How does water quality affect people's recreation behaviour and welfare in Alberta?

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Master of Science
In the Department of Agricultural and Resource Economics
University of Saskatchewan
Saskatoon

By

Nasim AminiKhezri

PERMISSION TO USE

DISCLAIMER

ABSTRACT

This thesis studies the effects of lakes' water quality on recreation demand across provincial parks of Alberta using a travel cost model. Canada has an extremely large number of lakes with fairly good water conditions. However, poor water quality is found in many Albertan lakes and population growth, agricultural pollution, and climate change may worsen water quality in the future. I employed revealed preference of over 70,000 individuals per year in conjunction with the beach advisories of 73 campsites for the years between 2014 and 2018.

This study measures beach advisories' economic impacts on campers' behaviour at the participation decision and site choice stages through a sequentially estimated two-stage Nested Logit (NL) model. I estimate a recreation demand model of combining single and multiple-day trips, calculating travel costs, using Alternative Specific Constants (ASCs) and Time Specific Constants (TSCs) to control for time-variant and unobserved sites characteristics.

The results show that participation and site decisions are formed independently, and campers would prefer to substitute the recreation destination with any other campsites rather than alter with different types of activities or staying at home. Campers were found to be negatively affected by the presence of beach advisory at the site. Our results demonstrate that campers are willing to pay \$15[1] more per trip for removing the beach advisory. I use the model to evaluate the welfare impacts of removing all beach advisories. The result can be useful to evaluate the non-market benefits of improving lake water quality.

[1] In the present study, the dollar sign (\$) is defined as a Canadian Dollar.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor Dr. Patrick Lloyd-Smith for his dedicated support and guidance. Further, I would like to thank Alberta Health Services and Alberta Environment and Parks for providing inputs and made this MSc programme happen. Many thanks also to all participants that took part in the study and enabled this research to be possible.

DEDICATION

Dedicated to my beloved family, who have been the source of inspiration and gave me strength when I thought I was giving up, who continuously provide their moral, spiritual, and financial support. Without whose support, I would not be where I am today.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACFT .... Alberta Centre for Toxicology
AEP .... Alberta Environment and Parks
ASC .... Alternative Specific Constant
CL .... Conditional Logit
CV .... Compensation Variation
EPA .... Environmental Protection Agency
GEV .... Generalized Extreme Value
HAB .... Harmful Algal Bloom
IIA .... Independence of Irrelevant Alternative
IIN .... Independence from Irrelevant Nests
LCM .... Latent Class Model
LONGWKD .... Long weekend
MLE .... Maximum Likelihood Estimation
MNL .... Multinomial Logit
MWTP .... Marginal Willingness to Pay
NL .... Nested Logit
RAP .... Reserve Alberta Parks
RNL .... Repeated Nested Logit
RP .... Revealed Preference
RUM .... Random Utility Model
RXL .... Repeated Mixed Logit
SP .... Stated Preference
TC .... Travel Cost
TSC .... Time Specific Constant
TCM.. Travel Cost Model
WQ .... Water Quality
WQI .... Water Quality Index
WTA .... Willingness to Accept
WTP .... Willingness to Pay

Chapter 1
Introduction

1.1 Introduction

Canada has the third most abundant annual freshwater resource, which makes up 9% of the world's freshwater while its population is not more than 1% of the world's total (Natural Resources Canada 2014). A wide variety of freshwater sources exists across Canada, including rivers, lakes, glaciers, tiny and large streams, which have provided numerous outdoor activities and pleasant sceneries across due to its beautiful nature. Statistics show that with 561 lakes larger than 100 km^2 (The Atlas of Canada - Lakes 1973), Canada has the most lake areas in the world (Dewar et al. 2010) and naturally serves as a recreational attraction.

Generally, most water bodies in Canada have good water quality (The Conference Board of Canada- Water quality index 2018) but, some areas suffer the impacts of poor water quality, such as blue-green algal blooms. A growing population can threaten water quality by consuming more water and producing more municipal waste. Due to industrial sewage, municipal waste, and fertilizer residuals used in the agriculture sector, the lakes' water-quality is at risk (The Conference Board of Canada- Water quality index 2018). Hence, due to human activities, the areas with poor water quality tend to be closer to human settlements and associated recreation activities.

Algal blooms, a biological indicator, can influence a traveller's perception visually by forming colonies and affect water clarity and colour. Besides, when the algal blooms die, microorganisms feeding and decomposing them. Therefore, available oxygen in the water consumed and aquatic animals encounter with oxygen scarcity. The effect of algal blooms has received scant attention due to the difficulty of data collection. Moreover, cyanobacteria, which has some characteristics of bacteria and some characteristics of algal blooms (World Health Organization 2003), produces a toxin called Microcystin- LR or blue-green algal.

In the past decades, the earth's climate has been changed. Climate change has extended the summer length and consequently has provided more opportunities for some activities we can do in summer (Mendelsohn and Neumann 2004). Meanwhile, the amount of blue-green algae increases as the temperature increases (Mooij et al. 2005) and can impair water quality. Blue-green algae can change the lakes' colour and odor, and produce some toxins that pose some health risks to the public from skin and eye irritation to liver and nerve toxins ("Cyanobacterial ('Blue-Green Algae') Blooms and Toxicity" 2015). Furthermore, the quality of water resources might influence visitors' well-being, particularly for the activities that have direct interaction with water like swimming.

There are 600 lakes and over 500 sites in Alberta. These lakes offer a diversity of recreational opportunities, including both activities that have direct contact with water (e.g., swimming, water skiing), and indirect contact (e.g., fishing and boating). Recently, water quality issues have become more significant due to public awareness. In 2012 and 2013, the Alberta Cyanobacteria Beach Monitoring program was established to inform the visitors about the water quality condition for specific purposes (Alberta and Alberta Health 2014). The program started in Alberta because of the prevalent of cyanobacteria across lakes along with its potential health hazards.

Thus, parks' abundance, lakes' variety, and water quality issues have made Alberta an interesting area for studying recreation demand in Canada. Likewise, climate change and municipal growth encouraged me to investigate the effect of water quality on recreational demand to provide better information. The present study aims at investigating the effect of lakes' water quality on the visitors' behaviour and the level of their welfare.

However, assessing the economic value of water quality as a public good is not easy because its value is not reflected in the market. Therefore, assessing an individual's willingness to pay for receiving a higher quality or quantity, and willingness to accept for giving up a given quality or quantity of environmental goods are often hard since they are not revealed in market prices. A close interaction between water quality and human life, as well as population growth and climate change, has made water quality vital for public health and ecosystem habitats. This study aimed to investigate the effect of lakes' water-quality on recreation behaviour.

This paper studies the relationship between lake water-quality measures and recreation demand between 2014 and 2018 across provincial parks of Alberta. Numerous surface water resources across Alberta have provided lots of recreational attraction for people as a public good. On average, 76,000 two-person reservations were conducted through Reserve Alberta Parks' website for visiting the Provincial Parks of Alberta annually. Hence, lake water-quality and its effects on recreation demand are essential since it involves around 3.5% of Alberta's population.

An economic evaluation is conducted to answer the first underlying question of whether a lake's water quality is important for people or not, and whether campers really suffer from the lake's water quality degrading? Then, this thesis intends to determine the extent to which lake water quality affects individuals' welfare and behaviour. The purpose of this study is to evaluate the cost of bad water-quality or the benefit of improving water quality across the lakes.

The revealed preferences of a substantial number of campers and beach advisories, as the lakes water quality criteria, are set out to discover the relationship between lakes' water quality and recreation demand. A study showed that 70% of the Great Lakes' campers used the information provided the beach advisories in 1998 (Murray, Sohngen, and Pendleton 2001). The estimated percentage might have been increased during the last decades since public awareness and information flow have been improved.

Generally, lakes are open for recreation services between May and September. Recreationists choose a campground during the open water season (late May to early September) among a set of alternatives, including a stay at home option. In the meanwhile, beach advisory signs are posted to inform the visitors about the water quality conditions for specific purposes. Therefore, I use a Travel Cost Model (TCM) to conduct this behavioural study, where decisions are assessed based on actual behaviour, travel cost, site attributes, and time-variant variables.

I monetize the welfare change to answer the second question and provide an understandable and comparable intuition about the visitors' well-being change in response to water quality improvement. I employ a compensation variation notion to report how much people are willing to pay to improve the lakes' water quality and assess yearly welfare changes. Measuring the total welfare change is possible under several scenarios. Here, I conduct a scenario in which the lakes' water quality improves to the point that there are no beach advisories at the sites. This paper addresses the economic benefits of improving lake water quality based on the mentioned scenario.

Generally, most of the recreation demand studies conducted in the United States (Wolf, Georgic, and Klaiber 2017; Zhang and Sohngen 2018; Egan et al. 2009) and demonstrated a negative relationship between water quality degrading and individuals' welfare or behaviour. To the best of our knowledge, this will be the first study in Canada, which uses a travel cost model to assess water quality changes on recreational demand based on revealed preferences. This is the first contribution of this study and would be helpful to provide a viewpoint on water quality and recreation demand. Johnston and Thomassin (2010) found differences in Canadian and the U.S willing to pay for water quality improvement. Based on their multinational meta-analysis, Canadians have a lower value for water quality improvements compared to the U.S (Johnston and Thomassin 2010).

The rich dataset of revealed preferences of the considerable number of travellers and its combination with the beach advisories records employed in this study. At the same time, most pre-existing studies in the field of recreation demand have focused on stated preferences or the small size of the human decisions datasets, which meant a drawback. Most recreation demand studies (Wolf, Georgic, and Klaiber 2017; Egan et al. 2009; Ji and Keiser 2016) had only focused on the total number of trips taken during a month, season or year. However, this paper deals with the exact day of the trips. Therefore, I can link individuals' behaviour with the lakes' water condition more precisely.

The employed water quality measure plays a significant role in assessing individuals' behaviour and welfare. Literature has focused on different water quality measurements and employed such measures in different ways because accessing a competent water quality dataset is not possible in most cases. Numerous studies (Zhang and Sohngen 2018; Wolf et al. 2019; English et al. 2018; Van Houtven et al. 2014) have shown that travellers' welfare is influenced by water quality. It has also been demonstrated that the number of visits increases as the water clarity increases (Keeler et al. 2015). Similarly, algae prevalence decreases the number of anglers at least by 10% (Wolf, Georgic, and Klaiber 2017).

The overall structure of the study takes the form of six chapters, beginning with a brief background of water in Canada. I will then go on to present an overview of the related works in chapter two, and in chapter three, I will describe the data employed. The fourth chapter is concerned with the methodology used for this study. The fifth division presents the research findings, focusing on the estimated relationship and welfare measure under a given scenario. Section six includes a summary of this study, with conclusions and limitations, applications to management, and future research.

Chapter 2
Literature Review

2.1 Introduction

A basic introduction to water resources in Canada and the summary of cyanobacteria side effects, providing an elementary understanding of the water quality indicator, was presented in chapter 1. Various water quality indices, different types of recreational activities, and data availability have brought about numerous studies (Wolf, Georgic, and Klaiber 2017; Wolf et al. 2019; Keeler et al. 2015; Zhang and Sohngen 2018; Egan et al. 2009), in which a considerable amount of them were conducted in the United States.

This section reviews the previous water quality and recreation demand literature to identify the gaps in this research area, conflicts in previous studies, and open questions from other research. First, I will review related studies in the aspect of the type of data they used, including recreational data and water quality data.

2.2 Revealed Preference and Stated Preference

Non-market valuation is based on individuals' choices. People make their decisions' by evaluating the possible options and trade-offs they are willing to make between a set of alternatives since the access cost and individuals' preferences are dissimilar. The term 'value' refers to these trade-offs in economic theory. Ultimately, each choice illustrates a combination of different views and reveals the individual's preference. In environmental valuation, individuals' preferences are used to form a behavioural model and welfare calculations (Segerson 2014).

Most non-market valuation studies of water quality using the stated preference (SP) method (Zhang and Sohngen 2018; Van Houtven et al. 2014). In the SP method, a hypothetical situation is described to survey respondents. The SP model's reliability depends on whether the respondents understand and commit to the described condition and respond to the tasks honestly (Abdullah et al. 2011).

The accuracy of the SP method is highly dependent on the survey method and the questionnaires. Furthermore, strategic behaviour or hypothetical bias is one of the most common challenges associated with the SP method. In this case, respondents might overestimate or underestimate their willingness to pay (WTP) or willingness to accept (WTA) under a given scenario if they feel they might incur a cost or lose a facility by being truthful. Thus, this method's validity could be improved by increasing incentives for truth-telling, avoiding having a cheap talk, and asking how much respondents are certain about their stated preferences.

The revealed preference (RP) method studies an individual’s actual behaviours, in which only existing alternatives are portrayed. Hence, using the RP method has the potential to provide more reliable observations under the real-world situation in comparison with the SP method. Therefore, RP methods do not suffer from the hypothetical nature of survey questions. Unlike conducting a survey, sometimes there are some limitations to access further required information, such as campers’ preferred recreational activity, income, or any other sociodemographic characteristics.

An on-site sample is a technique of collecting data from people who are present at zonal. The recreation site demand surveys use on-site samples fail to engage the travellers who did not visit the campsite and only collect the campers' observations which were at the site. Moreover, the campers who travel more often have a higher chance of participating in the survey than the travellers who occasionally visit a site (D. Shaw 1988). Van Houtven et al. (2014) conducted an on-site survey to assess an individual’s willingness to pay for water quality improvement, while selection bias could be a potential concern. However, in the RP method, the selection bias issue is not a big deal.

In some cases, the dataset has a small size in the SP methods. Zhang and Sohngen (2018) conducted a mail survey to 3000 fishing licence holders whose final sample had only 767 anglers. A major advantage of the non-survey data used in this study is associated with a noticeable number of observations. However, by conducting a survey, I would have the opportunity of collecting all the required information, and I have control over them.

Some studies used the RP method to assess the relationship between water quality and recreation demand (Egan et al. 2009; Wolf, Georgic, and Klaiber 2017; Keeler et al. 2015). The present research explores the effects of the lake's water quality on recreation demand by employing a rich RP dataset of over 70,000 people per year.

2.2.1 Revealed Preference Studies

A recreational water quality valuation was done in Iowa by studying the actual behaviour of fewer than 4,000 people in 2002 (Egan et al. 2009). The weather condition in Iowa does not restrict lake usage during the year, unlike our study area, where frozen lakes lead to limit most water activities from October to April since doing most water activities is not possible during winter. Thus, one trip per week was considered for each household for the whole year. The authors had information about the total number of hours of each household to approximately 130 lakes in Iowa, along with sociodemographic information. The RP data set was also used by other researchers for the years 2002 to 2005 and 2009 (Ji and Keiser 2016). In the same way, the choice occasion was 52 and stay at home option was included in the model to provide a complete choice set.

Likewise, the Harmful Algal Blooms (HAB) impact and angler's actual behaviour in Ohio were studied for around three years (Wolf, Georgic, and Klaiber 2017). The authors had information about households' zip codes and the monthly fishing permit sale, not the day in which the fishing trip was taken. The RP of more than 700 households in Ohio counties was also used to assess the impact of the HAB and E.coli on recreational behaviour during the summer of 2016 (Wolf et al. 2019). The respondents were asked about their typical day-trip as well as the total number of trips to Lake Erie for more than three months.

Most of the RP studies did not have the information about the exact trip dates to each campsite, unlike research that was done in the United States in 2015. Keeler et al. (2015) employed Flicker users' sociodemographic data as well as geotagged photographs that were taken during 24 hours (Keeler et al. 2015). For the water quality section, water clarity, lake size, and other site attributes were considered. Due to data limitations, the authors calculated the average annual trips to each lake. Hence, the researchers could not use the advantage of employing the exact day of the tours in the data set practically.

Although extensive research has been carried out on water quality and recreation demand, there is still a gap in evidence for this relationship in consideration of the exact day of the trips. This study aims to contribute to this growing area of research by incorporating the exact day of the trips rather than the total annual, seasonal, or monthly number of tours. Employing the travel dates helps figure out the relationship between biological indicators and an individual's behaviour since these water quality measures are responsive to various environmental situations and might change day by day.

2.3 Water Quality

Generally, in water quality and recreation demand, data sets are composed of two independent components; individual's behaviour, as discussed, and water quality indices. There are several water quality indicators, including biological indicators, e.g. algal blooms and cyanobacteria, conventional indicators, e.g. pH and suspended solids, temperature, and nutrients. In the following, I will review how these components can affect travellers' welfare or behaviour and how researchers employed different water quality indicators in their studies.

