.92)	77.69*** (19.96)	-	-
	age	5.02** (2.53)	5.08** (2.53)	5.90** (2.64)	6.33** (2.65)
	eduyear	-22.05* (11.67)	-20.99* (11.65)	-21.58* (12.19)	-19.44* (12.25)
	income	1.19*** (0.44)	1.21*** (0.44)	1.27*** (0.47)	1.28*** (0.47)
	distance	-	-	0.18 (0.15)	0.25 (0.14)
Respond	distance	0.003*** (0.001)	-	0.004*** (0.001)	-
	Log(distance)	-	0.12** (0.06)	-	0.15*** (0.53)
Log-likelihood value		-1338.4	-1340.7	-1343.0	-1346.2
AIC		2718.8	2723.4	2728.0	2734.3
BIC		2797.5	2802.2	2806.7	2813.0

Note: *10% significance level; **5% significance level; *** 1% significance level;

Only distance variables and significant variables are shown here; all other insignificant variables maintain same as Table 3-6.

Appendix G

Spending By Categories in Essay 3

Appendix G Provides detailed information of Michigan beachgoers' spending by categories. For day trips, Table G-1 shows that beachgoers spent most on gas and oil, followed by groceries and take-out food or drink. Restaurants and bars take the third place in beachoger' spending for day trips.

Table G-1 The spending per party by categories for day trip

Spending categories	Mean ($)	Standard Deviation	Min	Max
Gas and Oil	12.32	13.52	0.00	90.00
Car rental, airfare, taxi, bus	0.66	5.78	0.00	60.00
Restaurant and Bar	8.69	15.20	0.00	75.00
Groceries and Take-out food/drink	9.92	20.85	0.00	200.00
Park Access fee	3.14	5.27	0.00	30.00
Entertainment fees	1.21	9.66	0	100
Sporting goods	2.41	16.61	0	150
Clothing	3.38	19.83	0	175
Souvenirs/gifts/postcards	0.79	5.35	0	50
Others	0.88	4.27	0	35

For overnight trips spending within 35 miles of the destination,, Table G-2 shows that beachgoers spent the most on hotels, followed by restaurants and bars. Spending on groceries and take-out food/drink is half the amount of the spending at restaurants and bars. Spending on gas and oil, and on campgrounds also take a relatively important portion of the total spending per party.

Table G-2 Spending per party within 35 miles of the destination by categories for overnight trips

Spending categories	Mean ($)	Standard Deviation	Min	Max
Hotel	281.17	363.44	0	1250
Campground	50.07	114.44	0	430
Gas and Oil	42.77	29.95	0	100
Car rental, airfare, taxi, bus	0.00	0.00	0	0
Restaurant and Bar	115.00	119.59	0	400
Groceries and Take-out food/drink	56.70	57.43	0	150
Park Access fee	3.20	4.99	0	12
Entertainment fees	12.17	30.95	0	150
Sporting goods	2.83	8.48	0	40
Clothing	26.67	77.39	0	400
Souvenirs/gifts/postcards	26.00	48.31	0	200
Others	2.83	9.62	0	40

Economic impact studies typically focus on the money spent in the local region only, in order to measure the corresponding impacts for that region. We followed Stynes (1997) to differentiate the overnight trips' spending into two types, one is the expenditures spent with 35 miles of the destination, which has shown in Table G-3; the other is the expenditures spent outside 35 miles of the destination. A comparison of Table G-2 and Table G-3 shows that beachgoers make most of their expenditures in the local region.

Table G-3 Spending per party outside 35 miles of the destination by categories for overnight trips

Spending categories	Mean ($)	Standard Deviation	Min	Max
Hotel	4.50	24.65	0	135
Campground	1.17	6.39	0	35
Gas and Oil	29.67	36.41	0	120
Car rental, airfare, taxi, bus	0.00	0.00	0	0
Restaurant and Bar	26.17	92.91	0	500
Groceries and Take-out food/drink	27.43	63.28	0	300
Park Access fee	0.33	1.83	0	10
Entertainment fees	0.00	0.00	0	0
Sporting goods	1.00	4.03	0	20
Clothing	5.83	27.61	0	150
Souvenirs/gifts/postcards	0.00	0.00	0	0
Others	3.00	16.43	0	90

Appendix H

2014 Michigan Beach Visitor Spending Survey

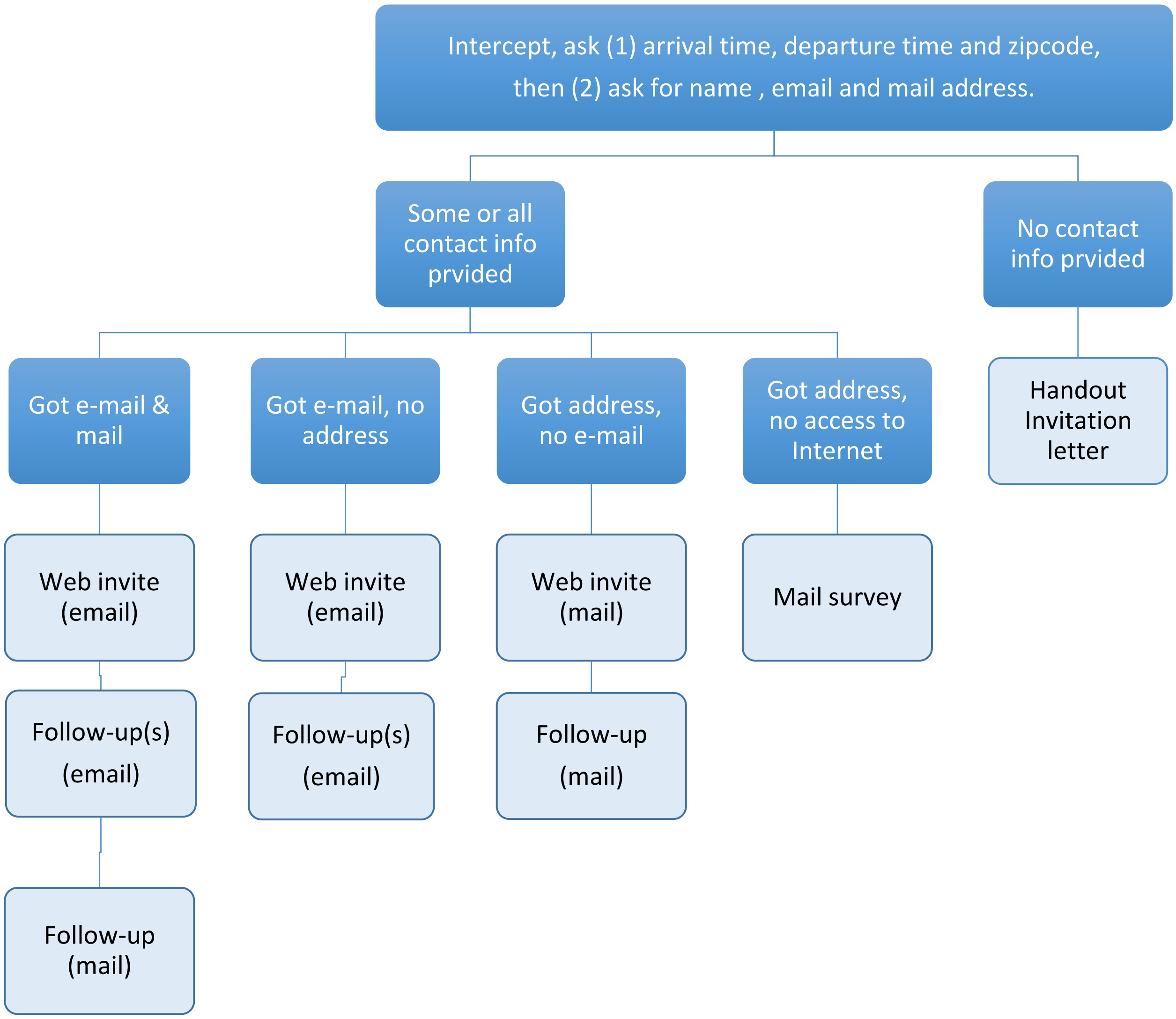
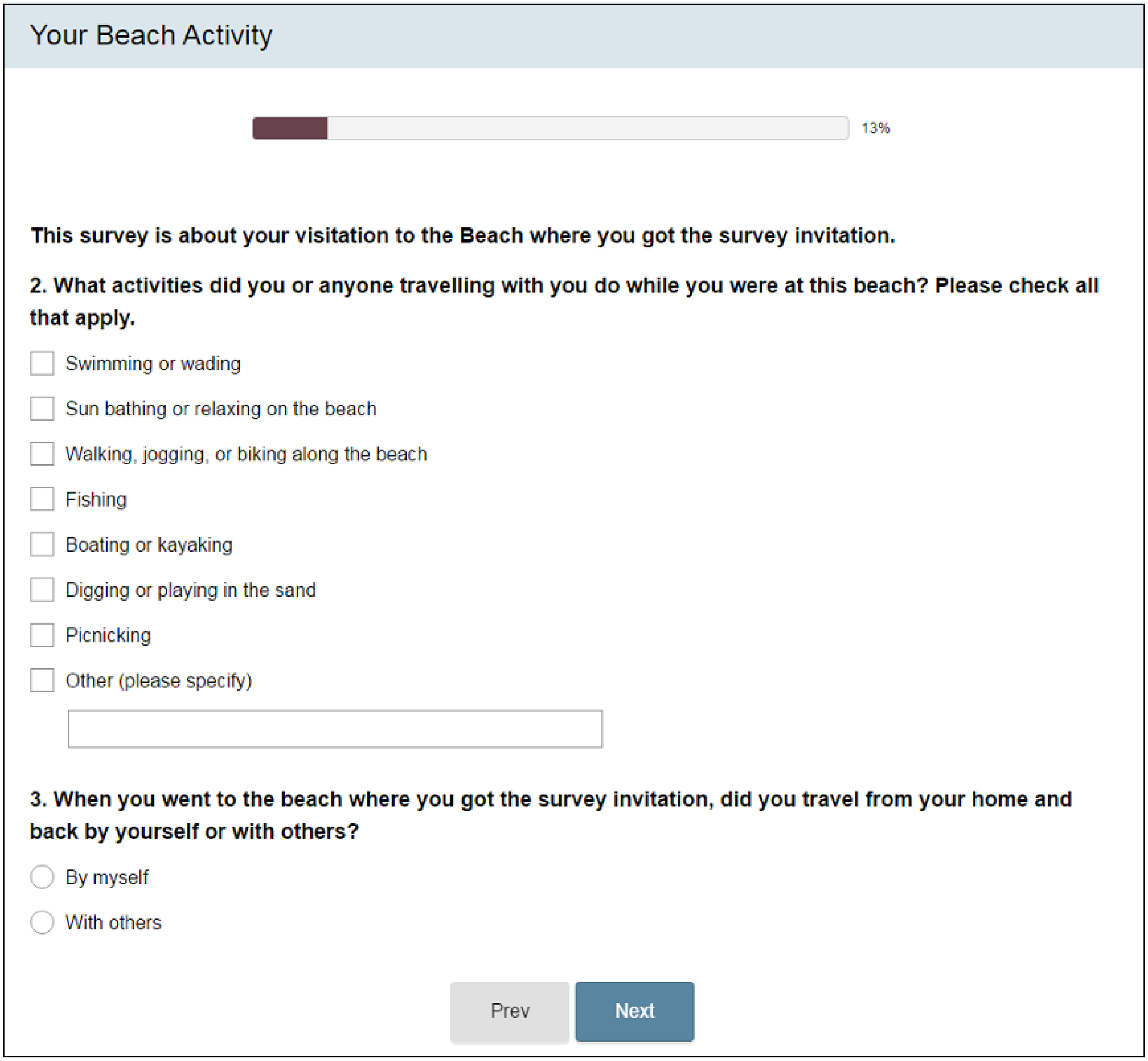


