Combining Revealed and Stated Preference Methods for Valuing Water Quality Changes to Great Lakes Beaches[1]

Li Cheng

Michigan State University

chengli5@msu.edu

Frank Lupi

Michigan State University

lupi@msu.edu

Selected Paper prepared for presentation for the 2016 Agricultural & Applied Economics Association, Boston, MA, July 31-August 2

[1] Financial support for this research was provided by NOAA Coastal Oceans Program, Michigan Sea Grant and Michigan State University's AgBioResearch and Michigan Applied Public Policy Research Grants.

1. Introduction

The water quality of the Great Lakes is of interest to policy makers and the public. Legislative efforts and government regulation, such as Clean Water Act (CWA, 1970, 1972) and Great Lakes Water Quality Agreement (GLWQA, 1972, 1978, 1987, 2012), aim to restore and enhance Great Lakes water quality. 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 when 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, three key challenges for valuing water quality changes include measurement, comprehension and data issues. First, unlike air quality, which has a comparatively small number of accepted measures of quality, water quality can be measured 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.

Noting that some of the strengths of RP approaches can be 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. We use web survey data that 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 Cheng (2016), 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 paper 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 Cheng (2016), 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 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. Our nesting process can be visualized as groupings of the *M* nests, *M*= {*Trip, No trip*}, the *L* lakes in the nest *Trip*, *L* = {*Lake Erie, Lake St. Clair, Lake Huron, Lake Michigan}*, and the *J* beaches at one of the lakes *l.* The nesting tree is illustrated in the figure below:

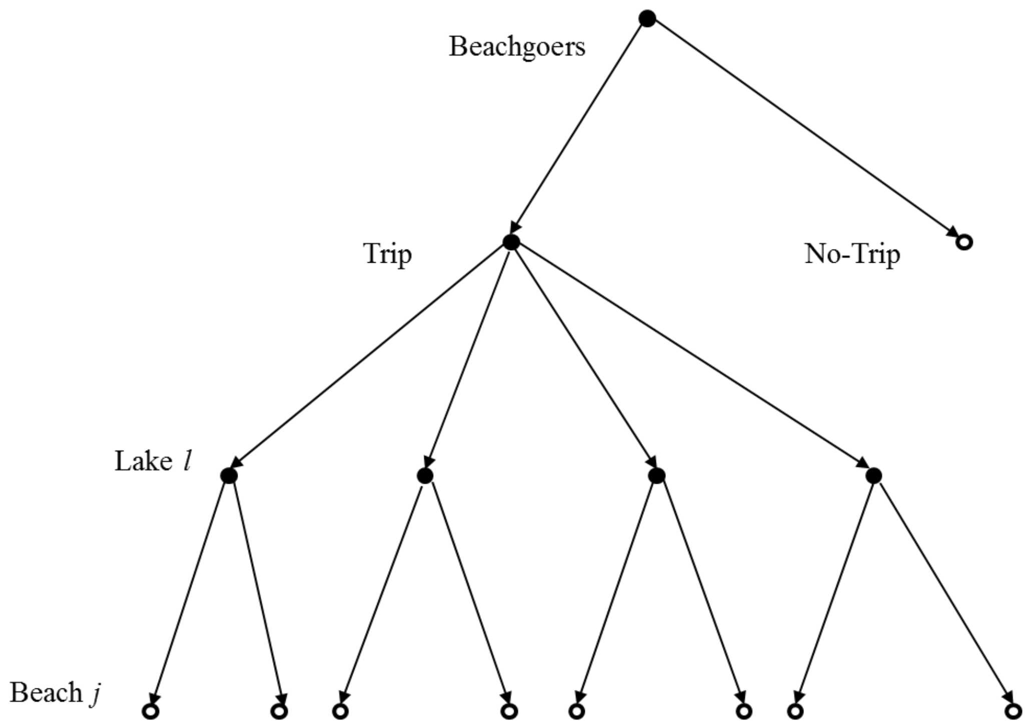

Figure 1. Repeated three level decision tree of beach recreation trip

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

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

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:

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

where w_n is the weight of person *n*. There are three purposes of the weight (See Cheng 2016, Appendix A). 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 Cheng (2016), 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 Cheng (2016) to handle trips with partial information. Denoting the grouped area as a, the log-likelihood function for this sample of "grouped beaches" is:

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

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 & 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 random utility framework (e.g., Adamowicz et al., 1994; 1997; Earnhart, 2001; Von Haefen & Phaneuf, 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 & 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 & Louviere, 1993; Bradley & Daly, 1997; Hensher, Louviere, & Swait, 1998; Louviere, Hensher, & Swait, 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[2], 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 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.

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

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%. Cheng (2016) 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.

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.

Table 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 Cheng (2016), 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 reports descriptive statistics for both individual characteristics and site attributes in the RP data.

Table 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
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 levels of water quality (see Table 3), and $\boldsymbol{\delta}$ is a vector 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 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 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:[3]

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

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 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

[3] 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).

Table 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 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 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 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[4]. 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

[4] 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 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 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 yielding 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 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 & 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:

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

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 8.

Table 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:

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

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

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

where

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

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:

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

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:

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

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

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

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

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

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

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 Cheng (2016), 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 8 and Table 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 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 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 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).[5] Trips increase slightly

[5] 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.

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 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 Cheng (2016) 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 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 11: Estimated Trips and Welfare Measures of Shifting Half of Sites' Water Quality up 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 *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 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 paper 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.

References:

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.

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.

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.

Ben-Akiva, Moshe, Bradley, Mark, Morikawa, Takayuki, Benjamin, Julian, Novak, Thomas, Oppewal, Harmen, & Rao, Vithala. (1994). Combining revealed and stated preferences data. *Marketing Letters, 5*(4), 335-349.

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.

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.

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

Cheng, Li. (2016). *Measuring the Value and Economic Impacts of Changes in Water Quality at Great Lakes Beaches in Michigan.* PhD Dissertation, Michigan State University.

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.

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.

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.

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.

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.

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.

McFadden, Daniel. (1978). *Modelling the Choice of Residential Location*: Institute of Transportation Studies, University of California California.

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.

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

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.

Train, Kenneth E. (2009). *Discrete Choice Methods with Simulation*: Cambridge university press.

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. (2011). Unpublished data. Beach Sanitary Surveys: Great Lakes Sanitary Surveys. https://www.epa.gov/beach-tech/beach-sanitary-surveys

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.

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, 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.

Wooldridge, Jeffrey M. (2010). *Econometric analysis of cross section and panel data*: MIT press