2.3.1 Water Interaction

Some studies in surface water quality areas addressed recreational activities and welfare change in response to water quality alterations (Wolf, Georgic, and Klaiber 2017; Zhang and Sohngen 2018; Feather, Hellerstein, and Tomasi 1995; Alvarez et al. 2014). There is a growing body of literature that recognizes the importance of water quality for travellers in the United States, in which anglers, beachgoers, and campers' welfare were estimated(Keeler et al. 2015; Egan et al. 2009; Wolf et al. 2019; Ji and Keiser 2016).

Early works in water quality and individual's welfare were mainly based on changes in catch rate, as an indicator of the water quality, and anglers' reactions. In 1982, Russell and Vaughan assessed the benefits of controlling water quality in the United States by observing changes in anglers' participation in response to improving water conditions under four different scenarios. Findings illustrate an overall positive relationship between improving fishable water by implying more strict laws for both warm and cold water. Likewise, the result was supported by another study in Minnesota, which revealed a positive relationship between anglers' participation and water quality improvement based on the RP method in 1989 (Feather, Hellerstein, and Tomasi 1995).

Recently, Zhang and Sohngen (2018) defined this positive relationship as a monetary value. Based on the study conducted in 2013, anglers' were willing to pay 8 USD to 10 USD more per trip to experience one mile fewer boating through HAB in Ohio (Zhang and Sohngen 2018). Other assessments have shown fishing permit sales will be dropped by 10% to 13% if algal blooms pass the standard threshold of 20,000 cells/mL (Wolf, Georgic, and Klaiber 2017).

Studies of swimming, a recreational activity that has a direct interaction with water, have also given some evidence for the positive relationship between water quality and recreational demand. Based on Bockstael, Hanemann, and Kling's study across 30 beach sites in Boston, perception of water quality plays a significant role in recreational behaviour (Bockstael, Hanemann, and Kling 1987). Along with this declaration, there is some evidence that suggests using perception in non-market evaluations, rather than objective attributes, will provide a model with better performance (Adamowicz et al. 1997).[2] However, there might be some discriminations on the intensiveness of perception effect in the different environmental quality areas, and researchers have not treated perception in much detail in the water quality and recreation demand area.

A recent study was done in Lake Erie based on travellers' actual behaviour (Wolf et al. 2019). The researchers classified recreation activities into two different groups; activates with direct and indirect interactions with water and figured out some information about the kind of activities that the travellers might have. Then, a latent class model was applied to evaluate the impacts of HAB and E.coli on the individual's behaviour. Findings demonstrated that E.coli and HAB affect traveller's behaviour in different ways; beachgoers are more sensitive to E.coli, while anglers are more responsive to HAB. Therefore, findings show that although different users will respond to various water quality measures, water quality degrading affects all travellers negatively.

2.3.2 Water Quality Measure

Finding water quality data is an additional difficulty for researchers. Using Iowa lakes as a case study is a unique opportunity for researchers since several physical water quality measures are available. Moreover, due to a wide variety of environmental conditions, researchers can

[2] A study of chronic wasting disease (CWD) and recreational hunters' behaviour in Alberta did not support this idea (Zimmer, Adamowicz, and Boxall 2012). The comparison between the hunter's perception and objective measures of CWD did not represent a significant difference.

characterize respondents' characteristics and evaluate the associated impacts on recreational demand (Egan et al. 2009). Hence, most of the studies in this area are done in Iowa Lakes (Egan et al. 2009; Jeon and Herriges 2010; Ji and Keiser 2016), and some of them are done in Lake Erie (Wolf et al. 2019; Zhang and Sohngen 2018).

Assessing the effect of water quality on recreational demand and individuals' behaviour might vary across different water quality indices. In other words, figuring out the most effective water index which drives an individual's behaviour is still challenging. In 1982, Russell and Vaughan considered fishery type catch rate as a representative of low dissolved oxygen, inappropriate pH, and solid soils, which seems reasonable due to the data limitation (Russell and Vaughan 1982). Probably catch rate looked attractive for researchers since it directly affects fisher's perception of water quality. Besides, the catch rate data, as a water quality indicator, was available for the researchers. However, the catch rate might be influenced due to several factors other than changes in water quality, including; changes in the food chain, interactions with other species, weather patterns, and spawning.

Using the catch rate as a representative of water quality changes kept this question ambiguous whether the outcome was due to changes in other water quality characteristics. In the following years, angler's choice was assessed among two regions in Minnesota (Feather, Hellerstein, and Tomasi 1995). Unlike other studies, average Secchi depth, littoral zone, and lake's specifications were used as water quality criteria. Also, the natural logarithm of these variables was employed to avoid correlation with lake size and capture its non-linear effect.

Gradually, the catch rate was accompanied by more accurate water quality measurement, algal blooms. In the study of recreational angler's welfare in Ohio, in addition to the expected catch rate, water clarity, boating time to the fishing zone, and the distance of boating through algal blooms area were considered as water quality indicators (Zhang and Sohngen 2018). This online survey studied sampled anglers' preferences in response to six scenarios that vary across the mentioned variables as well as the driving distance between angler's location and boat ramp.

Technology helped the study that was done in Ohio and Lake Erie based on fishermen's actual behaviour (Wolf, Georgic, and Klaiber 2017). The authors monitored the water condition through remote sensing data and calculated a ten-day composite of algal blooms. Then, a dummy variable was employed to show whether cyanobacteria was greater than 20,000 cells/ml or not. Three closest locations to the angler's zip code were selected to record monthly precipitation, temperature, and algal bloom readings. Dummy variables were also employed by researchers to display campsites attributes (Keeler et al. 2015; Wolf, Georgic, and Klaiber 2017).

To determine the effect of the water quality on recreation demand, Parsons, Helm, and Bondelid (2003) categorized water quality into three stages (i.e., low, medium, and high). This classification was based on biological oxygen demand, total suspended solids, dissolved oxygen, and fecal coliforms (Parsons, Helm, and Bondelid 2003). In some cases, the authors used the averages of watershed water quality for coastal water quality. The findings of this study suggested that a larger recreational benefit increase is associated with significant water quality improvement.

Likewise, various water quality measurements were used by Egan et al. (2009), including; Secchi depth, Chlorophyll, nutrients level (nitrogen and phosphorus), suspended solids, and cyanobacteria. These water quality measurements were collected three times per year. The findings showed that recreationists' behaviour is mainly influenced by Secchi transparency, cyanobacteria and nutrients level (Egan et al. 2009).

An initiative use of Water Quality Index (WQI) was done by aggregating six different standardized water quality sub-indexes, i.e., turbidity, dissolved oxygen, total nitrate, phosphorus, solids, and pH (Ji and Keiser 2016). Although I expected a correlation between turbidity, included in WQI, and Secchi depth, results were not influenced, and the authors compared the effectiveness of these two water quality indices on recreational use. WQI can make the analysis simpler since a single number represents the overall quality of water. Also, the challenge associated with selecting effective water quality indices might be alleviated when researchers use a number as a representative of numerous criteria.

Existing research recognized that although various water quality measures will affect recreationists' welfare differently, water quality degrading affects all travellers negatively. Probably because of missing values of water quality observations, as well as not having information for the trip days to adjust with water quality records, the majority of water quality studies conducted using monthly, seasonally, or yearly average and index for most water quality readings. A more comprehensive study would include trip days and, consequently, more accurate water quality estimations. The second and the third column of table A.1 presents a summary of the used variables in the pre-existing studies.

2.4 Recreation Demand Modelling

Publications that concentrate on the surface water quality and recreation demand more frequently adopt a Random Utility approach. Random Utility Model (RUM) is one of the most popular tools for modelling the non-market valuation of discrete choice methods. Therefore, assessing the probability of choosing one alternative rather than the other alternatives become feasible with the aid of this model.

The RUM has two components; the first is the systematic factor (V_{ij}) and the second is the stochastic or the random portion (ε_{ij}). The systematic factor is quantifiable for researchers, while the random part is only observable for the decision-makers. Equation 2.1 represents a fundamental of the RUM.

$$U_{ij} = V_{ij} + \varepsilon_{ij} \tag{2.1}$$

Various types of discrete choice models are presently in practice for assessing the probability of choosing an alternative by a person. Scientists use the results to forecast individuals' behaviour in a hypothetical situation. Initially, discrete choice models are classified based on the number of available alternatives. The binomial choice model suits the choice sets with two alternatives, and the multinomial choice model is appropriate for three and more options. Some recognized models in recreation studies are Conditional Logit (CL), Latent Class (LCM), and Nested Logit (NL), Repeated Mixed Logit (RXL) models. Each of these models is appropriate and applicable in different cases.

Different authors have measured recreation demand for water quality improvement in a variety of ways, and each of the employed methods has some pros and cons. The CL model is popular when the alternatives have mutual attributes, and the random portion has i.i.d distribution. The LCM is used to address systematic heterogeneity in the RUM. Based on this model, individuals are categorized into different groups based on psychological or sociodemographic characteristics commonality, and a membership function is required for each group (Boxall and Adamowicz 2002). In recent years the LCM employed in different studies to assess biological indicators importance for the travellers across lake Erie (Zhang and Sohngen 2018; Wolf et al. 2019).

Feather, Hellerstein, and Tomasi (1995) studied anglers' recreational demand in Minnesota. The authors divided surface water into two regions; the southern has warm water, and the northern has cold water. The authors innovated a novel model, the NL model, to estimate fishermen's choices. The findings were varied across the two regions, with a more significant increase in welfare and participation for the southern part in comparison to the northern one (Feather, Hellerstein, and Tomasi 1995). Morey, Rowe and Watson (1993) also used this method to assess the participation and site decisions of anglers. Recently, this model was also employed to evaluate the economic effect of losing the Gulf Coast regions' shoreline by considering single and multiple day trips (English et al. 2018).

The conditional indirect utility function for the various alternative is the same in RXL and Repeated Nested Logit (RNL). The differences arise from the distribution of the vector of the unknown parameters and the random portion. RXL assumes that the random component has i.i.d distributed over the individuals, and the unknown parameters are randomly changing among the choice occasions and individuals (Herriges and Phaneuf 2002). In other words, the RXL model is the extended version of the CL model by letting the unknown factor to be randomly distributed. Although the RXL model supports a vast substitution among alternatives and individuals' heterogeneity (Herriges and Phaneuf 2002), computational issues count as a limitation for employing this method with substantial observations.

Egan et al. (2009) employed the RXL method to assess water quality value for recreationists in Iowa lakes. However, the RXL model failed to forecast non-participants behaviour for conditional welfare estimation (Egan et al. 2009). From the unknown attributes component, Ji and Keiser (2016) conducted the RXL model with and without ASC, which captures the effect of impressive variables that are not included in the model (ASCs) to assess the relationship between eater quality indices and water-based recreational demand. The results revealed that addressing the ASCs and the random portion issues would provide a better understanding of the relationship between lakes' water quality and recreational behaviour (Ji and Keiser 2016).

2.5 Travel Cost Calculation

Travel cost computation can affect estimations. Travel cost plays a role in determining which campground traveller might visit since it reflects the resources an individual has to give up to visit a site (Parsons 2014). Different methods have been proposed to estimate the travel cost per km. Regarding the driving expenses, some authors took no account of depreciation on a car (Zhang and Sohngen 2018) in travel cost calculation since they believe that it could be a neglectable factor over time. However, in some studies, depreciation was included in estimations (Egan et al., 2009; English et al., 2018; Wolf et al., 2019). Hang et al. attempt to evaluate the impact of depreciation on an individual's decision. This study that conducted on the United States' Households revealed preferences illustrates depreciation has a slight effect on their choices (Hang et al. 2016). Therefore, depreciation is considered a part of travel costs in this study. Moreover, unlike the previous studies that calculated the travel cost for a specific type of car (English et al. 2019), the estimated total annual driving cost is operated for three different car categories.

Accessing recreation sites is possible through different types of transportation, i.e. train, car, or flight. The study of Flickr geotagged photographs considered twelve closest locations to Minnesota, and it was assumed that travel beyond the nearest locations used other transportation methods rather than driving and withdrew them from the estimations (Keeler et al. 2015). In a study of the oil spill in the Gulf of Mexico and recreation lost, a weighted average of both air travel and road travel was considered in travel cost estimation (English et al. 2018). Since people are more likely to take a road trip across a province for visiting campsites, a road trip is considered a preferred type of transportation.

The second challenge in travel cost calculation arises from the inconsistency associated with the value of time. In other words, different ratios of wage rate or households' annual income were considered to estimate the value of the time. In some studies, one-third of wage rates (Ji and Keiser 2016; Wolf et al. 2019; Zhang and Sohngen 2018), and in some cases one-quarter of hourly wage (Zimmer, Adamowicz, and Boxall 2012) employed. However, seasonal camping demand evaluation of 71,000 recreational trips revealed that the value of travelling time is much greater than the conventional ratio and close to the full wage (Lloyd-Smith and Becker 2020). Lloyd-Smith and Becker used the median annual household after-tax income to estimate the travel cost in two bounds. The higher bound (i.e., two-third of income) fits the revealed preferences better in comparison with the lower bound (i.e., one-third of income)(Lloyd-Smith and Becker 2020). The advantage of the employed approach for assessing the value of time in this study is associated with selecting this accurate wage rate. I will define the procedures and methods used in this investigation in chapter 4.

2.6 Welfare Measure

Collectively, findings of published papers in the United States revealed that water quality degrading influences people's welfare or behaviour negatively. However, the magnitude of this effect might vary across different studies since they employed different data, methods, and scenarios. Impaired water quality has contributed to a decline in fishing permit sales per month if algal blooms pass the threshold of 20,000 cells/mL (Wolf, Georgic, and Klaiber 2017). This finding was confirmed by the evidence from the behavioural study of anglers in Ohio (Zhang and Sohngen 2018).

Also, site closure due to algal blooms would cost 2.3 USD and 23.5 USD per trip for beachgoers and anglers, respectively (Wolf, Georgic, and Klaiber 2017). Meanwhile, alleviating algal blooms to one less mile boating could increase angler's surplus up to around 6.8 million per year (Zhang and Sohngen 2018).

It is vital to anticipate households' welfare change in response to policies that improve water quality to provide a viewpoint for policymakers. The consequence of decreasing phosphorus loading by 40% in Lake Erie was studied from two perspectives (Wolf et al. 2019; Zhang and Sohngen 2018). These studies shared 4 USD to 6 USD and 4.3 USD million annual welfare gain. There are likely causes for the differences between the results, using RP and SP, using different indicators for algal bloom, and the model they applied.

Moreover, the effects of TMDL (i.e., put an annual limitation on runoffs) for the Chesapeake Bay, was assessed across lakes in Virginia (Van Houtven et al. 2014). Households are willing to pay 60 USD per year to decrease the mean of total nitrogen, phosphorus, and Chlorophyll up to 18% (Van Houtven et al. 2014).

Another article studied the effects of water clarity improvement that influence individuals' perceptions directly (Keeler et al. 2015). Individuals are willing to incur greater travel costs by travelling to further destinations to experience one more meter water clarity (Keeler et al. 2015). On the other hand, one meter of water clarity can give rise to the number of annual trips for around 1390 (Keeler et al. 2015).

HAB and water clarity differ not only in physical attributes but also in the way in which they drive an individual's' behaviour. Whereas anglers are willing to pay between 8 USD and 10 USD for boating in around 1.6 kilometres less algal bloom water (Zhang and Sohngen 2018), lake users are willing to pay 22 USD per trip for one more meter of water clarity (Keeler et al. 2015). Anglers tend to incur higher costs for improving the water quality in comparison to beachgoers (Wolf et al. 2019). Therefore, put these finding together indicate that water clarity affects an individual's welfare more and consequently has a greater willingness to pay in comparison with harmful algal blooms.

Murray et al. (2001) assessed the visitors' welfare in response to removing beach advisories. The estimated willingness to pay was a function of how the beach visitors were informed about the beach advisory status and ranged between 24 USD and 38 USD per person for controlling beach advisory during a season. Hence, the travellers who used the media to monitor the water condition (presence or absence of the beach advisory at the site), gained less benefit in comparison to the travellers who check the status at the place for removing one beach advisory (Murray, Sohngen, and Pendleton 2001).

Recent research has revealed that the value of lake resources are not distributed equally among different lakes (Egan et al. 2009). The findings of Egan et al. (2009) study suggest the benefits of a significant improvement in the water quality of a few lakes will be as great as the benefit of a modest improvement in the water quality of all lakes. Thus, for policymakers, it is vital to recognize the priority of lakes water quality improvement. The last column of table A.1 shows the findings of some of the most related studies in the recreation demand area.