Figure H-1 Interview flow chart

Appendix I

Beach spending web survey instruments



Your Beach Activity

13%

This survey is about your visitation to the Beach where you got the survey invitation.

2. What activities did you or anyone travelling with you do while you were at this beach? Please check all that apply.

- [ ] Swimming or wading
- [ ] Sun bathing or relaxing on the beach
- [ ] Walking, jogging, or biking along the beach
- [ ] Fishing
- [ ] Boating or kayaking
- [ ] Digging or playing in the sand
- [ ] Picnicking
- [ ] Other (please specify)

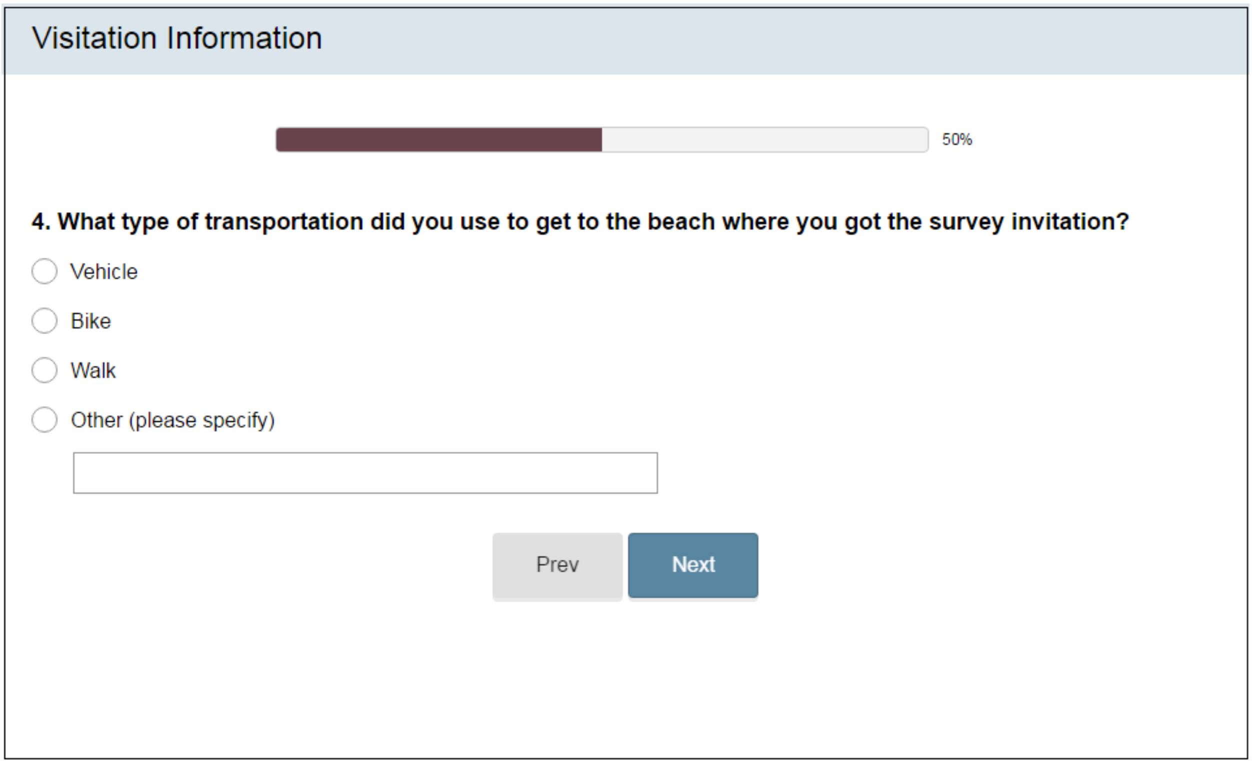
3. When you went to the beach where you got the survey invitation, did you travel from your home and back by yourself or with others?

- By myself
- With others

Prev Next

Figure I-1 Beach spending web survey

Figure I-1-(cont'd)