2.7 Summary

Generally, studies published to date show a positive relationship between recreation demand and water quality. This might be because of the 'publication effect,' or that null results are not published. In light of published studies in valuing water quality and recreational demand, it is becoming extremely difficult to ignore the existence of a positive relationship between water quality and recreational demand. Table A.1 presents a summary of some of the discussed studies in consideration of the variables and model they employed as well as their findings.

Since assessing the effect of water quality on recreational demand and individuals' behaviour might vary across different water quality indices, households' willingness to pay for controlling each of these indices might be different as well. Households' welfare and willingness to pay for improving the used water quality measures will be reviewed in chapter four.

Chapter 3
Data

Recreational demand and travel cost modelling and its associated relationship with individuals' behaviour has not been reflected by a lot of studies in Canada. However, some studies documented recreation demand and various water quality measurements in the United States. The following part of this paper moves on to describe the data I considered and ultimately used.

The study area for this research includes 55 provincial parks in Alberta. Alberta's lakes vary from glacial lakes to small shallow lakes and large lakes with sandy beaches. The lakes' distribution is not randomly distributed across the province, and the majority of them are in the northern half of the province. Figure 3.1 illustrates the distribution of parks and lakes across Alberta. Alberta's lakes provide different recreation opportunities, including swimming, boating, fishing different fish species, and amazing sceneries. Campgrounds are equipped with different amenities to provide better accommodation for travellers. Whether a campground has a boat lunch, beach, equestrian, and any other feature has offered various options for the campers.

In this research, the data sets played a significant role in determining how the study was structured. Data for this project are composed of two principal sources; recreational data and water quality data for years between 2014 and 2018. Along with the substantial records of recreational data, the various types of lakes and recreation activities have made Alberta a competent case of study since it includes different tastes of preferences.

Figure 3.1Parks and Lakes across Alberta (Alberta and Parks 2019)

3.1 Recreational Data

Investigating individuals' behaviour is one of the components of studying the relationship between lake water quality and recreation demand. In this study, I use records of travellers' trips, travellers' locations, trip starting points, site locations, destination, all of which are essential to assess the cost of travelling. Therefore, knowing the individual's postal code and the selected site as a recreation destination are vital. However, more variables than the spatial factors could help study the relationship between the water quality and individual's behaviour in consideration of their gender category or age group.

The present study used the RP dataset collected by an online reservation system that launched in 2010 by Alberta Environment and Parks (AEP) to facilitate recreation trips. Individuals can read the features of campgrounds and see some photos of them online. Available dates and costs of staying per night are presented on this website. Usually, most of the campgrounds are available to book between May and September.

The recreational data were obtained from the online reservation system provided by Reserve Alberta Parks (RAP) for provincial campsites across Alberta. This dataset was derived from individuals' actual decisions and gathered by asking travellers to report their postal codes, number of campers and stay duration while they were reserving a campsite. The original data include actual reservations to over 200 provincial campgrounds from over 70,000 people per year.

Unique identification keys assigned to each person allow us to link individuals across choice occasions, track their decisions over time. From around 1,175,500 total trips in the database, a subsample of more than 930,000 trips was drawn from screening the revealed preferences between late May and September. Not only were the majority of campsites booked during this period, called the open water season, but the recreation water condition also monitors during this time, which coincides nicely.

One of the limitations with existing recreational activity datasets is that the researchers did not have information on the date that individuals travelled to a campground. In most cases, researchers had information on the number of trips per year, and in a few instances, they had records of trips per season or month. The benefit of the RP dataset used in this study here is that it provides records of camping trips along with their exact times, including day, month, and year. Lakes' water quality is volatile during the summer. Beach advisories are usually issued between late May and September and pose limitations on water consumption periodically. Knowing if an individual had a trip early in May when there was no beach advisory at the site or before issuing an advisory in August is important for accurately identifying the effect of advisories in trip behaviour. Therefore knowing the exact date of the trips is useful for linking the individuals' actions to the lakes' water quality conditions, and so, it is critical for assessing individuals' behaviour.

The data provide records to various campground types, including group, comfort, and equestrian campsites. I used the following exclusions on the destinations by filtering campgrounds with special amenities. Equestrian campgrounds offer special facilities, riding trails to the travellers who have horses to enjoy their accommodation. Comfort camping is a convenient alternative for traditional camping, and its fee is between \$55 and \$170 per night. Group camping supports a large number of visitors. Approximately 2.6% of observations identified as group, comfort, and equestrian travels were removed from the sample in the screening.

People usually go camping on weekends or national holidays. Thus, for the evaluation of water quality and individual behaviour, a weekly span is considered. Regarding the first day of the week, since there are four long weekends between May and September (i.e. fixed on Mondays), Tuesday, Wednesday, Thursday, and Friday could be competent for the beginning of the weeks. However, Canada day has a fixed date (first of June), so the holiday was on different weekdays every year (i.e. Tuesday, Wednesday, and Friday). Therefore, in order to capture the effect of this national holiday, Wednesday and Thursday are the two best choices for the week starting day. In this study, Wednesday is considered the first day of the week to capture the effect of long weekends. 2.9% of the reservations that were made for more than seven nights were removed from the data set.

Although I could estimate the access cost of reaching a campsite through various methods, I calculate the travel cost for one car with a maximum capacity of five people with the help of the household's postal code and its linked annual income. This type of calculation is employed since assigning individual travel costs with big groups would be complicated and challenging. Thus, 3.5% of reservations with more than five people were removed from the dataset.

After eliminating these observations, unavailable records of the number of nights and the number of campers were replaced by their mean values. Table 3.1 presents a summary of the trip's information use in this study. Reservations with missing postal codes or missing travel distances or annual household income were removed from the sample, and that is the reason why travel counts declined in 2017 and 2018. In general, each individual had one trip in a year.

Table 3.1 Summary of Trip Information

Variable	2014	2015	2016	2017	2018
Total Number of Trips	87,161	97,284	94,148	55,701	46,978
Average Number of Trips (per person)	1.15	1.14	1.13	1.18	1.18

In this study, I assume that each person could only have one trip per week. Thus, if an individual had more than one reservation in a week, I randomly picked one of them. Approximately 217,000 reservations were removed from the sample at this stage. Since not all sites were open during the open water season of all these five years, I only consider the campgrounds that had at least one reservation per year. I impose the same attitude to individuals' availability, persons with no reservation in a year were removed from the relevant year's sample. These incorporations improve the model accuracy in terms of eliminating persons who have moved or stop camping due to any reason, and campsites closure. In addition, for assessing the opportunity cost of time and access cost, I assume that trips were not multipurpose, for instance, not for visiting family while they were travelling. After these modifications, I start the analysis with more than 830,000 trips to 73 campsites.

3.2 Water Quality

The second portion of the data consists of water quality measurements for the campsites. Lake water quality can be assessed based on various indices, and each of them can be a collection of different elements or a single factor. Blue-green algal blooms, a type of bacteria named cyanobacteria, may pose potential health risks to humans, animals, and also affect scenery. Blue-green algae spread quickly, mainly when it is sunny, or the temperature is high. Once the cyanobacteria propagate, the colour changes to red, light blue, or tan. The colour alteration, as well as the transparency, might be evident visually for the campers ("Blue-Green Algae | AEP - Environment and Parks" n.d.).

Figure 3.2 illustrates the use of remote sensing technology to measure the qualitative parameters of water bodies. Chlorophyll concentration is one of the most commonly measured water quality parameters. In this figure, Chlorophyll, as an indicator of algal biomass, shows blue-green algae growth across Pigeon Lake, one of the most popular recreational destinations in Alberta in 2018. High algal biomass became apparent in the shoreline in mid-July and spread all across the Lake in September. Then, by ending the open water season and changes in weather conditions, it started to disappear by October. This figure shows how the lakes' water condition changes over several days and asserts the importance of employing the exact day of the trips.

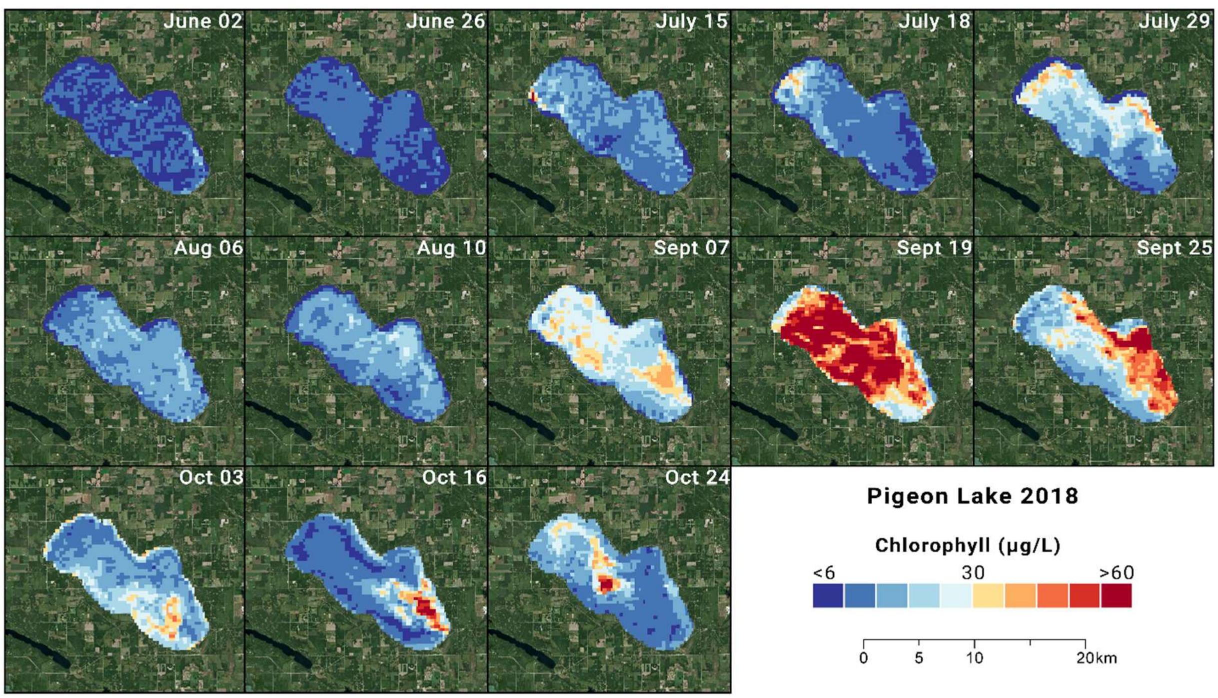

Figure 3.2 Blue-green algae growth in Pigeon Lake, 2018

In this study, the water quality data is composed of two different water quality representatives; beach advisory and cyanobacteria cell count from late May to September, but ultimately only I used one as explained below. Both of these indicators are concerning with cyanobacteria in different ways, and it is one of the components of the beach advisory.

3.2.1 Beach Advisory

Alberta Health Services issues advisories when the water quality does not meet Alberta's recreational water standards, including sample results, photographs, local conditions, complaints and cyanobacterial blooms ("Alberta Health Services," 2019). Advisory signs remove once the first acceptable sample results are received. Otherwise, it remains in place until the end of the open water season or November as blooms can disappear when the weather condition changes ("Alberta Health Services," 2019).

The advisory signs are posted on the beach or water body to inform beachgoers that the presence of blue-green algae can pose a risk to human health, and the water situation is not suitable for particular purposes like swimming. All active advisories messaging are posted online at www.albertahealthservices.ca (Health Advisories section) as well. Campers could check out the Alberta Health Services website to obtain information about the water quality condition. Beach advisories can derive individuals' behaviour by officially informing people both at the site and online. Therefore, I can examine how the beach advisories drive an individual's decision.

However, this water quality indicia represents a binary condition for water; it is suitable when there is no beach advisory, and it is not appropriate when there is a beach advisory. Therefore, it is not possible to have a continuous report of the water condition, which counts as a limitation associated with this water quality measurement. Here, in figure 3.3, is the sample of advisory signage (Alberta and Alberta Health 2014).

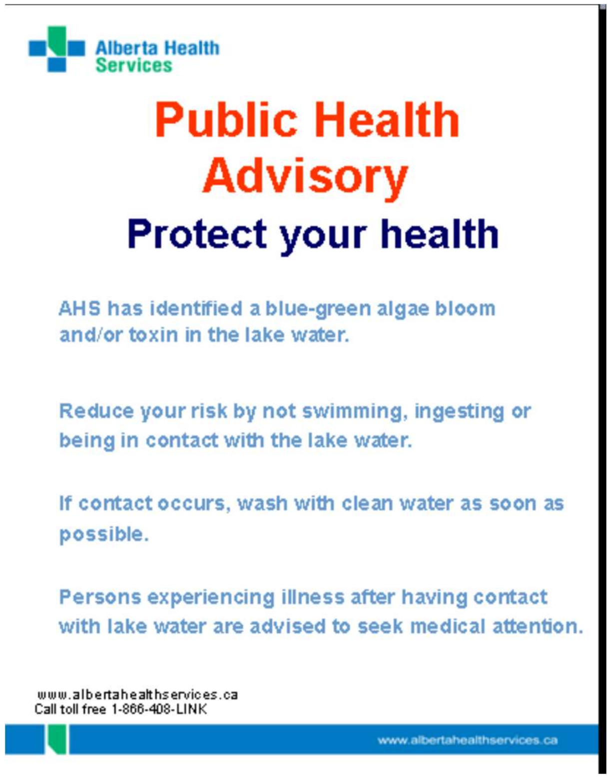

Figure 3.3 Advisory Signage

I had a complete dataset of whether there was an advisory on the beach or not. Hence, I assigned a dummy variable as a representative of a beach advisory presence for each week. Figures 3.4 and 3.5 represent the total number of weeks that beach advisory and the number of campsites with beach advisories were present during the years between 2014 and 2018, respectively.

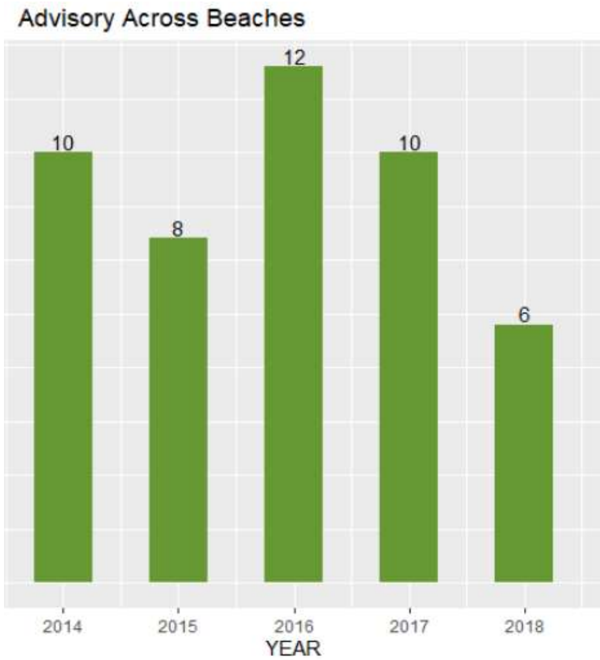

Figure 3.5 Number of Campground with an Advisory each Year

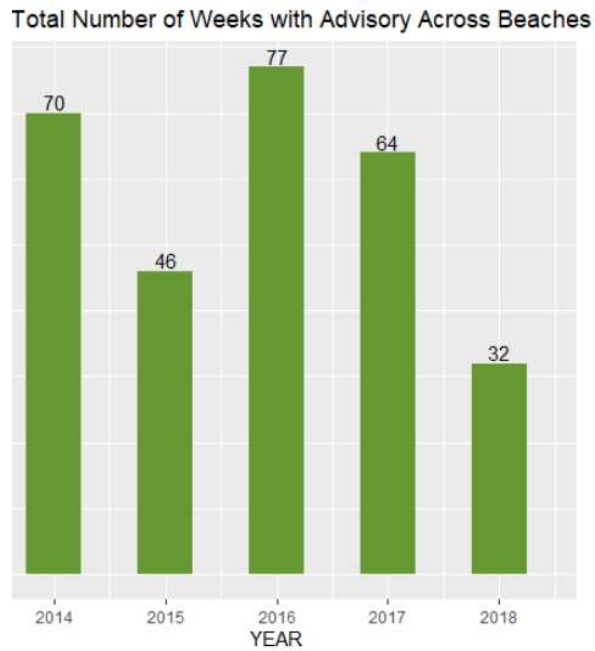

Figure 3.4 Total Number of Weeks with Beach Advisories in each Year

Alberta Health Services conducted Alberta's provincial parks water quality readings, and I merged this information with site attributes provided by Alberta Environment and Park (AEP).

3.2.2 Cyanobacteria Cell Count

The open water season often begins with the May long weekend, Victoria Day, and ends with September long weekend, Labour Day, at recreational areas in Alberta. Alberta Health Services reports biological characteristics of cyanobacteria cell count (cell/mL) for some of the lakes across Alberta throughout the open water season, between the months of May or Jun and August or September, which coincides nicely with the camping dataset.

Cyanobacteria can be loosely described as blue-green algae. When the algal blooms die, microorganisms feeding and decomposing them. Therefore, available oxygen in the water consumed and aquatic animals encounter with oxygen scarcity ("Facts About Water in Alberta" 2010). Moreover, in the case of human contact with cyanobacteria, irritation, intestinal side effects, or severe illness may occur ("Blue-Green Algae | AEP - Environment and Parks" n.d.).

The literature on the effect of water quality and recreational demand has highlighted several water quality measurements. In the absence of a determined single best criterion, Zhang and Sohngen suggested HAB as a significant water quality indicator (Zhang and Sohngen 2018).

In this study, I employed cyanobacteria as a microorganism that can produce HABs, and as I expected, there is a positive correlation between beach advisory and cyanobacteria cell count. Still, nevertheless, the impact of the blue-green algae has received scant attention due to the difficulty of data collection. Alberta Health Services collected ten samples along the length of a beach with a depth of 1 meter. Then these ten samples were mixed and shipped to the Alberta Centre for Toxicology (ACFT) for further analysis (Alberta and Alberta Health 2014).

In the dataset, one of the challenges associated with cyanobacteria cell count is that I had a lot of missing values for this continuous variable, and I did not have any information for more than half of the lakes. Figure 3.6 illustrates the weekly mean of the logarithm of cyanobacteria at different levels. This graph was formed on linear imputation of the weekly mean for the lakes with at least two cyanobacteria records during the open water season. The second challenge associated with using cyanobacteria cell count is that it is a noisy variable and might alter day by day since different environmental factors, including rainfall, temperature, wind, and sunlight, affect its propagate.

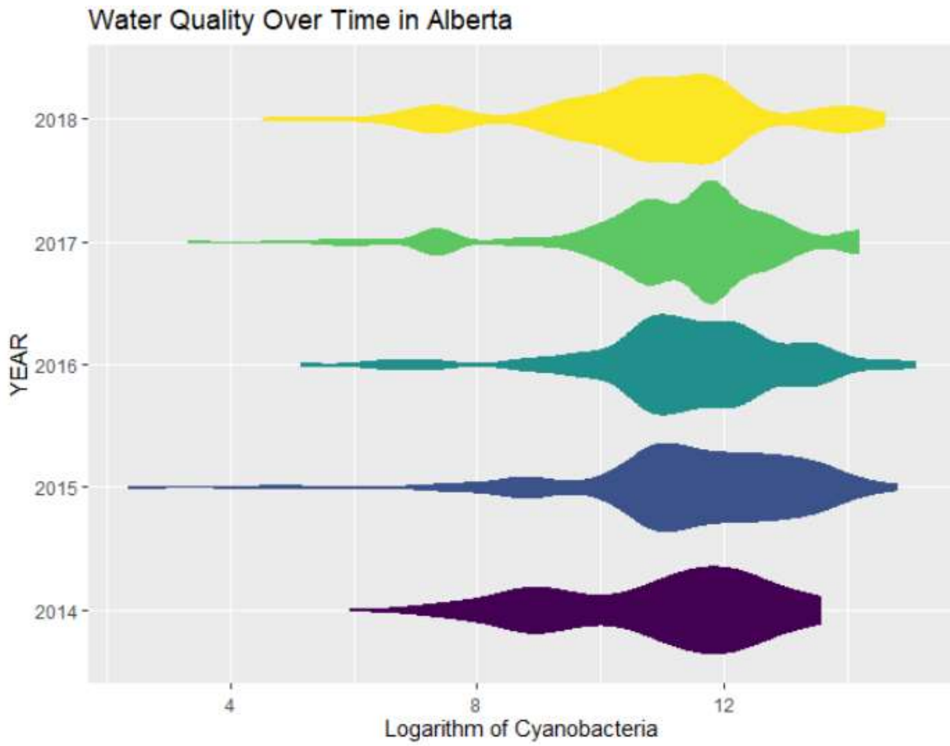

Figure 3.6 Logarithm of Cyanobacteria over time in Alberta

The logarithm of cyanobacteria is used to transform this noisy variable with skewed records to approximately conform to normality for all lakes with at least two logs in a year. This figure shows the distribution of the logarithm of cyanobacteria at different levels.

Figure 3.7 shows that there is a considerable variation in this variable during each week and why it was hard to impute the missing values in consideration of data availability and unpredictable pattern. I ran different imputations to predict the missing values for cyanobacteria cell count and tried to find an accurate interpolation. Since there are a lot of missing values, I couldn't use it in the model. Figure A.1 in the Appendix section represents the available data of cyanobacteria for the lakes with less than nine weekly records in Alberta.

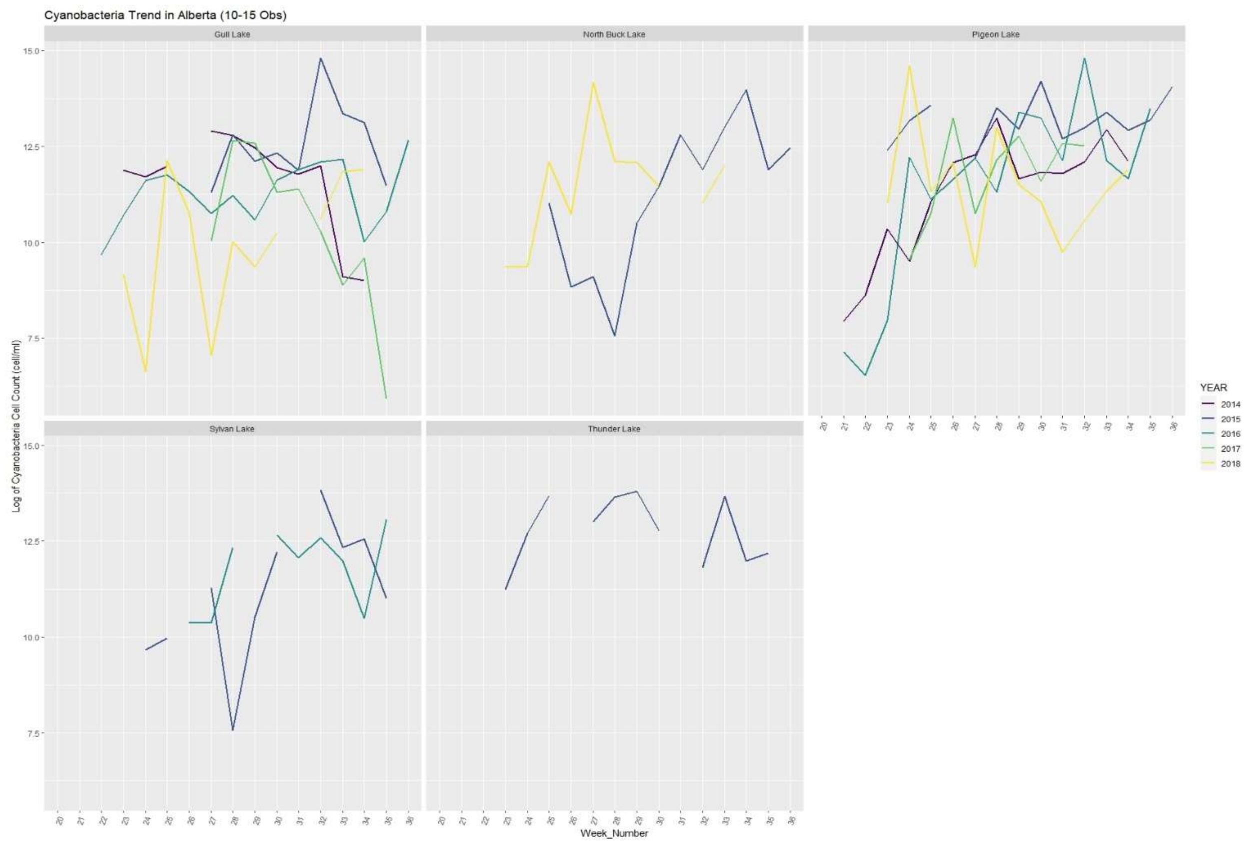

Figure 3.7 Cyanobacteria Trend for the Lakes with at Least 10 to 15 weekly records in Alberta

3.3 Summary

Initially, I wanted to know what water quality measurement, cyanobacteria cell count or beach advisory, affect recreational behaviour. However, having numerous missing values for the cyanobacteria did not allow us to use this variable in the analysis. Moreover, the volatility of this variable made the interpolation complicated in a way that was beyond the scope of this study. Therefore, for the rest of the analysis, I only focus on beach advisory.

By combining RP on recreation behaviour with data on beach advisories, I created a novel dataset that allows us to estimate recreation demand and water quality across Alberta. Eleven campsites do not have a lake, 37 have one, and nine parks have more than one campsite, and among them, four of them have more than one lake. Table 3.2 provides a summary of the sites' specifications that use in this study for more than 380,000 trips that had travel costs from 154,000 people to 73 campgrounds and 48 lakes. Beach advisories were issued for 6% of the choice occasions, which is pretty rare. All data management and analysis were performed using R statistical software.

Table 3.2 Summary of The Sites Used in Analysis

Variable	Total Number			
Number of Campers	154,653			
Parks	55			
Lakes	48			
Campground	73			
	Mean	Std. Dev.	Min.	Max.
Beach Advisory (Dummy)	0.06	0.24	0	1

Chapter 4
Methodology

4.1 Introduction

In order to evaluate the impact of various water quality indices on recreation demand, I have to translate the factors that affect individuals' decisions or behaviours into meaningful economic values. Welfare and utility could be the two critical elements in terms of making a decision, which is constrained by exogenous environmental conditions, budget, time and some other factors. The economic theories that help us to evaluate the mentioned values will be explained in the following section. First, I will describe the conceptual model and then the empirical approach.

4.2 Conceptual Model

4.2.1 **Random Utility Model**

The travel cost model can be employed to assess the economics of recreation demand. The demand for travelling to specific campsite increases as the price or travelling cost decreases. Traditional TCM estimates demand to visit a single-site by considering the cost of travelling and the frequency of the trips and does not reflect the welfare associated with each site (Parsons 2014). However, the number of trips which might be equal to one or greater than one, and a discrete number of visits were problematic in TCM. The problems were solved by the count model approach in on-site recreation demand studies (Englin and Shonkwiler 1995). For outdoor recreation demand models, it is essential to consider that how does the model assess behaviour in the baseline condition, and predict behaviour changes in response to quality changes when imposing different scenarios (von Haefen and Phaneuf 2003). In this study, I have multiple sites, and therefore the substitution effect is going to be an essential element in predicting an individual's behaviour. However, the count data demand system model does not explain the substitution probabilities between alternatives (Englin and Shonkwiler 1995). Kuhn-Tucker framework has resolved both weaknesses associated with the TCM model, and it is applicable for seasonal trip data considering the total number of trips to each campground (Haefen and Phaneuf 2005). This model is not feasible to employ in this study because of water conditions and, consequently, beach advisories that change over time, and I can not use the seasonal demand model.

Discrete choice RUM has improved TCM by incorporating qualitative differences across sites that would influence campers' decisions. Moreover, the seasonal trips get divided into smaller choice occasions where an individual makes an independent choice on each of them (Haefen and Phaneuf 2005). I use RUM as one of the most popular frameworks in non-market valuation in this study area since this model helps to assess how different factors influence recreation demand with a large number of alternatives. Hence, I can examine the importance of lakes' water quality to recreationists within a set of known alternatives.

Early examples of research into the discrete choice model emerged in transportation preferences (Domencich and McFadden 1975; Ben-Akiva, Lerman, and Lerman 1985). In this study, I am assessing an individual's preferences in taking a recreation trip to a campground or substituting it with any other activity. I employed a repeated discrete choice model framework that includes a stay at home option. Based on this model, I assume that the open water season t is divided into weekly periods in which each person can have at most one trip and decides among alternatives to obtain the maximum utility (Morey, Rowe, and Watson 1993). In fact, I improved the discrete choice model by adding the stay at home option to the choice set J. Therefore, each person can choose from a choice set of J options, while the alternatives are mutually exclusive, as discussed in chapter 3, I assumed each person had one trip in a week. So, an individual can select one of the alternatives in a given choice occasion. Stay at home option is a representative of any activity other than travelling to a campground. Thus, the choice set is comprised of two major categories, including the stay at home option or visiting one of the 73 provincial campsites, which are denoted by J.

Based on the RUM, individuals decide between staying at home or travelling to a campsite based on the attributes of the campsites, time-varying variables, and their obtained utility. The underlying concept is that people tend to maximize their utility U_{ijt} $(Q_{i1t},\ Q_{i2t}, \dots, Q_{ijt})$, where Q_{i1t}, $Q_{i2t}, \dots, Q_{ijt}$ are the campsites' attributes. Based on equation 4.1, the person i will travel to campground j once he obtains a higher level of satisfaction by visiting campground j rather than k in choice occasion t.

$$U_{ijt} \geq U_{ikt}, \quad where\ j, k \in choice\ set\ J, j \neq k \tag{4.1}$$

The utility of an alternative is described as

$$U_{ijt} = U(Q_{ijt}, TC_{ijt}, \varepsilon_{ijt}) \tag{4.2}$$

Where

- U_{ijt} is composed of Q_{ijt}, a vector of attributes for alternative j;
- TC_{ijt} is the travel cost for a person i to access alternative j in a choice occasion t;
- ε_{ijt} a random vector of tastes for a person i embodied in alternative j in a choice occasion t, that is known to the individual but not to the researchers.

As explained earlier, the RUM has two components; a systematic factor that is observable for campers and also for us as researchers and a random portion that is observable only for the campers. Therefore, individuals have more information in comparison with the researchers. The systematic factor is deterministic for the researchers and the campers; in this study, it depends on the cost of travelling, the water quality, and some other attributes. The stochastic term has made the utility as a random variable for the researchers with the form of

$$U_{ijt} = V_{ijt} + \varepsilon_{ijt} \tag{4.3}$$

Where

- U_{ijt} is the utility that person i obtain by travelling to the campground j in choice occasion t;
- V_{ijt} is the systematic factor for person i to travel to the campground j in choice occasion t;
- ε_{ijt} is the random portion for person i to travel to the campground j in choice occasion t.

I can assess the campers' tastes and preferences, which relate to the systematic factor. However, evaluating the taste variation associated with the stochastic portion is not applicable (Train 2009a). As reflected in equation 4.4, each individual chooses a site through which a higher level of utility is provided.

$$V_{ijt} + \varepsilon_{ijt} > V_{ikt} + \varepsilon_{ikt}, where\ j, k \in choice\ set\ J, j \neq k \tag{4.4}$$

Since the systematic factor is observable, the probability of choosing site j rather than site k is

$$\varepsilon_{ijt} - \varepsilon_{ikt} > V_{ikt} - V_{ijt} \ . \tag{4.5}$$

Therefore, the probability that an individual will visit site j is equal to the probability that the random portion of the utility from the site j is higher than the random portion utility to all of the other sites in the choice set J, including alternative k. Equation 4.6 shows the probability of choosing one site rather than the other site.

$$prob_{jt} = prob(\varepsilon_{ijt} - \varepsilon_{iKt} > V_{ikt} - V_{ijt}). \tag{4.6}$$

The following equation shows the choice probability of choosing an alternative j among the choice set J implied by the simple MNL[3]. Thus, the probability of choosing one alternative j among various alternatives is equal to the expected utility of travelling to j by the summation of obtained expected utility of going to other campsites.

$$P_{ijt} = \frac{\exp(V_{ijt})}{\sum_J \exp(V_{iJt})} \tag{4.7}$$

As mentioned in chapter 2, several types of discrete choice models exist in practice. The logit model, as one of the most popular discrete choice models, assumes the distribution of the stochastic part is independently, identically distributed (i.i.d) extreme value. Also, the unobserved factors are not correlated with different alternatives, and they have the same variance (Train 2009b). Equations 4.8 and 4.9 show the density function for the type I extreme value error and the cumulative distribution, respectively.

$$f(\varepsilon_{ijt}) = \theta e^{-\theta \varepsilon_{ijt}} e^{-e^{-\theta \varepsilon_{ijt}}} \tag{4.8}$$

[3] For further discussion on a derivation of this probability please see (Haab and McConnell 2002).

$$f(\varepsilon_{ijt}) = \theta e^{-e^{-\theta \varepsilon_{ijt}}} \tag{4.9}$$

Where the size of θ reflects how much information researchers have on the systematic factor in comparison to the stochastic portion. As θ increases, the variance also increases, and usually, it is normalized to one for convenience (Bockstael and McConnell 2007).In some cases, it is possible to estimate different values for θ, which is called the NL model, and I will discuss it in the following subsection.