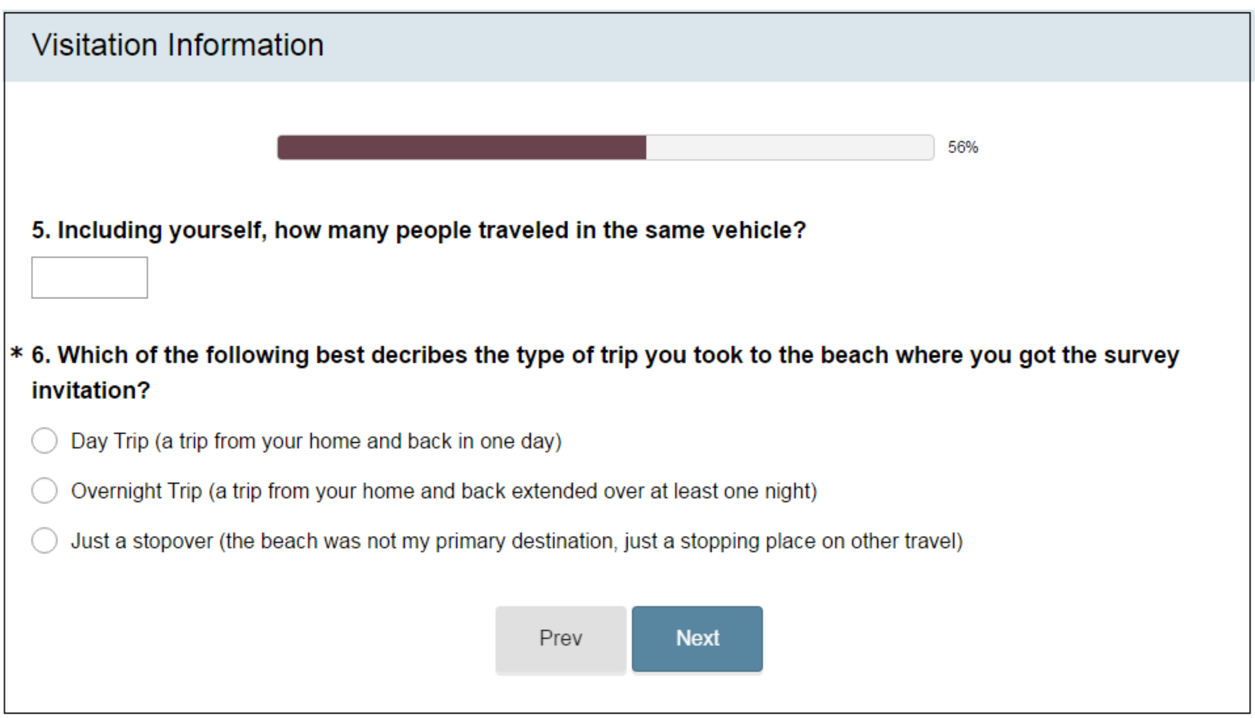
Visitation Information

50%

4. What type of transportation did you use to get to the beach where you got the survey invitation?

- ◯ Vehicle
- ◯ Bike
- ◯ Walk
- ◯ Other (please specify)

Prev Next



Visitation Information

56%

5. Including yourself, how many people traveled in the same vehicle?

*** 6. Which of the following best decribes the type of trip you took to the beach where you got the survey invitation?**

- ◯ Day Trip (a trip from your home and back in one day)
- ◯ Overnight Trip (a trip from your home and back extended over at least one night)
- ◯ Just a stopover (the beach was not my primary destination, just a stopping place on other travel)

Prev Next

Figure I-1-(cont'd)

Expenditure for Day Trip

63%

* **7. Please enter the dollar amount of expenditures <u>you and the people that traveled in the same vehicle as you</u> made for this trip to <u>the beach where you got the survey invitation</u>. Do this for each category below. Please be as accurate as possible - If unsure, provide your best estimate. If you made no expenditures for a spending category, please enter a "0".**

Gas and oil (auto, RV, boat, etc)

Car Rental, Airfares, Taxi, Bus

Restaurants and Bars

Groceries and Take-out Food/Drink

Park Access Fee

Entertainment fees (boat renting, fishing, etc)

Sporting goods (volleyball, life vest, etc)

Clothing

Souvenirs/gifts/postcards

Other

Prev Next

Figure I-1-(cont'd)

Your Trip Frequency

88%

8. Please think about the beach where you got the survey invitation.

Including the time we met, how many trips have you taken to this beach since Memorial Day weekend (May 24th, 2014)?

9. Please think about any Great Lakes beaches in Michigan.

Including the time we met, how many trips have you taken to any Great Lakes beaches since Memorial Day weekend (May 24th, 2014)?

Prev Next

About You

94%

10. Who is filling out this survey?

- The person the invitation was given/addressed to
- Another household member
- Someone else

11. What is your gender?

- Male
- Female

12. In what year were you born?

Figure I-1-(cont'd)

13. What is the highest degree or level of schooling you have completed?

- ◯ Less than High School
- ◯ High School or equivalent
- ◯ Some College, no degree
- ◯ Associate's degree
- ◯ Bachelor's degree
- ◯ Graduate or Professional degree

14. What is the zip code of the place you live?

15. What is your current employment status?

- ☐ Employed Full Time
- ☐ Employed Part Time
- ☐ Unemployed
- ☐ Stay at home parent
- ☐ Retired
- ☐ Student

Figure I-1-(cont'd)

16. Do any of the following live in your household?
(check all that apply)

- [ ] Spouse or significant other
- [ ] Children age 5 and under
- [ ] Children age 6-17
- [ ] Other immediate family
- [ ] Extended family or other adults
- [ ] None of these

17. What is your approximate annual household income?
(check one)

- ○ Less than $24,999
- ○ $25,000 to $34,999
- ○ $35,000 to $49,999
- ○ $50,000 to $74,999
- ○ $75,000 to $99,999
- ○ $100,000 to $149,999
- ○ $150,000 to $199,999
- ○ $200,000 or more

Prev Next

Thank You

100%

Thank you for completing this survey.

Prev Done

Figure I-1-(cont'd)

Lodging for Overnight Trip

69%

7. Where did you stay during your visit?

- ◯ Stayed at hotel or motel
- ◯ Rented a house, condo, or apartment
- ◯ Campground
- ◯ Stayed in your own property
- ◯ Stayed at family or friends' place
- ◯ Other (please specify)

Prev Next

Figure I-1-(cont'd)

Expenditure for Overnight Trip

75%

* 8. **Expenditures made *Within 35 Miles* of the Beach Destination**

For the beach where you got the survey invitation, please enter the dollar amount of expenditures you and the people that traveled in the same vehicle as you made Within 35 miles of the beach destination. Do this for each category below. Please be as accurate as possible - If unsure, provide your best estimate. If you made no expenditures for a spending category, please enter a "0".

Hotels, motels, Cabins

Campground Fees

Gas and oil (auto, RV, boat, etc)

Car Rental, Airfares, Taxi, Bus

Restaurants and Bars

Groceries and Take-out Food/Drink

Park Access Fee

Entertainment fees (boat renting, fishing, etc)

Sporting goods (volleyball, life vest, etc)

Clothing

Souvenirs/gifts/post cards

Other

* 9. **Expenditures made *Outside 35 Miles* of the Beach Destination**

For the beach where you got the survey invitation, please enter the dollar amount of expenditures you and the people that traveled in the same vehicle as you made Outside 35 miles of the beach destination. Do this for each category below. Please be as accurate as possible - If unsure, provide your best estimate. If you made no expenditures for a spending category, please enter a "0".

*If you traveled less than 35 miles to reach the beach destination, enter a "0" for all categories.

Hotels, motels, Cabins

Campground Fees

Gas and oil (auto, RV, boat, etc)

Car Rental, Airfares, Taxi, Bus

Restaurants and Bars

Groceries and Take-out Food/Drink

Park Access Fee

Entertainment fees (boat renting, fishing, etc)

Sporting goods (volleyball, life vest, etc)

Clothing

Souvenirs/gifts/post cards

Other

Prev Next

Figure I-1-(cont'd)

Expenditure for Your Stopover

44%

* 6. **Expenditures made *Within 35 Miles* of the Beach**

For the beach where you got the survey invitation, please enter the dollar amount of expenditures you made Within 35 miles of the beach. Do this for each category below. Please be as accurate as possible - If unsure, provide your best estimate. If you made no expenditures for a spending category, please enter a "0".

Hotels, motels, Cabins	
Campground Fees	
Gas and oil (auto, RV, boat, etc)	
Car Rental, Airfares, Taxi, Bus	
Restaurants and Bars	
Groceries and Take-out Food/Drink	
Park Access Fee	
Entertainment fees (boat renting, fishing, etc)	
Sporting goods (volleyball, life vest, etc)	
Clothing	
Souvenirs/gifts/postcards	
Other	

Prev Next

MICHIGAN STATE
UNIVERSITY

«DATE»

Your help is needed with a study of Great Lakes beaches. The study is being conducted by Michigan State University. The results from this survey will help agencies make beach management decisions that better reflect the needs of people that visit beaches in Michigan.

You are part of a small sample of people being asked about their beach activities. Your answers are needed to help ensure the results accurately represent the people who go to beaches in Michigan. Your answers are strictly confidential.

Please use the link provided below to visit the survey website. The web-based survey will take less than **5** minutes to complete.

To access the survey on the internet, please type the following into the address bar of your web browser:

Web address: www.research.net/s/GLbeach?c=id

(For help finding the address bar of your browser, see the back of this letter.)

If you have any questions about this study, contact me at 301b Agriculture Hall, MSU, East Lansing, MI 48824; 517-355-1692, MIstudy@msu.edu.