4.2.2 **Nested Logit Model (NL)**

In some cases, the Independence of Irrelevant Alternative (IIA), counts as an essential issue in the MNL model. Based on this assumption, the probability of choosing between two alternatives is not affected by the third option. Suppose several campgrounds are closed due to poor water quality, travellers to those campgrounds will get divided by the available campsites. Therefore, I would expect the probability of travelling to available campsites with close features would increase more in comparison with the probability of staying at home. The IIA assumption seems implausible since some people might not give up camping to substitute their activity with any other sort of activities which are captured by a stay at home option. In other words, the probability of replacing a closed campsite with another campsite would be different from substituting it with staying at home. For that reason, in some cases, the IIA only holds across the alternatives with mutual characteristics.

The Generalized Extreme Value (GEV) model has been introduced to help relax the IIA assumption issue by assuming generalization of the extreme value distribution and leads to the NL model. GEV model categorizes alternatives with the same correlation into groups called nests (Train 2009b). The advantage of the GEV model in comparison with the MNL one is that while the IIA still holds for the alternatives of the same nest, it does not hold across the nests; the Independence from Irrelevant Nests (IIN) holds across the nests (Train 2009c). Hence, different correlations exist across alternatives in the GEV model, and it is consistent with jointly distributed errors (Bockstael and McConnell 2007).

The NL model categorizes alternatives into different groups, called nests, with common attributes. Depending on the case study, the number of stages might vary between two or higher, and the alternatives are classified in a diagram that looks like a tree with branches. Figure 4.1 represents a simple two-stage NL model tree diagram. Here, individuals make a binary choice between staying at home and travelling to a campsite, which is called the upper level of the decision. And in the second level, which is called the lower level, the individuals choose among a choice set of campsites conditional on selecting camping alternative. Therefore, the whole choice set is divided into two non-overlapping nests (n), where $n = 1$ implies travelling and denoted by B_k, and $n = 2$ means staying at home option and denoted by B_l. IIA assumption exists across the lower level ($j_1, j_2, \dots, j_{73}$), while it does not hold across all alternatives (stay at home and travelling to campsite choices), where j_{74} represents the staying at home option. Figure 4.2 represents the decision tree of the MNL model in which all alternatives are treated equally, and the IIA assumption holds.

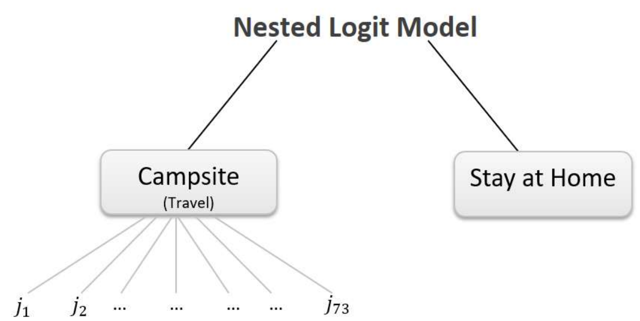

Figure 4.2 Two-Stage Nested Decision Tree

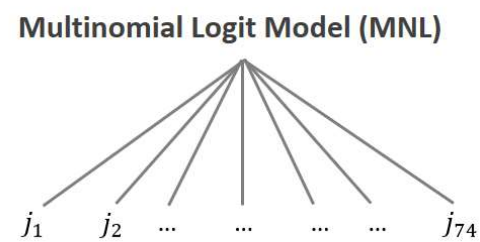

Figure 4.1 MNL Decision Tree

In the NL model, it is assumed that the vector of $\varepsilon_{it} = (\varepsilon_{i1t}, \ \varepsilon_{i2t}, \dots, \varepsilon_{ijt})$ has a cumulative distribution, which is shown in equation 4.10 (Train 2009c; Bockstael and McConnell 2007).

$$F\left(\varepsilon_{i1t},\ \varepsilon_{i2t},\dots,\varepsilon_{ijt}\right) = \exp\left(-\sum_{n=1}^{N}\left(\sum_{j_n=1}^{j_n}\exp\left(-\frac{\varepsilon_{ij_nt}}{\theta_{nt}}\right)\right)^{\theta_{nt}}\right) \tag{4.10}$$

Where N is the number of nests, j_n is an alternative within the n nest, θ_{nt} measures the degree of independence in unobserved utility across the nests. Or I can say $1-\theta_{nt}$ shows a correlation within the nests. $\theta_{nt}=1$ represents complete independence or no correlation across all J alternatives, and the model collapses to the MNL model. However, $\theta_{nt}<1$ represents correlation exists, and the NL is a competent model (Train 2009c). Based on the IIA assumptions:

$$\begin{cases} cov\left(\varepsilon_{ijt},\varepsilon_{imt}\right)=0, & \{\forall\, j\in(n=1), m\in(n=2), j\neq m\} \\ cov\left(\varepsilon_{ijt},\varepsilon_{ikt}\right)\cong 1-\theta_{nt}\ , & \{\forall\ j,l\in(n=1), j\neq k\ \} \end{cases}. \tag{4.11}$$

Since IIA exists within the nests and removing an alternative in the nest increases the probability of travelling to other campsites then $\theta_{nt}=1$. Therefore, the cumulative distribution for the disturbances in equation 4.12 changes to

$$F\left(\varepsilon_{i1t},\ \varepsilon_{i2t},\dots,\varepsilon_{ijt}\right) = e^{-e^{-\varepsilon_{ij_nt}}}. \tag{4.12}$$

Each person decides whether to go camping and where to go or not, and I need to estimate an individual's decision in two levels, participation choice and site choice. The basic idea of this model is to group comparable alternatives and then structure a choice set of each group. How the nests are formed plays a significant role in the results. I categorized the potentially closed alternatives in the united nest. The nested logit model partitions the choice set of all campsites and stay at home into two major categories of "camping" and "staying at home" as the first level, participation decision, meanwhile "camping" divided into subsections of several alternatives for each campsite, site decision. In other words, I used the NL model since I can divide and examine the sample into different stages. In the two-level NL model, each nest is a composite of alternatives, which is depicted in figure 4.3.

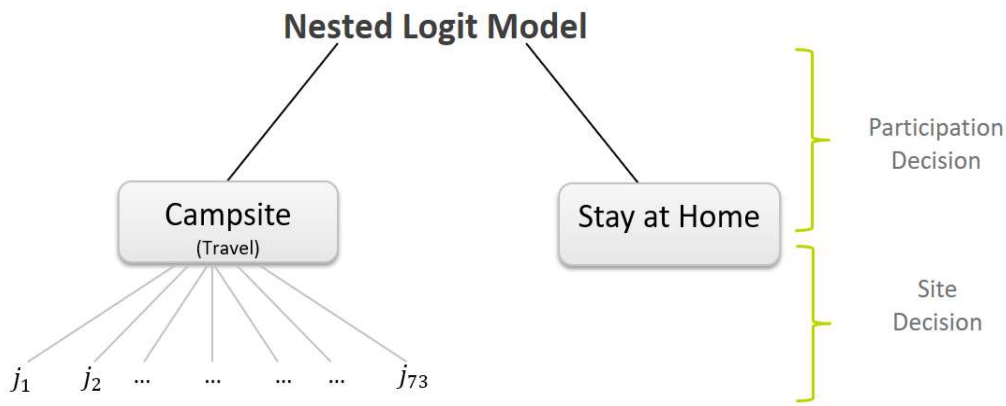

Figure 4.3 Tree Diagram Two-stage Nested Logit Model

Equation 4.13 shows the probability of choosing a campsite in the two-stage NL model that is a probability of two standard logit models (Train 2009c). Probability in the NL model can be divided into two levels; the probability of selecting the nest with alternative j ($p_{ijt|B_k}$) and the probability of selecting j in that nest (p_{n,B_k}) based on Bayes theorem[4].

$$p_{ijt} = p_{ijt|B_k} * p_{i,B_k} \tag{4.13}$$

Actually, equation 4.7 is the product of equation 4.14, where $\theta_{nt} = 1$.

$$P_{ijt} = \frac{\exp\left(\frac{V_{ijt}}{\theta_{nt}}\right)\left(\sum_{j\in B_k}\exp\left(\frac{V_{ijt}}{\theta_{nt}}\right)^{\theta_{nt}-1}\right)}{\sum_{n=1}^{N}\left(\left(\sum_{j\in B_l}\exp\left(\frac{V_{iJt}}{\theta_{nt}}\right)^{\theta_{nt}}\right)\right)} \tag{4.14}$$

Therefore, equation 4.15 represents the conditional probability of the lower level (B_k).

---

[4] $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

$$p_{ijt|B_k} = \frac{\exp(^{V_{ijt}}/_{\theta_{nt}})}{\sum_{j \in B_k} \exp(^{V_{iJt}}/_{\theta_{nt}})} \qquad , \; j \in \{j_1, j_2, \dots, j_{73}\} \tag{4.15}$$

Where V_{ijt} is the expected utility from travelling to a campsite in consideration of the beach advisory and the travel cost. Therefore, equation 4.15 reflects the probability of choosing a site, including the variables that vary across the campgrounds, such as water quality, travel cost, and unobserved factors.

Inclusive value helps us to join the conditional probability of choosing j in the nest B_k with the marginal probability of choosing alternative j in the nest B_k. Inclusive value (IV_{it}), also named log-sum of branch B_k, is the logarithm of the denominator of the conditional probability (equation 4.15) and the expected utility for visiting a campsite.

$$IV_{it} = \ln\left(\sum_{j \in B_k} \exp(^{V_{iJt}}/_{\theta_{it}})\right) \tag{4.16}$$

Therefore, the marginal probability, as an upper level, where include the stay at home option is

$$p_{i,B_k} = \frac{exp(\, Z_{it}\, \alpha + \theta_{nt}\, IV_{it})}{\sum_l exp(Z_{it}\, \alpha + \theta_{nt}\, IV_{it})} \tag{4.17}$$

Where Z_{it} is the expected utility person i obtain by staying at home in choice occasion t and α is a vector of coefficients on variables in the upper level. Equation 4.18 shows if $\theta_{nt} = 1$ then I should use the CL model and if $\theta_{nt} < 1$ then I should use the NL model. Therefore, the NL model estimates an additional parameter compared to the conditional logit model. θ_{nt} the coefficient of the inclusive value assists in choosing the competent model if its value is one or less than one.[5] The NL model can be estimated using the Maximum Likelihood Estimation (MLE).

[5] This discussion largely follows that in (Train 2009c).

$$If \begin{cases} \theta_{nt} = 1, & Conditional\ logit\ Model \\ \theta_{nt} < 1, & Nested\ logit\ Model \end{cases} \tag{4.18}$$

4.2.3 Travel Cost Calculation

Assessing the economic value of each person's trip to different destinations can facilitate quantifying economic impacts. Access cost or the monetary cost and the opportunity cost of time are both captured in travel cost modelling in equation 4.19. Figuring out access cost is possible by computing the distance between a household's stated postal code and the campsite's location. And the opportunity cost of time will be attainable through annual household income.

$$Travel\ Cost\ =\ Monetary\ Costs + Opportunity\ Costs\ of\ Time. \tag{4.19}$$

In addition to the actual trips dataset and water quality indices, other sources of data were used to calculate the travelling cost. Driving costs were obtained from the Canadian Automobile Association's driving cost calculator ("CAA National" n.d.). The estimated total annual driving cost was operated for three different categories of vehicles, i.e., Compact car, Sport Vehicle utility (SUV), and pickup truck. The driving cost per km includes fuel, depreciation, maintenance, licence and registration, and insurance costs, as well as a monthly car payment (Lloyd-Smith and Becker 2019).

I use the monthly average retail prices for gas across Alberta (Government of Canada 2018) to calculate the average gas price for the month between May and September. The calculation was conducted for the years between 2014 and 2018 to improve the estimation. Then, the average gas prices were plugged into the CAA driving cost calculator. Table 4.1 shows a summary of annual access cost.

Table 4.1 Summary of Access cost Variables

Variable	2014	2015	2016	2017	2018
Average Gas Price (per litre)	\$1.19	\$1.04	\$0.93	\$0.96	\$1.28

Driving Cost (per km)	\$0.47	\$0.46	\$0.44	\$0.45	\$0.48
Total Travel Cost	\$21,020,715	\$23,171,425	\$22,679,951	\$13,127,769	\$11,261,945
Average Travel Cost (per trip)	\$241	\$238	\$241	\$236	\$240

The access cost is estimated by considering road trips. The overall structure of the travel cost formula for a round trip has the following format

$$TC_{ikt} = \left(\frac{D_{ik} * 2 * DC_{it} + CC_{ikt}}{NC_{ik}}\right) + 2 * \left(\frac{2}{3} * \frac{I_i}{2080}\right) * h_{ik}. \quad (4.20)$$

Where

- D_{ik} is the distance between person i and the campground k (km);
- DC_{it} is the driving cost for person i in the year t (km);
- CC_{ikt} is the total cost of travelling for person i to campground k in year t;
- NC_{ik} is the number of campers that reserved by person i to campground k;
- $\frac{2}{3}$ is the opportunity of cost ratio;
- I_i is the median annual household after-tax income of person i for 2016;
- 2080 is the average hours worked hours in a year;
- h_{ik} is the hours take that person i travel to campground k, and both sentences are multiplied by two to capture the costs of around trip.

I used a two-third ratio based on the previous work that found that this assumption fits the data better than the one-third wage rate (Lloyd-Smith and Becker 2020). Due to the data limitation and the absence of sociodemographic information, I did not have any information on an individual's annual income to calculate the second component of the travel cost (equation 4.20). To overcome this drawback, I used the annual household after-tax income on the postal code level provided by the Canada Census.

Based on the preferred approach, household income is required to calculate the opportunity cost of time. With the help of the Canada census, the median annual household after-tax income for 2016 (“Statistics Canada: Canada’s National Statistical Agency” 2016) was linked to the reported postal codes to overcome the absence of household income. The median household income after tax was around $87,600.

Besides the inconsistency associated with the ratios of wage rate, there remains a debate in the literature on whether household or individual income should be used for travel cost calculations. Fezzi et al. (2014) studied individuals’ preferences between a faster route with a toll to save time or a slower route without a toll to save money. Based on individuals’ income, the authors found a reasonable value of travel time with the tree-quartered ratio of the wage rate (Fezzi et al. 2014). Since individuals’ income is lower than households’ income (English et al. 2018), incorporating a greater ratio would be reasonable. English et al. (2018) showed that the tree-quartered ratio of personal income is close to one-half of the household's income.

Besides this information, the entrance fees for each campground provided by AEP, and individuals have to pay the toll while they are making the reservation. Moreover, the campers charged $12 when making a reservation regardless of the reservation length. The number of campers the number of nights they spent in the campground, entrance and reservation fees were included in calculating the travel cost for each person to different destinations that are represented in equation 4.21.

$$CC_{ikt} = CF_k * NN_{ikt} + R_k \tag{4.21}$$

Where

- CC_{ikt} is the camping costs for person i to book campground k in year t;
- CF_k is the fee of campground k per night;
- NN_{ikt} is the number of nights that person i stayed in the campground k in year t;
- R_k is the reservation fee for the campground k.

The distance travelled, and the time spent was estimated by considering different combinations of the users’ stated home postal codes and campground locations (Lloyd-Smith and Becker 2019). By determining these variables as inputs and plugging them in equation 4.20, the travel cost for each person to each campground was computed. After all these calculations, I divided $\boldsymbol{TC_{ikt}}$ by 100 to help the convergence and joined the travel costs dataset with the revealed preference dataset. A summary statistics of the travels used in this study presented in table 4.2. Generally, on average, the campgrounds were booked by between 2 and 3 campers for 2 or 3 nights.

Table 4.2 Summary Statistics of The Travels Used in Analysis

Variable	Mean	Std. Dev.	Min.	Max.
Number of Campers (per trip)	2.48	2.30	1	5
Number of Nights (per trip)	2.67	1.01	1	7
Annual Household Income after Tax	$89,181	$28,835	$19,936	$355,328

4.2.4 **Model Estimation**

I study individuals' behaviour and the water conditions weekly, over the 17 weeks of the camping season. I use a repeated discrete choice model where each person has 17 choice occasions per year, and each person decides whether and where to camp for each week. The number of choice occasions was the same for all individuals, and an individual was included in each year's observation if that person at least made one reservation.