Thank you,

Professor Frank Lupi

DEPARTMENT OF AGRICULTURAL, FOOD AND RESOURCE ECONOMICS

Frank Lupi
Professor

301b Agriculture Hall
East Lansing, MI 48824-1039

517-355-1692
MIstudy@msu.edu

MSU is an affirmative-action, equal-opportunity employer

Figure I-2 Invitation letter (distributed on site if no contact information was received)

Answers to Frequently Asked Questions

How was I selected?
Three public beach sites were selected to represent Michigan's Great Lakes beaches. You are part of a small group selected to participate in the survey about Great Lakes beaches.

Will I be contacted about other surveys from you?
No, this is the only survey we will ask you to take. We know that you are busy and greatly appreciate your help with this important research project.

Why does this survey matter?
There is very little sound information about the economic benefits of beach recreation. Yet, decisions must get made about how to better manage public beaches. The information this survey gathers on beach activity will facilitate the protection and development of Great Lakes beaches.

Who sees my answers?
Your responses are saved directly into a database that does not contain your name or address. Personal information is only used to manage the mailing of survey invitations.

How is my privacy protected?
Your answers are kept separately from our mailing list. Our mailing list and data are stored on password protected computers in locked offices. Everyone who works on the survey has completed training and signed an oath saying they will not share any information. The addresses will be destroyed when the survey is complete.

The web address does not work
The survey web address needs to be typed exactly as printed on the letter, and it needs to be typed in the address bar of your browser. (To find the address bar see next question.)

How do I find the address bar?
The address must be typed into the address bar of your web browser. The survey is on a secure website which is not available through an internet search. So you need to type the exact address in the address bar, not in a search bar. For a brief video on the difference between doing a search and the using the address bar, search the internet for the following "*youtube where is the address bar*"

How do I get help with web survey access or other problems?
If you have trouble accessing the web survey or if you have other technical issues you can contact our research team by email (MIstudy@msu.edu) or by phone (517-355-1692).

Figure I-3 Invitation letter_Back page

DATE

Dear «Beach name» Beachgoer,

It was a pleasure to meet you on the beach this week. At that time we asked for your help with a web survey on Great Lakes beaches. If you have already completed the web survey, *thank you very much!* If not, please do so today.

The survey should take less than 5 minutes to complete. Your answers are strictly confidential.

To access the survey on the internet, please click on the following link or type it directly into the address bar of your web browser:

Web address: «web»

Your answer will help agencies make beach management decisions that better reflect the needs of people that visit beaches in Michigan. Because you are part of a small sample of people being asked about their beach activities, we need your answers to be sure our results accurately represent beachgoers.

If you have any questions about this study, contact me at 301b Agriculture Hall, MSU, East Lansing, MI 48824; 517-355-1692, MIstudy@msu.edu.

Thank you very much for helping with this important study.

Sincerely,
Professor Frank Lupi
Michigan State University

MICHIGAN STATE
UNIVERSITY

Figure I-4 Follow-up email reminder: First wave

DATE

Dear «Beach name» Beachgoer,

Two days ago, we sent you an email with a link to a web survey on Great Lakes beaches. To our knowledge, we have not yet received your completed web survey. *If you have responded, thank you!*

We are writing to you again because **your input is vital to our study**! Your answer will help agencies make beach management decisions that better reflect the needs of people that visit beaches in Michigan. Because you are part of a small sample of people being asked about their beach activities, we need your answers to be sure our results accurately represent beachgoers.

The survey should take less than **5** minutes to complete. Your answers are strictly confidential.

To access the survey on the internet, please click on the following link or type it directly into the address bar of your web browser:

Web address: «web»

If you have any questions about this study, contact me at 301b Agriculture Hall, MSU, East Lansing, MI 48824; 517-355-1692, MIstudy@msu.edu.

Thank you very much for helping with this important study.

Sincerely,
Professor Frank Lupi
Michigan State University

MICHIGAN STATE
UNIVERSITY

Figure I-5 Follow-up email reminder: Second wave

DATE

Dear «Beach name» Beachgoer,

Last week, we met you on the beach and sent you two emails with a link to a web survey on Great Lakes beaches. *If you have responded and completed the survey, thank you very much!* If not, please do so today, as our web survey will close soon.

We have undertaken this study to help agencies make beach management decisions that better reflect the needs of people that visit beaches in Michigan.

Your input is important, because you are part of a small sample of people being asked about their beach activities. Your answers are needed to help ensure the results accurately represent the people who go to beaches in Michigan.

The survey should take less than **5** minutes to complete. Your answers are strictly confidential.

To access the survey on the internet, please click on the following link or type it directly into the address bar of your web browser:

Web address: «web»

If you have any questions about this study, contact me at 301b Agriculture Hall, MSU, East Lansing, MI 48824; 517-355-1692, MIstudy@msu.edu.

Thank you very much for helping with this important study.

Sincerely,
Professor Frank Lupi
Michigan State University

MICHIGAN STATE
UNIVERSITY

Figure I-6 Follow-up email reminder: Third wave

DATE

Dear «Beach name» Beachgoer,

Last month, we met you on the beach and sent you emails with a link to a web survey on Great Lakes beaches. *If you have completed the survey, thank you very much!* If not, we still need your help.

This is the last time we will be contacting you. Your answers are vital for us to accurately represent people who go to beaches in Michigan. The survey should take less than **5** minutes to complete. Your answers are strictly confidential.

To access the survey on the internet, please click on the following link or type it directly into the address bar of your web browser:

Web address: «web»

If you have any questions about this study, contact me at 301b Agriculture Hall, MSU, East Lansing, MI 48824; 517-355-1692, lupi@msu.edu.

Thank you for helping with this important study.

Sincerely,
Professor Frank Lupi
Michigan State University

MICHIGAN STATE
UNIVERSITY

Figure I-7 Follow-up email reminder: Fourth wave

MICHIGAN STATE
UNIVERSITY

DEPARTMENT OF
AGRICULTURAL,
FOOD
AND RESOURCE
ECONOMICS

Frank Lupi
Professor

301b Agriculture Hall
East Lansing, MI 48824-1039

517-355-1692
MIstudy@msu.edu

DATE

«street»
«city», «state» «Zipcode»

Dear «Beach name» Beachgoer,

It was a pleasure to meet you on the beach last week. At that time we asked for your help with a web survey on Great Lakes beaches. If you have already completed the web survey, *thank you very much!* If not, please do so today.

The survey should take less than **5** minutes to complete. Your answers are strictly confidential.

To access the survey on the internet, please type the following into the address bar of your web browser:

Web address: «web»

(For help finding the address bar of your browser, see the back of this letter.)

Your answer will help agencies make beach management decisions that better reflect the needs of people that visit beaches in Michigan. Because you are part of a small sample of people being asked about their beach activities, we need your answers to be sure our results accurately represent beachgoers.

If you have any questions about this study, contact me at 301b Agriculture Hall, MSU, East Lansing, MI 48824; 517-355-1692, MIstudy@msu.edu.

Thank you very much for helping with this important study.

Sincerely,

Professor Frank Lupi

MSU is an affirmative-action, equal-opportunity employer

Figure I-8 Follow-up mail reminder: First wave

MICHIGAN STATE
UNIVERSITY

DATE

«street»
«city», «state» «Zipcode»

Dear «Beach name» Beachgoer,

Last month, we mailed you an invitation letter with a link to a web survey on Great Lakes beaches. To our knowledge, we have not yet received your completed web survey. *If you have responded, thank you!*

We are writing to you again because **your input is vital to our study**! Your answer will help agencies make beach management decisions that better reflect the needs of people that visit beaches in Michigan. Because you are part of a small sample of people being asked about their beach activities, we need your answers to be sure our results accurately represent beachgoers.

The survey should take less than **5** minutes to complete. Your answers are strictly confidential.

To access the survey on the internet, please type the following into the address bar of your web browser:

Web address: «web»

(For help finding the address bar of your browser, see the back of this letter.)

If you have any questions about this study, contact me at 301b Agriculture Hall, MSU, East Lansing, MI 48824; 517-355-1692, MIstudy@msu.edu.

Thank you very much for helping with this important study.

Sincerely,

Professor Frank Lupi

DEPARTMENT OF AGRICULTURAL, FOOD AND RESOURCE ECONOMICS

Frank Lupi
Professor

301b Agriculture Hall
East Lansing, MI 48824-1039

517-355-1692
MIstudy@msu.edu

Figure I-9 Follow-up mail reminder: Second wave

MICHIGAN STATE
UNIVERSITY

«DATE»

«street»
«city», «state» «Zipcode»

Dear «beach» Beachgoer,

Last month, we met you on the beach and mailed you two invitation letters with a link to a web survey on Great Lakes beaches. *If you have responded and completed the survey, thank you very much!* If not, please do so today, as our web survey will close soon.