Beach advisories are issued based on different factors (i.e. local conditions, complaints, cyanobacteria, photographs, and sample results) and impose some limitations on water usage. An individual's perception and experience of beach advisory might vary from different situations with the same advisory since there is a threshold for issuing advisories; in some cases, the water quality might be far above of the criteria and close to the beach closure, and in some cases, the water conditions might be a little away from no swimming criteria. Incorporating travellers' experiences might be difficult because the beach advisory's severity is not distinguishable for beach advisory. Moreover, individuals' perceptions are different based on the type of recreation activities they have. In the absence of such information, it is hard to assess experience's effect on their next trip behaviour. Due to the concerns associated with endogeneity if past behaviour is explaining current choices, I assumed that all travellers are alike, the effect of past experience did not consider in this study. However, a study incorporated fishing experience along with other sociodemographic characteristics (Morey, Rowe, and Watson 1993).

In this study, the expected utility an individual obtains from visiting a site is a function of the travel cost, water quality, and some other time-variant as well as the site characteristics variables. Since I did not have information about all site characteristics and time-variant variables that might influence campers' decisions, I assigned a constant to each site and each choice occasion. Murdock (2006) proposed the MLE for controlling unobserved characteristics. Based on the suggested approach, the model mitigates biased travel cost since the potential correlation between travel cost and unobserved characteristics incorporated in the model estimation; therefore, I would have a more precise valuation (Murdock 2006). In the present study, I do not follow her approach directly, but I include ASCs to help control for omitted variables.

An ASC for each of the campsites is included to help control for other site-specific characteristics such as site congestion, lake size, and site access, which are not included in the model and might influence individuals' decisions. Several site attributes are known to affect campers' choice, e.g. whether the site has a boat ramp or not, how big the lake is, or how crowded the site is. If the model does not include these factors, the effect of the neglected factors might be reflected in the coefficient of the water quality measurement, where independent variables are correlated. For example, as the site congestion increases, the water quality might decrease. Therefore, the coefficient of the water quality measurement would be overestimated and express greater prominence than it really has.

I use the same approach to the time-variant variables, such as weather conditions. Since I did not incorporate any information about temperature or any other impressive time-variant factors, I assigned a TSC to each choice occasions. For instance, some lake's characteristics like cyanobacteria increase as the temperature increase, while incorporating TSC would help us extract the effect of time-varying variables from the beach advisories' coefficient.

Thus, ASC and TSC help control for any unobserved campground or week specific effects and figure 4.4 represents how I use them at different levels. Assigned specific constants to weeks and campsites, as a representative of the relevant time-variant and site characteristics, helped to avoid biasness in estimated parameters for travel cost and the advisory's existence. Moreover, a dummy variable assigns to each week to show whether there was a long weekend on each choice occasion or not.

An infinite combination of constants is possible for assessing the probability of choosing one site rather than the other one. One of the ASCs has to be normalized to zero to overcome this issue. Therefore, 72 estimated ASCs are interpreted relative to the normalized ASC (Train 2009b). Following the same approach for the assigned TSCs to 17 choice occasions, I normalize one of the time-varying constants and estimate 16 TSCs.

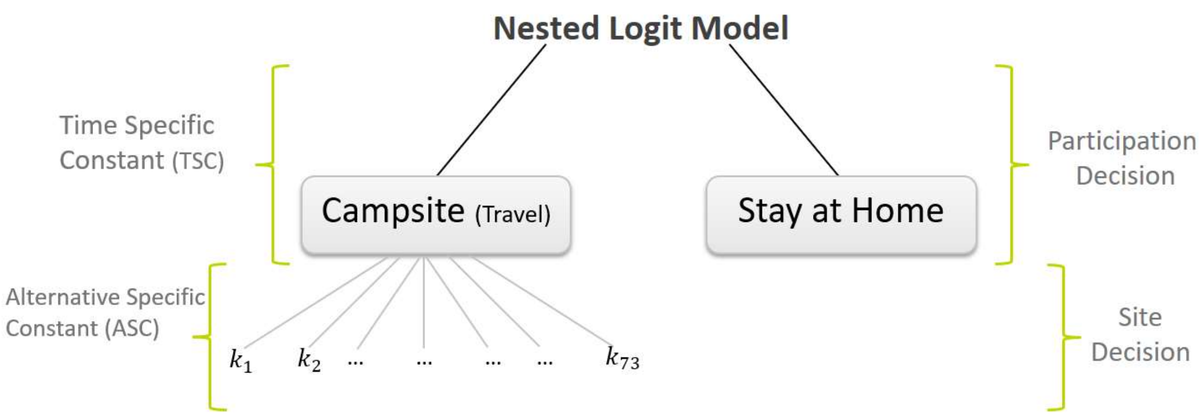

Figure 4.4 Tree Diagram Two-stage nested Logit Model and Specific Constants

Adding specific constants makes the RUM complex where the expected utility of choosing an alternative is a function of the obtained utility from each nest. The probability of selecting an alternative through each nest is represented in equation 4.22.

$$P_{choice} = P_{choice|site\ decision} * P_{choice|participation\ decision} \tag{4.22}$$

$P_{choic\ \ |site\ decision}$ depends on the variables vary across campsites, i.e., travel cost, water quality, and different attributes that are reflected in ASC. $P_{choice|participation\ decision}$ is a function of time-variant variables, i.e., long weekend and weather conditions.

The model parameter estimation has been performed using a newly released package in estimation software RStudio; Apollo package a powerful tool for the maximum likelihood estimation of discrete choice models. The dataset is required to have a specific wide format for estimating the model in Apollo. Moreover, lists of utility functions and campsite availability for each camper should be presented in the model specification file.

Working with a large dataset and assessing participation and site decisions of 4,400,000 and 380,000 campers, respectively, makes the simultaneous estimation complex. To solve this problem, I used a sequential approach based on which the NL model gets divided and run into different stages. In this case, since I am working with the two-stage NL model, I split the model into two logit models. First, I estimate the MNL model for travelling to a campsite, which represents the site decision stage, the lower-level. Then, I processed the results at an intermediate level and employed the result for estimation at the upper-level, the participation decision stage.

The sequential approach has both advantages and disadvantages. It has been found that the sequential approach is not efficient, yet it is consistent (Brownstone and Small 1989). Moreover, in some cases, some parameters might be found both in the upper level and the lower level estimations (Train 2009c), which is not a drawback in this study, since the two levels are entirely separate. Although the sequential pattern is not as efficient as the simultaneous estimation, the noticeable amount of observation that incorporated in this study help to repay the efficiency.

Eventually, the expected utility from travelling to a campsite and staying at home in the sequential approach gets the following form.

$$V_{it} = \begin{cases} V_{camping} = \theta_{nt} * \log(\sum exp(ASC_k + b_{TC} * TC_{ikt} + b_{ADV} * ADV_{kt}) * AV_{kt}), \\ \qquad k = 1, \dots, 73 \\ \\ V_{home} = ASC_h + b_{longwkd} * longwkd_t + (TSC_{20} * W_{20} + \dots + TSC_{36} * W_{36}), \\ \qquad if\ k = 74 \rightarrow h \end{cases} \tag{4.23}$$

Where the first stage of the sequential approach that reflects the site decision stage, included in the obtained expected utility from travelling to a campsite. For better clarification, equation 4.24 represents the first stage.

After removing the duplicate reservations and the observations without travel cost, I estimate the lower level of the model for more than 380,000 trips. Since working with a large dataset and estimating the NL model simultaneously was too complicated and took a long time, I used a nice trick and ran this model at two separate stages. In the NL sequential estimation, the evaluation starts from the last branch and use the results for the upper limbs.

In the first stage of the model, I estimate the following model to assess the effect of a posted advisory on the beach on the individuals' site decision (V_{ikt}). Equation 4.23 shows the expected utility from travelling to a campsite in consideration of beach advisory as a representative of lakes' water quality conditions.

$$V_{ikt} = (ASC_k + b_{TC} * TC_{ikt} + b_{ADV} * ADV_{kt}) * AV_{kt} \tag{4.24}$$

Where

- V_{ikt} is the systematic factor of the utility function that person i gains by travelling to a campsite k in choice occasion t;
- ASC_k captures all unobserved site characteristics for the campground k;
- b_{TC} is the coefficient for the travel cost that needs to be estimated;
- TC_{ikt} is a travel cost for person i incur by travelling to the campground k in choice occasion t;
- b_{ADV} is the coefficient for the presence of beach advisory that needs to be estimated;
- ADV_{kt} is a dummy variable that represents the presence of beach advisory for the campground k in choice occasion t;
- AV_{kt} is a dummy variable, denotes if a campsite was available on the campground k in choice occasion t.

As an intermediate level, the step through which I connected the two stages. I processed the results from the first stage to estimate the expected utility from camping by using the inclusive value. I assessed the utility that each individual obtained from travelling to all campsites in each choice occasion through the following formula.

$$IV_{it} = log\left(\sum (exp(ASC_k + b_{Tc} * TC_{ikt} + b_{ADV} * ADV_{kt}\right) * AV_{kt})) \tag{4.25}$$

Then, I started the second stage and estimated the participation decision by running a model for camping and staying at home that included more than 4,400,000 participation decisions. Equation 4.26 shows the expected utility from travelling to a campground.

$$V_{camping} = \theta_{nt} * \ IV_{it} \tag{4.26}$$

And the expected utility from staying at home is estimated by the following formula:

$$V_{home} = ASC_h + b_{longwkd} * longwkd_t + (TSC_{20} * W_{20} + \cdots + TSC_{36} * W_{36}) \tag{4.27}$$

Where $longwkd_t$ is a dummy variable assigned to the presence of long weekends and W_{20},…,W_{36} are the week numbers in each year. In equation 4.17, α coefficient captures the variables in the upper nest, including ASC for staying at home, long weekend, and TSCs.

4.2.5 **Welfare Measures**

So far, I have discussed how to construct the model. In this subsection, I will monetarily quantify the impacts of water quality improvement in terms of marginal and total welfare. The Marginal Willingness to Pay (MWTP) to avoid beach advisory with a beach advisory is possible by dividing the estimated coefficient for the beach advisory from the first stage by the estimated coefficient for the travel cost. Equation 4.28 represents the MWTP, where b_{TC} implies the negative of the marginal utility of income. MWTP shows how much money ($0.01) campers are willing to pay to improve the lakes' water quality and avoid beach advisory (in this case, to avoid beach advisory because it is a dummy variable).

$$MWTP = -\frac{b_{ADV}}{b_{TC}} \tag{4.28}$$

In addition to the marginal willingness to pay, I used a total welfare measure based on a scenario of removing all beach advisories. In this regard, I used Compensation Variation (CV) notion. CV calculates the difference in the expected utility after and before a change in the quality. Therefore, the CV shows the amount of money required to compensate the campers to bring them back to the initial utility level. In general, CV is a kind of payment that makes respondents indifferent between the status qua and after a change (Bockstael and McConnell 2007).

The CV will be negative if the water quality change makes the campers worse off, and they should be paid. The positive CV illustrates that campers are willing to give up an amount of money to obtain the second level of the utility. Equation 4.29 represents a simple form of CV when there is no uncertainty or random parameter.

$$V(P, Q^0, M) = \; V(P, Q^1, M + CV) \tag{4.29}$$

Where V is the expected utility, P is the price (in this study the travel cost), $\Delta Q = Q^0 - Q^1$and shows the water quality change and M is income.

The random component of the two-stage nested logit model estimation complexifies equation 4.29. Based on equations 4.15 and 4.17, I can write the expected value of the maximum utility function as below (Bockstael and McConnell 2007).

$$\tilde{V} = \ln\left(-\sum_{n=1}^{N} \left(\sum_{j_n=1}^{j_n} \exp\left(-\frac{V_{ij_nt}}{\theta_{nt}} \right) \right)^{\theta_{nt}} \right) + \bar{C} \tag{4.30}$$

Where $\bar{C}$ is an unrecoverable constant.

The CV measures the payment that should be made to make an individual indifferent between the original and after the change situation. Under this assessment, the individual's marginal utility of income (b_{TC}) would remain fixed and make the calculation easier since they would cancel out each other. Moreover, I had the same estimated coefficient for the beach advisory (b_{ADV}) for assessing the welfare change for each year. Since two stages are completely independent, and changes in the lower level do not drive behaviour at the upper level, I employed the MNL model of the RUM approach. The equation 4.31 represents the average per trip CV.

$$E(CV) = \left\{ \ln(\sum_{j \in B_k} \exp(V_{ijt}^1) - \ln(\sum_{j \in B_k} \exp(V_{ijt}^0) \right\} \Big/ -b_{Tc} \tag{4.31}$$

Where V_{ikt}^0 is the status quo (initial state) and V_{ijt}^1 is the level of utility under the given scenario (subsequent state). In this study, I studied a scenario in which the lakes' water quality had been improved in a way that there are no beach advisories, and $Q^0 < Q^1$.

Chapter 5

Results

The main findings of this study will be presented in this chapter to answer the research questions of whether and how lakes' water quality affects individuals' behaviour. Toward this end, over 380,000 trips from the revealed preference of more than 150,000 campers to 73 campsites between 2014 and 2018 were employed in the analysis. First, the results from the first stage of the NL model will be discussed to determine what water quality indices affect recreational decisions, followed by the challenges I had with the lake's water quality measurement. Then, the results were used to build up the second stage of the NL model to assess the participation decision, and the discussion will move to verify the NL model as a competent model for this study. Finally, I used these econometric results to evaluate the welfare measure, and I assessed the scenario of removing all beach advisories.

5.1 Site Choice Model (First stage)

The site decision was assessed in the first stage of the two-stage sequentially estimated NL model. In this level, the travel data to 73 campgrounds were modelled using the MNL model. As explained earlier, the disturbance terms within a nest are correlated, which violates the i.i.d assumption. Therefore, the estimated t- ratio and standard errors are not reliable anymore. For assessing the precision of the parameters, I used robust standard error since it estimates the correct standard error (White 1980; Jang et al. 2010). Table 5.1 shows the parameters estimated from the first stage of the NL model with a full set of ASCs, an interaction variable, travelling cost, and robust t-ratio.

The results of the first stage revealed negative coefficients for the presence of the beach advisory and travel cost. As table 5.1 represents the presence of a beach advisory provides a negative utility for the campers, and b_{ADV} is estimated precisely since the robust t–ratio is large. Regarding the estimated coefficient for the travel cost, which can also be interpreted as a negative of marginal utility of income, the result shows that an increase in income (decrease in travel cost) increases the obtained utility from travelling to a campsite. All parameters are significant at 1%.

72 ASCs were estimated at this stage, and ASC_1 was normalized to zero. Estimated ASCs with the negative signs represent that relative to the campground number 1 (the one with the ASC normalized to zero), the marginal utility is negatively affected by having a trip to that campsite where travel costs and lakes' water quality are held constant. And positive estimated ASCs represent the marginal utility is positively affected by having a trip to that campsite under the mentioned condition. The estimated coefficients of 71 ASCs are statistically significantly different from 0 at the 95% confidence level since the absolute values of robust t-ratios are greater than 1.96.

Table 5.1 Estimated Parameters – First Stage

Number of individuals	: 154653
Number of observations	: 381272
LL(0)	: -1605394
LL(final)	: -1209426
Rho-square (0)	: 0.2466
Adj.Rho-square (0)	: 0.2466

Coefficient	Estimate	Rob.t-ratio(0)	Coefficient	Estimate	Rob.t-ratio(0)
b_ADV	-0.136	-14.555	b_TC ($100)	-0.910	-257.864
Estimated ASCs					
ASC_1	0	[fixed]	ASC_2	-1.361	-36.597
ASC_3	-1.255	-37.538	ASC_4	-0.012	-0.465
ASC_5	-1.343	-27.035	ASC_6	0.103	4.149
ASC_7	-1.242	-24.292	ASC_8	-1.363	-45.923
ASC_9	-0.748	-32.246	ASC_10	-0.369	-18.232
ASC_11	-1.446	-49.199	ASC_12	-0.749	-32.827
ASC_13	-1.112	-48.622	ASC_14	-1.098	-47.119
ASC_15	0.140	5.547	ASC_16	-1.285	-53.479
ASC_17	-1.448	-43.773	ASC_18	-2.498	-51.904
ASC_19	-0.562	-22.871	ASC_20	-0.963	-35.998

ASC_21	-0.760	-25.191	ASC_22	-1.830	-43.208
ASC_23	-0.440	-15.860	ASC_24	-2.845	-41.971
ASC_25	-1.778	-30.617	ASC_26	-1.439	-44.077
ASC_27	-2.263	-60.474	ASC_28	-0.788	-34.558
ASC_29	-0.227	-9.801	ASC_30	-2.148	-71.041
ASC_31	-0.389	-13.581	ASC_32	-0.568	-24.586
ASC_33	-1.870	-40.957	ASC_34	-0.630	-25.690
ASC_35	-2.002	-56.902	ASC_36	-1.423	-43.931
ASC_37	-1.768	-54.541	ASC_38	-2.157	-34.958
ASC_39	-0.363	-16.042	ASC_40	-0.849	-36.291
ASC_41	-2.784	-84.121	ASC_42	-0.540	-24.863
ASC_43	-2.543	-77.720	ASC_44	-1.278	-47.905
ASC_45	-0.850	-36.244	ASC_46	-1.804	-71.379
ASC_47	-1.457	-32.652	ASC_48	-1.112	-35.609
ASC_49	-0.801	-28.674	ASC_50	-2.514	-42.839
ASC_51	-3.581	-57.076	ASC_52	-1.550	-48.790
ASC_53	-0.971	-41.616	ASC_54	-3.256	-62.762
ASC_55	-2.327	-73.118	ASC_56	-2.012	-70.943
ASC_57	-2.153	-48.911	ASC_58	-1.242	-37.871
ASC_59	-2.553	-66.767	ASC_60	-2.977	-40.005
ASC_61	-1.537	-35.367	ASC_62	-3.044	-33.693
ASC_63	-1.492	-39.908	ASC_64	-0.323	-10.814
ASC_65	-2.591	-51.495	ASC_66	-0.651	-19.081
ASC_67	-2.110	-63.529	ASC_68	-2.181	-56.539
ASC_69	-2.022	-35.469	ASC_70	-3.205	-41.190
ASC_71	-3.555	-25.641	ASC_72	-1.491	-49.936
ASC_73	-1.735	-29.228			

5.2 Participation Choice Model (Second stage)

Before moving to the second stage of the NL model, I estimated the utility that each individual obtained from travelling to campsites in each choice occasion (equation 4.25). The study of over 4,400,000 participation decisions helped us determine inclusive value, which represents the expected utility of going camping.

Table 5.2 shows the parameters estimated from the second stage of the NL model with a full set of TSCs, ASC for staying at home, a dummy variable for the long weekends, and θ_{nt}. This table reflects the participation choice level results. Since the standard errors of the estimated IV_{it} from the lower level (B_k), is not incorporated in the second stage calculation; the standard errors of the upper-level parameters are downward biased (Amemiya 1978). The standard errors in the second stage are biased downwards because of uncertainty over the estimated IV_{it} is not accounted for. Though the severity of this issue highly depends on the accuracy of the lower-level, generally, this issue gets critical when the model goes beyond three stages (Ben-Akiva, Lerman, and Lerman 1985).

From the results of the second stage, represented in table 5.2, the estimated θ_{nt} is close to zero and statistically different from 1, which would represent the MNL model. If θ_{nt} is between one and zero, it will be consistent with utility maximization (Train 2009c). The NL model is a generalized format of the CL model, where a correlation between unobserved utility is possible (Train 2009c). Based on equation 4.18 and the estimated θ_{nt} in table 5.1 the NL model is the best fit for this study relative to the CL model. Therefore, $1 - \theta_{nt} \cong 1$ and shows that IIA assumption holds within the nest. In other words, the two stages of the model are completely independent and degrading in lakes' water quality will not stop campers from camping. In real-world θ_{nt} might be different from person to person but in this model I consider θ_{nt} as a fixed parameter (Train 2009c).

ASC for staying at home is positive because people choose to stay home in more than 90% of the choice occasions. The estimated coefficient for having a long weekend with the negative sign represents that relative to travelling to a campsite, the marginal utility is negatively affected by staying at home where other time-variant variables are held constant. Therefore, people prefer camping on long-weekends relative to other weekends. Same as the first stage, all parameters are significant at 1%.

Table 5.2 Estimated Parameters – Second Stage

Number of individuals	: 154653
Number of observations	: 4462158
LL(0)	: -3092932
LL(final)	: -1290466
Rho-square (0)	: 0.5828
Adj.Rho-square (0)	: 0.5828

Coefficient	Estimate	Rob.t-ratio(0)	Description
θ_{nt}	$1.261e^{-5}$	2.64	Inclusive Value
ASC_Home	3.017	321.470	
B_LONGWKD	-0.261	-40.193	
Estimated TSCs			
TSC_20	0	[fixed]	The twenty-first week of the year is considered as the first week of the open water season in this study, and the seventeenth week or the last week of the open water season is the thirty-sixth week in a year. TSC_20 was normalized to zero.
TSC_21	-0.208	-18.315	
TSC_22	-0.097	-7.516	
TSC_23	-0.461	-38.296	
TSC_24	-0.383	-31.292	
TSC_25	-0.585	-49.044	
TSC_26	-0.717	-63.352	
TSC_27	-0.684	-58.484	
TSC_28	-0.831	-71.360	
TSC_29	-0.857	-73.586	
TSC_30	-0.866	-74.798	
TSC_31	-0.684	-57.261	
TSC_32	-0.799	-69.418	
TSC_33	-0.852	-73.343	
TSC_34	-0.712	-60.769	
TSC_35	-0.483	-41.389	
TSC_36	-0.153	-12.164	

The weekly time fixed effects, TSCs, were assigned to each week that was captured in this stage. Same as the first stage, one TSC was normalized to zero, and 16 TSCs that included in the stay at home equation were estimated. The negative coefficient, along with the robust large t-ratio, implies a precise calculation and shows that people are less likely to stay at home. For a better illustration, figure 5.1 represents the TSC trend over the weeks. A negative TSC shows that relative to other choice occasions, the marginal utility is negatively affected by staying at home. Therefore, based on figure 5.1, the absolute value of TSCs increases as the weather conditions become more suitable for camping. This graph represents that during week number between 28 and 33, the absolute value of the TSCs had higher records. These weeks reflects the late June and July when the weather condition is most suitable for camping and when people have holidays.

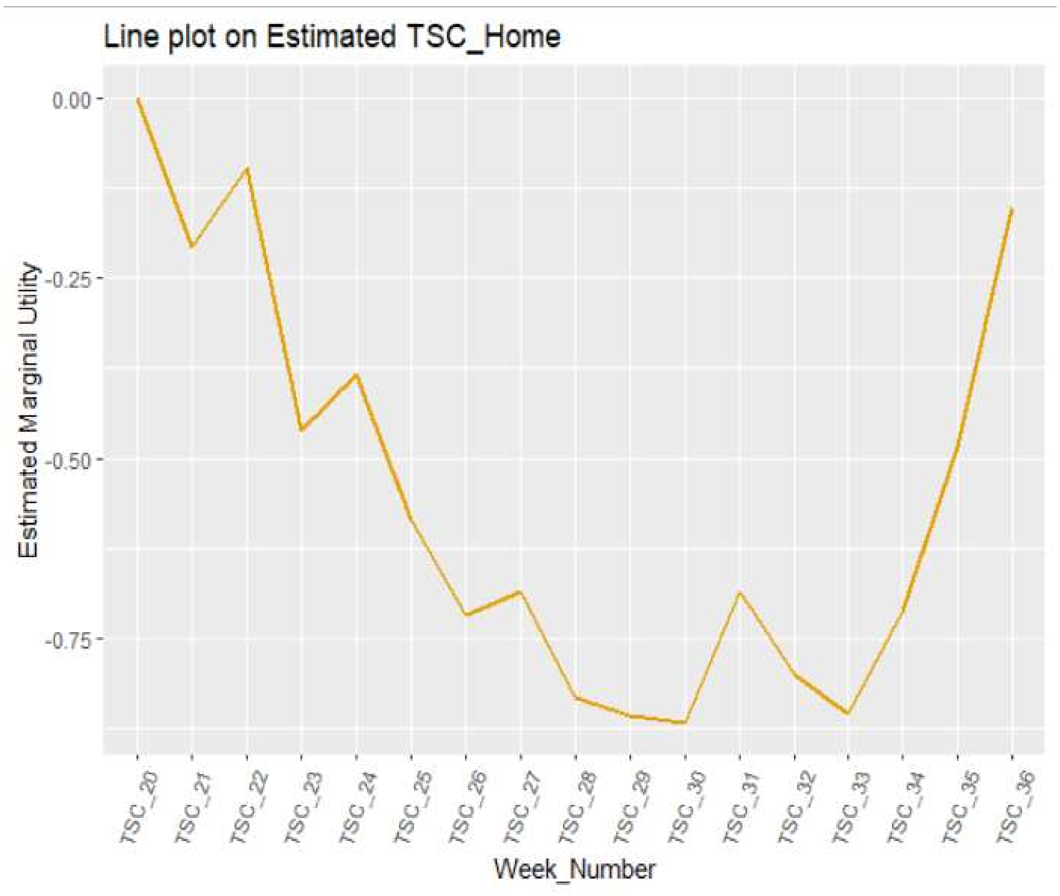

Figure 5.1 TSCs Trend Over the Weeks

5.3 Welfare Measures

This section discusses how the NL model is used to predict behavioural outcomes and measure the economic welfare impacts on the lakes' water quality. The parameters of the indirect utility function are used to calculate the welfare measures. The results will help us to understand how people are affected by the presence of a beach advisory as lakes' water quality indicium. In the following subsections, I will have a result of the marginal willingness to pay and the total welfare measure.

5.3.1 Marginal Welfare

I use the parameters estimated in the first stage of the NL model presented in table 5.1, and calculate a marginal willingness to pay (MWTP), based on formula 4.28 in section 4. Individuals are willing to pay $15 more per trip to travel to campsites without an advisory, while the average travel cost per trip is $239. In other words, on average, campers are willing to incur 6.3% more expenditure per trip to travel to a campsite without beach advisory. Table 5.3 represents the result of the MWTP per trip.

Table 5.3 MWTP to Avoid Beach Advisory

Coefficient	Estimate	Rob.std.err.	Rob.t-ratio(0)
MWTP	15	1.03	14.53

A study conducted by Murray et al. (2001) has demonstrated that each visitor is willing to pay 28 USD to remove beach advisory across the Great Lakes. The authors employed the SP of 800 on-site visitors while in this study, I used the RP of a substantial observation (Murray et al. 2001). Moreover, in the SP method, the questionnaire design plays a significant role in the reliability of the responses, as discussed in chapter 2. The previous studies were conducted in the United States on HAB and water clarity (Zhang and Sohngen 2018; Keeler et al. 2015). Zhang and Sohngen (2018) showed that recreational anglers are willing to pay 8 to 10 USD more per trip for one less mile of boating through HABs. Keeler et al. (2015) found that travellers are willing to pay 22 USD for every 1-meter increase in water clarity.

All in all, the results from the present study are in line with the findings of a great deal of the previous works in the recreation demand area. Data collection methods, different water quality measures, having information about recreation activities, as well as how to incorporate water quality measures, play a significate role in the results of each study. However, most of the published studies are consistent with the positive WTP for improving water quality.

5.3.2 Total Welfare

To predict an individual's behaviour and welfare in response to water quality changes in the beach advisories, I considered a scenario in which the lakes' water quality is improved in a way that all beach advisories are removed each year. So the baseline is different each year. Then, I simulated the scenario and completed the average per trip welfare analysis for each year using equation 4.31.

The inclusive value coefficient (θ_{nt}) that was estimated in the second stage of the NL model represents that the two stages are completely independent, and a correlation does not exist across the nests. Therefore, any changes in water quality will not influence campers' decisions at the participation level. Thus, to simulate behavioural responses and welfare impacts, I used the actual trip data that was employed in the site decision stage.

Table 5.4 reports the average per trip welfare changes from improving the lake's water quality in a way that there is no beach advisory at a site for each year. The policy of improving water quality to remove all beach advisories do not affect the sites that did not have advisory at all. To estimate the welfare change in response to this policy, the trips to the campgrounds that had beach advisory in the status quo. Therefore, affected trips show the number of trips to the campsites that had beach advisories, and due to the imposed water policy, the beach advisories are removed. Some campgrounds did not experience advisories at all.

Table 5.4 Proposed scenario to Calculate Compensation Variation for Each Year

Variable	2014	2015	2016	2017	2018
Average Welfare Measure (per trip)	$81,307	$59,490	$104,183	$53,513	$9,861
Number of Campgrounds with Beach Advisory	10	8	12	10	6
Affected Trips	4,740	4,022	6,543	3,507	627

Annual Welfare Change (per trips)	$17.15	$14.79	$15.92	$15.26	$15.73

The mentioned water treatment policy would improve the annual camper's welfare of provincial campgrounds. In table 5.4, it can be seen that by far, the greatest economic cost of beach advisories is around $104,000 for 2016. The discrepancy of economic costs of beach advisories could be attributed to the number of issued beach advisories and travel cost data limitation issue. To calculate the cost of travelling, I considered day trips and over-night trips. Generally, on-site time can be ignored for the travel cost calculation as it is endogenous, and travellers can substitute a one-day trip with a multiple-day trip (McConnell 1992). However, trip length has little impact on welfare measure (W. D. Shaw and Ozog 1999). It has been shown that on average, anglers' welfare change who travelled over-night is a bit higher than those who had day trips for increasing the catch rates(W. D. Shaw and Ozog 1999).

The travel cost was calculated based on the linked the households' annual income to postal codes 2016. Consequently, the welfare assessment is limited by the lack of information on visitors' income records for the years 2017 and 2018. Based on table 3.2, 6% of the choice occasions had beach advisories, which could be a reason why the estimated welfare is relatively low. It is important to bear in mind that this assessment is conducted based on the sample I had, not the whole population who travelled to the campgrounds. On average, the welfare change for each person would be almost two and a half times greater than the estimated values since, on average, the number of campers is 2.5.

Chapter 6
Conclusion

6.1 Summary

The goal of this research is to study whether lake water quality affects people's choice of recreation sites and to quantify their behavioural change in response to water quality improvement by using a recreation demand model. A secondary goal is to calculate welfare measures that can be used to understand the economic benefits of water quality improvements. Results can be employed by decision-makers in a benefit-cost analysis of water treatment programs to examine the economic efficiency associated with the program by analyzing the cost and the benefits. More broadly, this research aims to improve the understanding of how local environmental quality affects people's decisions.

This study studied over 70,000 campers' behaviour for five years and conducted across provincial parks of Alberta. The online reservation system allowed us to access campers' choices, including the campsite, number of campers, reservation length, and their postal codes. The model was built up by linking the RP dataset with the beach advisory reports as a representative of lakes' water quality conditions. Due to the substantial volume of observations, running the two-stage NL model at once was not feasible. Therefore, I used the sequential approach and estimated the model in two pieces and joined them via an intermediate step.

The major contribution of this work is associated with studying a noticeable number of revealed preference dataset that was covered in chapter 2. The main finding of this non-market valuation was that the beach advisories affect recreationists negatively. Individuals are willing to pay $15 more per trip to travel to a campsite with no advisory in place. The research has also shown the usefulness of non-survey data to understand recreation behaviour.

One of the important findings of this study is understanding that campers will not easily substitute their recreation activities, travelling to a site, with any other sort of activities that are captured by staying at home option in this study. The most likely cause of independency between site decision and participation decision is that most travellers had only one trip per year. Various recreation sites and numerous lakes across Alberta have provided a rich choice set of alternatives for the travellers.

6.2 Policy Implication

The results of this study can be used by policymakers to measure the benefits of water quality improvements in the recreational demand area. The monetary values in welfare changes and an individual's willingness to pay would be helpful for policymakers in designing the lakes' water quality management program in Alberta. To assess the economic efficiency of a water quality improvement program, policymakers can examine the benefits and the cost associated with the program and compare the results with the status quo. The result of this analysis depends on the priorities and goals of the program. The welfare measures can be used to compare to the costs of the preferred program to perform a cost-benefit analysis.

For the welfare result, I estimated the CV for a water quality management policy that would remove all 8 to 12 beach advisories each year across 73 provincial campsites. I used the individuals' actual behaviour indirectly to assess the economic value of water-quality changes. The estimated economic value of removing beach advisories seems to be small for a few reasons. The first reason only people who made reservations in advance through RAP were included in this study, not the day users or who used the first come first serve option. For instance, Pigeon Lake, a popular recreational destination, is close to Edmonton, and many people drive there for a day visit. Moreover, people who own a cabin on the lake were not considered.