We have undertaken this study to help agencies make beach management decisions that better reflect the needs of people that visit beaches in Michigan.

Your input is important, because you are part of a small sample of people being asked about their beach activities. Your answers are needed to help ensure the results accurately represent the people who go to beaches in Michigan.

The survey should take less than **5** minutes to complete. Your answers are strictly confidential.

To access the survey on the internet, please type the following into the address bar of your web browser:

Web address: «web»

(For help finding the address bar of your browser, see the back of this letter.)

If you have any questions about this study, contact me at 301b Agriculture Hall, MSU, East Lansing, MI 48824; 517-355-1692, MIstudy@msu.edu.

Thank you very much for helping with this important study.

Sincerely,

Professor Frank Lupi

DEPARTMENT OF
AGRICULTURAL,
FOOD
AND RESOURCE
ECONOMICS

Frank Lupi
Professor

301b Agriculture Hall
East Lansing, MI 48824-1039

517-355-1692
MIstudy@msu.edu

MSU is an affirmative-action, equal-opportunity employer

Figure I-10: Follow-up mail reminder: Third wave

Answers to Frequently Asked Questions

How was I selected?
Three public beach sites were selected to represent Michigan's Great Lakes beaches. You are part of a small group selected to participate in the survey about Great Lakes beaches.

Will I be contacted about other surveys from you?
No, this is the only survey we will ask you to take. We know that you are busy and greatly appreciate your help with this important research project.

Why does this survey matter?
There is very little sound information about the economic benefits of beach recreation. Yet, decisions must get made about how to better manage public beaches. The information this survey gathers on beach activity will facilitate the protection and development of Great Lakes beaches.

Who sees my answers?
Your responses are saved directly into a database that does not contain your name or address. Personal information is only used to manage the mailing of survey invitations.

How is my privacy protected?
Your answers are kept separately from our mailing list. Our mailing list and data are stored on password protected computers in locked offices. Everyone who works on the survey has completed training and signed an oath saying they will not share any information. The addresses will be destroyed when the survey is complete.

The web address does not work
The survey web address needs to be typed exactly as printed on the letter, and it needs to be typed in the address bar of your browser. (To find the address bar see next question.)

How do I find the address bar?
The address must be typed into the address bar of your web browser. The survey is on a secure website which is not available through an internet search. So you need to type the exact address in the address bar, not in a search bar. For a brief video on the difference between doing a search and the using the address bar, search the internet for the following "*youtube where is the address bar*"

How do I get help with web survey access or other problems?
If you have trouble accessing the web survey or if you have other technical issues you can contact our research team by email (MIstudy@msu.edu) or by phone (517-355-1692).

Figure I-11 Follow-up mail reminders: Back page

Appendix J

Beach sites choice for 2014 beach visitor spending survey

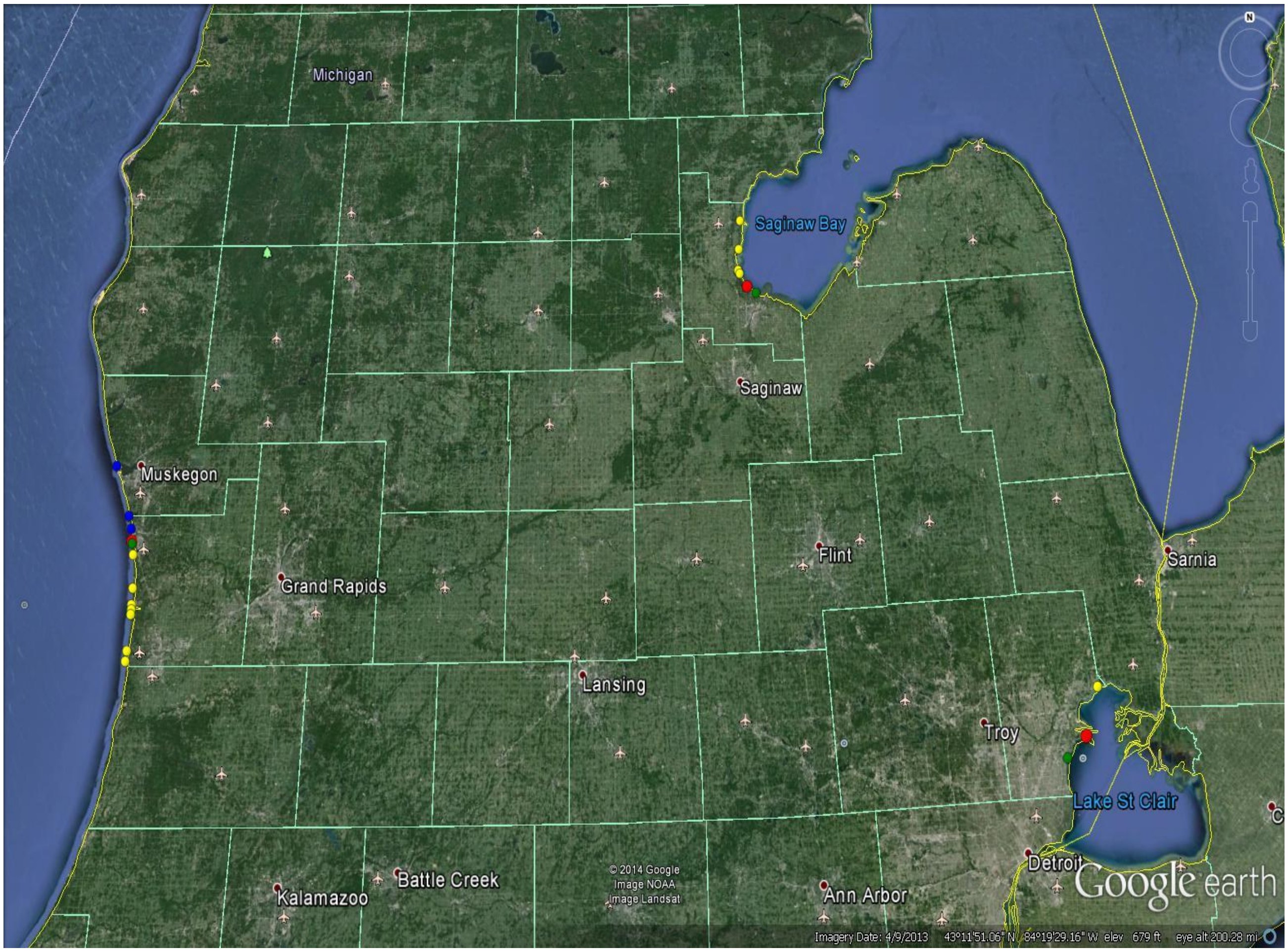


Figure J- 1 Beach sites choice for 2014 beach visitor spending survey

Note: Data sources: DEQ website http://www.deq.state.mi.us/beach/Default.aspx

Red dots on the maps are the targeted beach sites, from west to east, they are:

- *Grand Haven State Park*
- *H.C.M.A. - Lake St. Clair Metropark Beach*
- *Bay City State Recreation Area*

Green dots are the nearest next sites away from the targeted sites. Yellow and blue dots are other possible public beaches nearby. We pinned them on the map, so the surveyor had the choices to go nearby beaches to recruit more beachgoers. However, it turned out that going to multiple choices in the same day caused more time spending on commuting instead of interviewing more people, so the nearby beaches have never used.

.

Appendix J

Response Rate

Table J-1 describes the overall disposition of the sample

Table J-1 Response and disposition for the beach visitor survey

Stage	Disposition	Sub-categories	Totals
1, All	Total Intercepted		336
1	Stage 1 Refusal (refused to give zip and refused to take letter)		2
1	Stage 1 Response (gave zip code and took letter)		334
	Gave zip, but refused to take letter	8	
	Took invitation letter only	42	
	Gave e-mail only	234	
	Gave mail only	39	
	Gave e-mail and mail address	7	
	Gave mail only but had no internet access	4	
2	Stage 2 Respondents		170
	Responded (completed all spending & most demographic questions)	151	
	Responded (complete all spending but skipped demographic questions)	6	
	Responded (incomplete spending data)[19]	13	
2	Stage 2 Refusals (refused letter above)		8
2	Stage 2 Non-respondents		156
All	Overall Respondents		170
All	Overall Refusal + Non-respondents (=2+8+156)		166
All	Overall Response Rate (=170/336)		50.6%

Overall, the refusal rate for giving zip code was 0.6%, calculated as 2/336; the refusal rate overall, which includes not taking invitation letter was 2.98%, calculated as 10/336. 169 individuals logged on the web survey and answered some questions, and 1 individual sent back the mail survey, therefore, the overall survey response rate was 50.6% calculated as 170/336.

19 13 dropped out of before answering expenditures questions

Considering the non-responded items, 151 individuals completed all spending questions and most demographic questions. After imputed the missing demographic information(see Table J-7 Missing imputation method), we obtained 6 more useable responses, which leads to 157 effective observations, with an effect response rate for completed data of 47%, which was used in Essay 2. The statistical summary of the response rate per visit, per site, and per mode are presented in the following tables. Key results are:

- Average interview number is 30 for each visit. However, Grand Haven was more popular, and was easier to recruit beachgoers.

Table J-2 Response rate per visit for Grand Haven

Order of visit	1st	2nd	3rd	4th	5th
Interview date	8/4/2014	8/8/2014	8/28/2014	9/6/2014	9/8/2014
Total interviewees	25	70	35	33	38
Respondents	10	40	17	13	20
Response Rate	0.4000	0.5714	0.4857	0.3939	0.5263

Table J-3 Response rate per visit for Saginaw Bay

Order of visit	1st	2nd	3rd
Interview date	8/6/2014 Wed	8/10/2014 Sun	8/13/2014 Wed
Total interviewees	18	30	10
Respondents	8	9	2
Response Rate	0.4444	0.3000	0.2000

Table J-4 Response rate per visit for St. Clair Metro Park

Order of visit	1st	2nd	3rd
Interview date	8/25/2014 Mon	8/26/2014 Tu	8/27/2014 Wed
Total interviewees	38	21	8
Respondents	15	10	5
Response Rate	0.3947	0.4762	0.6250

- Across the three sites, Saginaw Bay had the lowest response rate at 32.76%, followed by St. Clair with 46.01%, with Grand Haven being the most populous and having the highest response rate of 50.75%.

Table J-5 Response rate per site

Site	Grand Haven	Saginaw Bay	St. Clair	Total
Total samples*	201	58	67	326
Total response	102	19	30	151
Response rate	0.5075	0.3276	0.4478	0.4632

*Excludes 10 refusals.

- For each survey mode, Email reminders had the highest response rate of 55.60%, Mail reminders has 34.78%. No one responded if they refused to provide address and only accepted an invitation letter.

Table J-6 Response rate per survey mode

Survey mode	Web survey: Email Reminder	Web survey: Mail Reminder	Mail Survey	Accept only Invitation letter on sites	Total
Total sample*	241	46	4	42	333
Total response	134	16	1	0	151
Response rate	0.556	0.3478	0.25	0	0.4535

*7 interviewers provided both Mail and Email addresses, which increases the sample size from 326 to 333. Strictly speaking, sample here means the way to contact interviewee.

Table J-7 Missing survey data imputation method

Variable	Missing	Missing Imputation method
Gender	6	mean of all other respondents
Age	9	Census: median age
Race	6	mean of all other respondents
Education	6	mean of all other respondent's education year
Employment	6	mean of all other respondents
Household	6	mean of all other respondents
Income	14	Census: median income

Appendix K

Comparison of Spending Prediction Using Heckman vs. OLS

This section compares the spending results using Heckman method to OLS method. We found the spending was underestimated without correcting for response bias for out-of-sample prediction (see Table K-1). Therefore, the total spending by region was also underestimated using OSL method (see Table K-2 and Table K-3). Our finding reinforced Messonnier et al (2000)'s study, which estimated the willingness to pay for aquatic plant management in Lake Guntersville, Alabama. They also found the amount of non-fishers' willingness-to-pay was underestimated without correcting for nonresponse bias.

Table K-1 Predicted spending using 2011 Great Lakes Beaches Survey if a visit were to be taken to each of the 451 sites in the recreation demand model choice set (out-of-sample)

predicted spending[20] (per party)	Obs	Mean	Std. Dev.	Min	Max
Heckman model	1,144,187	439.53	134.70	0.00	940.08
OLS model	1,144,187	332.24	134.54	0.00	844.53

[20] The reported value for the predicted spending is the estimated spending that would be made if a trip was taken to the sites (i.e., the figure is not yet weighted by the probabilities of visiting the sites).

Table K-2 Economic impacts of beach visitation in 2014 dollars per person per season.

	Number of Trips (per person per season)	Total Spending by Region (per person per season)	
		Heckman	**OLS**
Huron North	0.68	99.51	74.49
Huron South	0.69	96.55	69.99
St. Clair	0.42	54.57	37.46
Erie	0.27	35.92	25.07
Michigan North	1.59	229.92	170.15
Michigan Central	1.72	248.80	183.98
Michigan South	0.97	140.95	104.42

Table K-3 Economic impacts of total spending by region in 2014 dollars at state level

State level	Number of Trips (millions)	Total Spending by Region (millions)	
		Heckman	**OLS**
Huron North	2.86	420.78	314.96
Huron South	2.93	408.26	295.94
St. Clair	1.79	230.74	158.38
Erie	1.16	151.90	106.00
Michigan North	6.73	972.19	719.45
Michigan Central	7.27	1052.00	777.92
Michigan South	4.11	596.01	441.51

BIBLIOGRAPHY

BIBLIOGRAPHY

Adamowicz, Wiktor, Louviere, Jordan, & Williams, Michael. (1994). Combining revealed and stated preference methods for valuing environmental amenities. *Journal of Environmental Economics and Management, 26*(3), 271-292.

Adamowicz, Wiktor, Swait, Joffre, Boxall, Peter, Louviere, Jordan, & Williams, Michael. (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*(1), 65-84.

Alexander, J. (2011). *Pandora's Locks: The Opening of the Great Lakes-St. Lawrence Seaway*. MSU Press.

Armstrong, J. S., & Overton, T. S. (1977). Estimating nonresponse bias in mail surveys. *Journal of marketing research*, 396-402.

Austin, J, and A. Steinman. (2015). *Michigan's Blue Economy Report*. Retrieved from http://michiganblueeconomy.org/wp-content/uploads/2015/03/Michigan-Blue-Economy-Report.pdf

Azevedo, Christopher D, Herriges, Joseph A, & Kling, Catherine L. (2003). Combining revealed and stated preferences: consistency tests and their interpretations. *American Journal of Agricultural Economics, 85*(3), 525-537.

Bell, Frederick W, & Leeworthy, Vernon R. (1986). An Economic Analysis of the Importance of Saltwater Beaches in Florida. *Florida Sea Grant Report 82, pp. 1-166.* Gainesville, University of Florida.

Bell, Frederick W, & Leeworthy, Vernon R. (1990). Recreational demand by tourists for saltwater beach days. *Journal of Environmental Economics and Management, 18*(3), 189-205.

Ben-Akiva, Moshe, & Morikawa, Takayuki. (1990). Estimation of switching models from revealed preferences and stated intentions. *Transportation Research Part A: General, 24*(6), 485-495.

Bergstrom, J. C., Teasley, R. J., Cordell, H. K., Souter, R., & English, D. B. (1996). Effects of reservoir aquatic plant management on recreational expenditures and regional economic activity. *Journal of Agricultural and Applied Economics, 28*(2), 409.

Bockstael, N., Hanemann, W., & Strand, I. (1986). Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models, report presented to the Environmental Protection Agency under cooperative agreement CR&811043&01&0. *Washington, DC*.

Bockstael, Nancy E, Hanemann, W Michael, & Kling, Catherine L. (1987). Estimating the value of water quality improvements in a recreational demand framework. *Water Resources Research*, *23*(5), 951-960.

Bockstael, Nancy E, Strand, Ivar E, & Hanemann, W Michael. (1987). Time and the recreational demand model. *American Journal of Agricultural Economics, 69*(2), 293-302.

Bower, B. T., & Turner, R. K. (1998). Characterising and analysing benefits from integrated coastal management (ICM). *Ocean & Coastal Management, 38*(1), 41-66.

Bradley, Mark A, & Daly, Andrew J. (1997). Estimation of logit choice models using mixed stated preference and revealed preference information. *Understanding Travel Behaviour in an Era of Change*, Oxford, OX, UK: Pergamon. 209-232.

Cameron, Trudy Ann. (1992). Combining contingent valuation and travel cost data for the valuation of nonmarket goods. *Land Economics*, 302-317.

Cameron, T. A., Shaw, W. D., & Ragland, S. (1999). Nonresponse bias in mail survey data: salience vs. endogenous survey complexity. *Valuing recreation and the environment. Revealed preference methods in theory and practice*, Edward Elgar, Cheltenham, UK, 217-251

Champ, P. A., & Bishop, R. C. (1996). Evidence on the accuracy of expenditures reported in recreational surveys. *Journal of Agricultural and Resource Economics*, 150-159.

Chen, Min. (2013). *Valuation of public Great Lakes beaches in Michigan*, PhD Dissertation, Michigan State University

Cook, P.S. (2011). *Impacts of visitor spending on the local economy: Sleeping Bear Dunes National Lakeshore, 2009*. Natural Resource Report. National Park Service, Fort Collins, Colorado.

Deisenroth, D. B., Loomis, J. B., & Bond, C. A. (2013). Using Revealed Preference Behavioral Models to Correctly Account for Substitution Effects in Economic Impact Analysis. *Journal of Regional Analysis & Policy, 43*(2), 157.

Dorfman, M. H., Stoner, N., & Rosselot, K. S. (2009). *Testing the waters: a guide to water quality at vacation beaches*: Natural Resources Defense Council San Francisco. Retrieved from www.nrdc.org/water/oceans/ttw/ttw2009.pdf.

Dwight, R. H., Brinks, M. V., SharavanaKumar, G., & Semenza, J. C. (2007). Beach attendance and bathing rates for Southern California beaches. *Ocean & Coastal Management, 50*(10), 847-858.

Dwight, R. H., Catlin, S. N., & Fernandez, L. M. (2012). Amounts and distribution of recreational beach expenditures in southern California. *Ocean & Coastal Management, 59*, 13-19.

Earnhart, Dietrich. (2001). Combining revealed and stated preference methods to value environmental amenities at residential locations. *Land Economics*, *77*(1), 12-29.

Egan, Kevin J, Herriges, Joseph A, Kling, Catherine L, & Downing, John A. (2009). Valuing water quality as a function of water quality measures. *American Journal of Agricultural Economics, 91*(1), 106-123.

English, D. B., & Bergstrom, J. C. (1994). The conceptual links between recreation site development and regional economic impacts. *Journal of Regional Science, 34*(4), 599-611.

Englin, Jeffrey, & Shonkwiler, J Scott. (1995). Modeling recreation demand in the presence of unobservable travel costs: toward a travel price model. *Journal of Environmental Economics and Management*, *29*(3), 368-377.

English, D. B. (1997). Effects of sample selection on estimates of economic impacts of outdoor recreation. *The Review of Regional Studies*, *27*(3), 219.

Greene, W. H. (2003). *Econometric analysis*: Pearson Education.

Freeman III, A Myrick, Herriges, Joseph A, & Kling, Catherine L. (2014). *The measurement of environmental and resource values: theory and methods*. Oxon: RFF Press.

Great Lakes Commission. (2009). *Healthy Great Lakes and St. Lawrence Beaches: A National Priority*. Retrieved from http://www.glc.org/glinvestment/pdf/CitiesInvest-Beach-FINAL-1up-lowres.pdf

Griffiths, C., Klemick, H., Massey, M., Moore, C., Newbold, S., Simpson, D., . . . Wheeler, W. (2012). US Environmental Protection Agency valuation of surface water quality improvements. *Review of Environmental Economics and Policy*, *6*(1), 130-146.

Haener, Michel Karen, Boxall, Peter C, & Adamowicz, Wiktor L. (2001). Modeling recreation site choice: Do hypothetical choices reflect actual behavior? *American Journal of Agricultural Economics*, *83*(3), 629-642.

Hamel, C., Herrmann, M., Lee, S. T., Criddle, K. R., & Geier, H. T. (2002). Linking sportfishing trip attributes, participation decisions, and regional economic impacts in Lower and Central Cook Inlet, Alaska. *The Annals of Regional Science*, *36*(2), 247-264.

Hanemann, M, Pendleton, L, Mohn, C, Hilger, J, Kurisawa, K, Layton, D, & Vasquez, F. (2004). Using revealed preference models to estimate the effect of coastal water quality on beach choice in Southern California. *University of California at Berkeley, report to the US National Oceanic and Atmospheric Administration.*

Hartmann, H. C. (1990). Climate change impacts on Laurentian Great Lakes levels. *Climatic Change, 17*(1), 49-67.

Hausman, Jerry A, Leonard, Gregory K, & McFadden, Daniel. (1995). A utility-consistent, combined discrete choice and count data model assessing recreational use losses due to natural resource damage. *Journal of Public Economics, 56*(1), 1-30.

Hensher, David A, & Bradley, Mark. (1993). Using stated response choice data to enrich revealed preference discrete choice models. *Marketing Letters, 4*(2), 139-151.

Hensher, David, Louviere, Jordan, & Swait, Joffre. (1998). Combining sources of preference data. *Journal of Econometrics, 89*(1), 197-221.

Hoagland, P., Anderson, D., Kaoru, Y., & White, A. (2002). The economic effects of harmful algal blooms in the United States: estimates, assessment issues, and information needs. *Estuaries, 25*(4), 819-837.

Hoyos, David, & Riera, Pere. (2013). Convergent validity between revealed and stated recreation demand data: Some empirical evidence from the Basque Country, Spain. *Journal of Forest Economics, 19*(3), 234-248.

Huang, Ju-Chin, Haab, Timothy C, & Whitehead, John C. (1997). Willingness to pay for quality improvements: should revealed and stated preference data be combined? *Journal of Environmental Economics and Management, 34*(3), 240-255.

Huang, Ju-Chin, Poor, P Joan, & Zhao, Min Qiang. (2007). Economic valuation of beach erosion control. *Marine Resource Economics*, 221-238.

Hutt, C. P., Hunt, K. M., Steffen, S. F., Grado, S. C., & Miranda, L. (2013). Economic values and regional economic impacts of recreational fisheries in Mississippi reservoirs. *North American Journal of Fisheries Management*, *33*(1), 44-55.

Hynes, Stephen, Hanley, Nick, & Garvey, Eoghan. (2007). Up the proverbial creek without a paddle: Accounting for variable participant skill levels in recreational demand modelling. *Environmental and Resource Economics, 36*(4), 413-426.

Jeon, Hocheol. (2014). *Three Essays on Environmental Economics*. PhD Dissertation, Iowa State University.

Kannel, Prakash Raj, Lee, Seockheon, Lee, Young-Soo, Kanel, Sushil Raj, & Khan, Siddhi Pratap. (2007). Application of water quality indices and dissolved oxygen as indicators for river water classification and urban impact assessment. *Environmental Monitoring and Assessment, 132*(1-3), 93-110.

Keeler, Bonnie L, Polasky, Stephen, Brauman, Kate A, Johnson, Kris A, Finlay, Jacques C, O'Neill, Ann, Kovacs, kent, & Dalzell, Brent. (2012). Linking water quality and well-being for improved assessment and valuation of ecosystem services. *Proceedings of the National Academy of Sciences, 109*(45), 18619-18624.

King, P. (1999). *The fiscal impact of beaches in California*: Public Research Institute, San Francisco State University.

King, Philip G. (2002). *Economic analysis of beach spending and the recreational benefits of beaches in the city of San Clemente*: San Francisco State University.

Kneese, A. V., & Bower, B. T. (1968). *Managing water quality: economics, technology, institutions*: Baltimore: Resources for the Future, Johns Hopkins Press..

Lee, H.-C. (2001). Determinants of recreational boater expenditures on trips. *Tourism Management, 22*(6), 659-667.

Leeworthy, V. R., & Ehler, R. (2010). *Economic Contribution of Recreating Visitors to the Florida Keys, Key West 2007-08*: US Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service, Office of National Marine Sanctuaries.

Leeworthy, V. R., Loomis, D., & Paterson, S. (2010). Visitor Profiles: Florida Keys/Key West 2007-08. *Office of National Marine Sanctuaries, National Ocean Service, National Oceanic and Atmospheric Administration*.

Leeworthy, Vernon R, & Wiley, Peter C. (1993). *Recreational use value for three Southern California beaches*: Strategic Environmental Assessments Division, Office of Ocean Resource Conservation and Assessment, National Oceanic and Atmospheric Administration, US Department of Commerce.

Leeworthy, V. R., Wiley, P. C., English, D. B., & Kriesel, W. (2001). Correcting response bias in tourist spending surveys. *Annals of Tourism Research, 28*(1), 83-97.

Lew, Daniel K, & Larson, Douglas M. (2005). Valuing recreation and amenities at San Diego County beaches. *Coastal Management, 33*(1), 71-86.

Loomis, John B. (1997). Panel estimators to combine revealed and stated preference dichotomous choice data. *Journal of Agricultural and Resource Economics*, 233-245.

Louviere, Jordan J, Hensher, David A, & Swait, Joffre D. (2000). *Stated Choice Methods: Analysis and Applications*: Cambridge University Press.

Louviere, Jordan J, Meyer, Robert J, Bunch, David S, Carson, Richard, Dellaert, Benedict, Hanemann, W Michael, .Hensher, David, & Irwin, Julie. (1999). Combining sources of preference data for modeling complex decision processes. *Marketing Letters, 10*(3), 205-217.

Lupi, Frank, & Feather, Peter M. (1998). Using partial site aggregation to reduce bias in random utility travel cost models. *Water Resources Research, 34*(12), 3595-3603.

Massey, D. M., Newbold, S. C., & Gentner, B. (2006). Valuing water quality changes using a bioeconomic model of a coastal recreational fishery. *Journal of Environmental Economics and Management, 52*(1), 482-500.

McConnell, Kenneth E. (1992). On-site time in the demand for recreation. *American Journal of Agricultural Economics, 74*(4), 918-925.

McFadden, D. (1973). Conditional logit analysis of discrete choice behaviour. *Frontiers in Econometrics. Academic Press. New York.*

McFadden, Daniel. (1974). The measurement of urban travel demand. *Journal of Public Economics, 3*(4), 303-328.

McFadden, Daniel. (1978). *Modelling the choice of residential location*: Institute of Transportation Studies, University of California California.

Messonnier, M., Bergstrom, J., Cornwell, C., Teasley, R., & Cordell, H. (2000). Survey Response-Related Biases in Contingent Valuation: Concepts, Remedies, and Empirical Application to Valuing Aquatic Plant Management. *American Journal of Agricultural Economics, 82*(2), 438-450.

Morey, Edward R, Rowe, Robert D, & Watson, Michael. (1993). A repeated nested-logit model of Atlantic salmon fishing. *American Journal of Agricultural Economics, 75*(3), 578-592.

Morikawa, Takayuki. (1989). *Incorporating Stated Preference Data in Travel Demand Analysis.* Massachusetts Institute of Technology.

Murray, Chris, Sohngen, Brent, & Pendleton, Linwood. (2001). Valuing water quality advisories and beach amenities in the Great Lakes. *Water Resources Research, 37*(10), 2583-2590.

Nelsen, C., Pendleton, L., & Vaughn, R. (2007). A socioeconomic study of surfers at Trestles Beach. *Shore and Beach, 75*(4), 32.

Natural Resources Defense Council. (2009). *Testing the Waters: A Guide to Water Quality at Vacation Beaches*, 19th edition. Retrieved from http://www.nrdc.org/water/oceans/ttw/ttw2009.pdf

Parsons, George R, Massey, D Matthew, & Tomasi, Ted. (1999). Familiar and favorite sites in a random utility model of beach recreation. *Marine Resource Economics*, 299-315.

Parsons, George R, Helm, Eric C, & Bondelid, Tim. (2003). Measuring the economic benefits of water quality improvements to recreational users in six northeastern states: an application of the random utility maximization model. Working paper, University of Delaware, July

Paterson, S. K., Loomis, D. K., & Young, S. E. (2014). The human dimension of changing shorelines along the US North Atlantic Coast. *Coastal Management, 42*(1), 17-35.

Pendleton, Linwood, Mohn, Craig, Vaughn, Ryan K, King, Philip, & Zoulas, James G. (2012). Size matters: The economic value of beach erosion and nourishment in Southern California. *Contemporary Economic Policy, 30*(2), 223-237.

Phaneuf, D. J., & Smith, V. K. (2005). Recreation demand models. *Handbook of environmental economics, 2*, 671-761.

Shonkwiler, John Scott, & Shaw, W Douglass. (1996). Hurdle count-data models in recreation demand analysis. *Journal of Agricultural and Resource Economics*, 210-219.

Small, K. A., & Rosen, H. S. (1981). Applied Welfare Economics with Discrete Choice Models". *Econometrica* 49 (1), 105–130.

Sohngen, B., Lichtkoppler, F., & Bielen, M. (1999). *The value of lake Erie beaches*: Ohio Sea Grant College Program.

Starbuck, C. M., Berrens, R. P., & McKee, M. (2006). Simulating changes in forest recreation demand and associated economic impacts due to fire and fuels management activities. *Forest Policy and Economics, 8*(1), 52-66.

Stynes, D. J. (1997). Economic impacts of tourism: a handbook for tourism professionals. *Urbana, IL: University of Illinois, Tourism Research Laboratory*, 1-32.

Stynes, D. (1998). State and regional economic impacts of Michigan state park visitors. *Report to Public*.

Stynes, D. J., & Sun, Y.-Y. (2003). Economic impacts of national park visitor spending on gateway communities, systemwide estimates for 2001. *Final report to National Park Service.* East Lansing, MI: Department of Park, Recreation and Tourism Resources, Michigan State University.

Swait, Joffre, & Louviere, Jordan. (1993). The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research*, 305-314.

Tarrant, Michael A, Manfredo, Michael J, Bayley, Peter B, & Hess, Richard. (1993). Effects of recall bias and nonresponse bias on self-report estimates of angling participation. *North American Journal of Fisheries Management, 13*(2), 217-222.

Train, Kenneth E. (2009). *Discrete Choice Methods with Simulation*: Cambridge university press.

U. S. Environmental Protection Agency. (1989). *Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models: Part 1* (EPA Contract No. CR-811043-O1-O).

U.S. EPA. (2011). Unpublished data. Beach Sanitary Surveys: Great Lakes Sanitary Surveys. https://www.epa.gov/beach-tech/beach-sanitary-surveys

U.S. EPA. (2008). "Great Lakes Beach Sanitary Survey User Manual." Office of Water. EPA-823-B-06-001. May. Retrieved from http://water.epa.gov/type/oceb/beaches/upload/2008_05_29_beaches_sanitarysurvey_user-manual.pdf

U.S. EPA. (2009). *Fiscal Year 2010 EPA Budget in Brief*. United States Environmental Protection Agency Office of the Chief Financial Officer (2710A) Publication Number: EPA-205-S-09-001. May.

Verhougstraete, M., Rose, J., Byappanahalli, M., & Whitman, R. L. (2010). Cladophora in the Great Lakes: impacts on beach water quality and human health. *Water Science & Technology, 62*(1).

Von Haefen, Roger H, & Phaneuf, Daniel J. (2008). Identifying demand parameters in the presence of unobservables: a combined revealed and stated preference approach. *Journal of Environmental Economics and Management, 56*(1), 19-32.

Watson, A. E., Cole, D. N., Turner, D. L., & Reynolds, P. S. (2000). Wilderness recreation use estimation: a handbook of methods and systems. *Gen. Tech. Rep. RMRS-GTR-56. Ogden, UT: US Department of Agriculture, Forest Service, Rocky Mountain Research* Station.

Weicksel, Scott Arndt. (2012). *Measuring Preferences for Changes in Water Quality at Great Lakes Beaches Using a Choice Experiment.* Master Thesis, Michigan State University.

Whitehead, John. C. (1991). Environmental interest group behavior and self-selection bias in contingent valuation mail surveys. *Growth and Change, 22*(1), 10-20.

Whitehead, John C. (2005). Combining willingness to pay and behavior data with limited information. *Resource and Energy Economics, 27*(2), 143-155.

Whitehead, John C, Dumas, Christopher F, Herstine, Jim, Hill, Jeffery, & Buerger, Bob. (2008). Valuing beach access and width with revealed and stated preference data. *Marine Resource Economics*, 119-135.

Whitehead, John C, Edwards, Bob, Van Willigen, Marieke, Maiolo, John R, Wilson, Kenneth, & Smith, Kevin T. (2000). Heading for higher ground: factors affecting real and hypothetical hurricane evacuation behavior. *Global Environmental Change Part B: Environmental Hazards, 2*(4), 133-142.

Whitehead, John C, Haab, Timothy C, & Huang, Ju-Chin. (2000). Measuring recreation benefits of quality improvements with revealed and stated behavior data. *Resource and energy economics*, *22*(4), 339-354.

Whitehead, John C, Pattanayak, Subhrendu K, Van Houtven, George L, & Gelso, Brett R. (2008). Combining revealed and stated preference data to estimate the nonmarket value of ecological services: an assessment of the state of the science. *Journal of Economic Surveys*, *22*(5), 872-908.

Williams, Huw CWL. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and planning A*, *9*(3), 285-344.

Wilton, J. J. & Nickerson, N. P. (2006). Collecting and using visitor spending data. *Journal of Travel Research*, *45*(1), 17-25.

Wooldridge, Jeffrey M. (2010). *Econometric analysis of cross section and panel data*: MIT press.

Yeh, C. Y., Haab, T.C., & Sohngen, B.L. (2006). Modeling multiple-objective recreation trips with choices over trip duration and alternative sites. *Environmental and Resource Economics*, *34*(2), 189-209.