The second reason is that only the use-value, a part of total economic value, has been captured by the RP method. Meanwhile, travellers are enjoying and using the campgrounds' benefit with higher water-quality; some people out there are willing to pay for better water quality to use it in the future or even preserve the natural resources for the next generations. The dataset in this study did not incorporate the latter group. Consequently, the second group's economic value, called non-use value, is not reflected in the result. It is unfortunate that the study did not include any SP data to reflect both groups' preferences regarding the bequest value and existence value (Abdullah et al.

2011). Hence, the welfare measurements conducted in this study only depicted a private economic cost of the public good.

Maintaining and improving the water quality of all campgrounds could be costly. The entrance fee and the reservations fee are the two sources that could yield revenue to cover a part of maintenance or treatment costs. A further study could assess the welfare impact of changing these fees for each campground or increasing all prices by a constant percentage based on the applied behavioural model. Moreover, campers in Alberta would not replace travelling to provincial campsites with any other activity because the estimated coefficient for inclusive value (θ_{nt}), is almost zero and represents that the two nests are completely independent. Therefore, recreation activities are important for people and improving or maintaining suitable water conditions at recreation sites would be vital.

Alberta's population and economic growth and the global warming trend have raised the importance of water treatment policies. Therefore, beach advisories' economic costs, as a water-quality indicator, could help impose an appropriate approach. The agriculture sector is one of the water pollutant sources. Agricultural nutrient and pesticide run-offs will reach lakes and rivers and contaminate the water. Therefore, another important practical implication is controlling and reducing fertilizer and pesticide run-offs from the agriculture sector. I would suggest using some incentive policies and offer financial bonuses to encourage farmers to manage fertilizer and pesticide application and consequently mitigate the nutrient run-offs. Furthermore, allocating wetlands would help reduce the use of fertilizers and, as a result, lower agricultural run-off. In this case, the benefit-cost analysis of wetlands allocation is required. The economic cost of the beach advisories that are estimated in this study could be a financial source or budget of these diminishing run-offs programs. The money collected from people who are willing to pay to avoid beach advisories per trip, which will change travellers' welfare, represents the private part of this public good.

The Alberta government planned to fully or partially close some parks to optimize Alberta's parks("News & Events | Alberta Parks" n.d.). Parks have different economic values, and parks closure brings about different economic costs. The monetary values in welfare changes could be used beyond the water quality and recreation demand area. The applied approach could be extended to prioritize campground closure to lessen the economic cost associated with Alberta's park optimization program.

6.3 Limitations and Future Research

Some issues could be considered as further research that is specifically related to water quality and recreation demand. A limitation is not much individual-specific data on people. This research used the median household income at the postal code level, using the individual-specific information could help to assess the opportunity cost of time more precisely.

In addition, individual sociodemographic information could provide a better cognition of the sample, including their age group, gender, and recreational activities. This information would help us to classify campers into different groups based on their preferences and sociodemographic characteristics to see how different people have different WTP for water quality improvement. However, this study is unable to encompass the different recreational activities since the revealed preference dataset does not provide it, and I do not have any information about travellers' activities at the sites. Consequently, heterogeneity could be addressed to some extent and improve the current model. Validity is one of the most critical components of discrete choice models, which represents the reliability of the model. Since I was working on individuals' actual behaviour, validity might not be an issue; however, missing information about sociodemographic data and activity types exist.

The second limitation of this study is the endogeneity associated with the travellers' experience. The severity of beach advisories is not distinguishable because they are issuing based on several criteria and imposing the same limitation on water usage. Therefore, travellers' perception and experience would be different under the same situation of beach advisories' existence. Moreover, people might have a different experience of beach advisories based on their recreation activities. Due to the absence of such information and model limitation (i.e. assumed that the preferences of all individuals in the sample are identical), I did not consider the effect of past behaviour on the current behaviour.

The third limitation is the water quality dataset, including beach advisory and cyanobacteria cell count. This limitation arises from the difficulty of the data collection that caused a lot of missing values for most of the lakes. Moreover, predicting these missing values was not possible due to cyanobacteria's instability. Hence, incorporating this biological indicator was not applicable due to numerous missing values for the cyanobacteria as a noisy variable. More information on cyanobacteria cell count would help us to establish a greater degree of investigation on recreation demand and compare it with the beach advisories effect.

The study is limited by the lack of information on booking dates. Beach advisories might affect travellers' decisions while they are booking a campground for specific dates. However, this sort of information might not be posted on the Alberta Health Services when the reservation is processing some months earlier. Having complete information on the booking date and if the campers checked in the campground as they reserved would help assess beach advisory's effect on individuals' behaviour more precisely.

Although this study could reflect the importance of the lakes' water quality in recreation demand, further research may be performed to employ different water quality measurements, which may help better understand the campers' behaviour. Consequently, the result would be valuable for the policymakers since it gives them the idea of the most effective water quality measures to put it in the priority of recreation water acts.

There is a variety of ways to extend and build upon this study. Most of the pre-existing studies did not have access to the exact day of the trips, and therefore they considered monthly, seasonally, or a yearly number of trips in their studies. An aspect not examined in this study is the impact of considering the exact day of trips and the relative water condition, in comparison to considering the total number of trips. The results would play a significant role in assessing the accuracy of the current assessments.
Further research could expand on the study by including weather effects to assess the benefit-cost of climate change. Since climate change will likely degrade the lakes' water quality, it is vital to impose several scenarios to estimate the welfare change. A study assed the visitation behaviour of National parks under hypothetical climate conditions (Loomis and Richardson 2006). On the other hand, climate change would extend the summer length and associated recreation activities (Mendelsohn and Neumann 2004). Therefore, it is vital to employ this model and develop to calculate the benefit-cost of climate change and recreation demand. The study could also be

extended in imposing and examining different scenarios, like the extent to which welfare change across campsites in Alberta Province and prioritize the most beneficial ones.

REFERENCES

Abdullah, Sabah, Anil Markandya, and Paulo A.L.D. Nunes. 2011. "Introduction to Economic Valuation Methods." In *Research Tools in Natural Resource and Environmental Economics*, by Amitrajeet A Batabyal and Peter Nijkamp, 5:143–87. WORLD SCIENTIFIC. https://doi.org/10.1142/9789814289238_0005.

Adamowicz, Wiktor, Joffre Swait, Peter Boxall, Jordan Louviere, and Michael Williams. 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. https://doi.org/10.1006/jeem.1996.0957.

Alberta, and Alberta Health. 2014. *Alberta Cyanobacteria Beach Monitoring*. http://www.health.alberta.ca/documents/Cyanobacteria-Algae-Monitoring-2014.pdf.

Alvarez, Sergio, Sherry L. Larkin, John C. Whitehead, and Tim Haab. 2014. "A Revealed Preference Approach to Valuing Non-Market Recreational Fishing Losses from the Deepwater Horizon Oil Spill." *Journal of Environmental Management* 145 (December): 199–209. https://doi.org/10.1016/j.jenvman.2014.06.031.

Amemiya, Takeshi. 1978. "On a Two-Step Estimation of a Multivariate Logit Model." *Journal of Econometrics* 8 (1): 13–21. https://doi.org/10.1016/0304-4076(78)90086-6.

Ben-Akiva, Moshe E., Steven R. Lerman, and Steven R. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press.

"Blue-Green Algae | AEP - Environment and Parks." n.d. My Wild Alberta. Accessed April 17, 2020. https://mywildalberta.ca/fishing/advisories-corrections-closures/blue-green-algae.aspx.

Bockstael, Nancy E., W. Michael Hanemann, and Catherine L. Kling. 1987. "Estimating the Value of Water Quality Improvements in a Recreational Demand Framework." *Water Resources Research* 23 (5): 951–60. https://doi.org/10.1029/WR023i005p00951.

Bockstael, Nancy E., and Kenneth E. McConnell. 2007. *Environmental and Resource Valuation with Revealed Preferences*. Vol. 7. The Economics of Non-Market Goods and Resources. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-1-4020-5318-4.

Boxall, Peter C, and Wiktor L Adamowicz. 2002. "Understanding Heterogeneous Preferences in Random Utility Models: A Latent Class Approach." *Environmental and Resource Economics*, 26.

Brownstone, David, and Kenneth A. Small. 1989. "Efficient Estimation of Nested Logit Models." *Journal of Business & Economic Statistics* 7 (1): 67–74. https://doi.org/10.2307/1391839.

"CAA National." n.d. Accessed May 28, 2020. https://www.caa.ca/.

"Cyanobacterial ('Blue-Green Algae') Blooms and Toxicity." 2015. May 15, 2015. https://open.alberta.ca/dataset/9780778576068/resource/f1c0c6b8-bfe9-4d04-be28-ad94cf74b3dd.

Dewar, Heather, François Soulard, Robby Bemrose, Alison Clark Milito, Cindy DeCuypere, Gordon Dewis, Mark Henry, et al. 2010. "Human Activity and the Environment," 61.

Domencich, T. A., and D. McFadden. 1975. "URBAN TRAVEL DEMAND - A BEHAVIORAL ANALYSIS." https://trid.trb.org/view/48594.

Egan, Kevin J., Joseph A. Herriges, Catherine L. Kling, and John A. Downing. 2009. "Valuing Water Quality as a Function of Water Quality Measures." *American Journal of*

Agricultural Economics 91 (1): 106–23. https://doi.org/10.1111/j.1467-8276.2008.01182.x.

Englin, Jeffrey, and J. S. Shonkwiler. 1995. "Estimating Social Welfare Using Count Data Models: An Application to Long-Run Recreation Demand Under Conditions of Endogenous Stratification and Truncation." *The Review of Economics and Statistics* 77 (1): 104–12. https://doi.org/10.2307/2109996.

English, Eric, Roger H. von Haefen, Joseph Herriges, Christopher Leggett, Frank Lupi, Kenneth McConnell, Michael Welsh, Adam Domanski, and Norman Meade. 2018. "Estimating the Value of Lost Recreation Days from the Deepwater Horizon Oil Spill." *Journal of Environmental Economics and Management* 91 (September): 26–45. https://doi.org/10.1016/j.jeem.2018.06.010.

"Facts About Water in Alberta." 2010, December, 68.

Feather, Peter, Daniel Hellerstein, and Theodore Tomasi. 1995. "A Discrete-Count Model of Recreational Demand." *Journal of Environmental Economics and Management* 29 (2): 214–27. https://doi.org/10.1006/jeem.1995.1042.

Fezzi, Carlo, Ian J. Bateman, and Silvia Ferrini. 2014. "Using Revealed Preferences to Estimate the Value of Travel Time to Recreation Sites." *Journal of Environmental Economics and Management* 67 (1): 58–70. https://doi.org/10.1016/j.jeem.2013.10.003.

Government of Canada, Statistics Canada. 2018. "Monthly Average Retail Prices for Gasoline and Fuel Oil, by Geography." May 18, 2018. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1810000101.

Haab, Timothy C., and Kenneth E. McConnell. 2002. *Valuing Environmental and Natural Resources: The Econometrics of Non-Market Valuation*. Edward Elgar Publishing.

Haefen, Roger H., and Daniel J. Phaneuf. 2005. "Kuhn-Tucker Demand System Approaches to Non-Market Valuation." In *Applications of Simulation Methods in Environmental and Resource Economics*, edited by Riccardo Scarpa and Anna Alberini, 135–57. The Economics of Non-Market Goods and Resources. Dordrecht: Springer Netherlands. https://doi.org/10.1007/1-4020-3684-1_8.

Haefen, Roger H. von, and Daniel J. Phaneuf. 2003. "Estimating Preferences for Outdoor Recreation: A Comparison of Continuous and Count Data Demand System Framework." *Journal of Environmental Economics and Management* 45 (3): 612–30. https://doi.org/10.1016/S0095-0696(02)00024-4.

Hang, Derrick, Daniel McFadden, Kenneth Train, and Ken Wise. 2016. "Is Vehicle Depreciation a Component of Marginal Travel Cost?" *Journal of Transport Economics and Policy* 50 (May): 19.

Herriges, Joseph A., and Daniel J. Phaneuf. 2002. "Inducing Patterns of Correlation and Substitution in Repeated Logit Models of Recreation Demand." *American Journal of Agricultural Economics* 84 (4): 1076–90. https://doi.org/10.1111/1467-8276.00055.

Jang, Kitae, Shin Hyoung Park, Sungbong Chung, and Ki Han Song. 2010. "Influential Factors on Level of Injury in Pedestrian Crashes: Applications of Ordered Probit Model with Robust Standard Errors." In , 19.

Jeon, Yongsik, and Joseph A. Herriges. 2010. "Convergent Validity of Contingent Behavior Responses in Models of Recreation Demand." *Environmental and Resource Economics* 45 (2): 223–50. https://doi.org/10.1007/s10640-009-9313-5.

Ji, Yongjie, and David Keiser. 2016. "Water-Based Recreation and Water Quality Indices: A Revealed Preference Approach." In . Vol. 37. Economics Presentations, Posters and

Proceeding. https://lib.dr.iastate.edu/econ_las_conf/37/.
Johnston, Robert J., and Paul J. Thomassin. 2010. "Willingness to Pay for Water Quality Improvements in the United States and Canada: Considering Possibilities for International Meta-Analysis and Benefit Transfer." *Agricultural and Resource Economics Review* 39 (1): 114–31. https://doi.org/10.1017/S1068280500001866.
Keeler, Bonnie L, Spencer A Wood, Stephen Polasky, Catherine Kling, Christopher T Filstrup, and John A Downing. 2015. "Recreational Demand for Clean Water: Evidence from Geotagged Photographs by Visitors to Lakes." *Frontiers in Ecology and the Environment* 13 (2): 76–81. https://doi.org/10.1890/140124.
Lloyd-Smith, Patrick, and Marcus Becker. 2019. "Seasonal Recreation Demand Modeling and Non-Market Impacts of Natural Disasters: An Application to Camping in Alberta, Canada," 41.
———. 2020. "The Economic Value of Camping Using Administrative Data." *Journal of Agricultural and Resource Economics* 45 (3): 445–461. https://doi.org/10.22004/AG.ECON.302498.
Loomis, John B., and Robert B. Richardson. 2006. "An External Validity Test of Intended Behavior: Comparing Revealed Preference and Intended Visitation in Response to Climate Change." *Journal of Environmental Planning and Management* 49 (4): 621–30. https://doi.org/10.1080/09640560600747562.
McConnell, K. E. 1992. "On-Site Time in the Demand for Recreation." *American Journal of Agricultural Economics* 74 (4): 918–25. https://doi.org/10.2307/1243189.
Mendelsohn, Robert, and James E. Neumann. 2004. *The Impact of Climate Change on the United States Economy*. Cambridge University Press.
Mooij, Wolf M., Stephan Hülsmann, Lisette N. De Senerpont Domis, Bart A. Nolet, Paul L. E. Bodelier, Paul C. M. Boers, L. Miguel Dionisio Pires, et al. 2005. "The Impact of Climate Change on Lakes in the Netherlands: A Review." *Aquatic Ecology* 39 (4): 381–400. https://doi.org/10.1007/s10452-005-9008-0.
Morey, Edward R., Robert D. Rowe, and Michael Watson. 1993. "A Repeated Nested-Logit Model of Atlantic Salmon Fishing." *American Journal of Agricultural Economics* 75 (3): 578–92. https://doi.org/10.2307/1243565.
Murdock, Jennifer. 2006. "Handling Unobserved Site Characteristics in Random Utility Models of Recreation Demand." *Journal of Environmental Economics and Management* 51 (1): 1–25. https://doi.org/10.1016/j.jeem.2005.04.003.
Murray, Chris, Brent Sohngen, and Linwood Pendleton. 2001. "Valuing Water Quality Advisories and Beach Amenities in the Great Lakes." *Water Resources Research* 37 (10): 2583–90. https://doi.org/10.1029/2001WR000409.
Natural Resources Canada. 2014. "Water." December 31, 2014. https://www.nrcan.gc.ca/earth-sciences/geography/atlas-canada/selected-thematic-maps/16888.
"News & Events | Alberta Parks." n.d. Alberta Parks. Accessed October 14, 2020. https://www.albertaparks.ca/albertaparksca/news-events/.
Parsons, George R. 2014. "Travel Cost Models." In *A Primer on Nonmarket Valuation*, by Patricia A Champ, Thomas C Brown, and Kevin J Boyle. Vol. 6. Economics of Non-Market Goods and Resources 13. http://web.b.ebscohost.com.cyber.usask.ca/ehost/ebookviewer/ebook/bmxlYmtfXzEzNTM4ODNfX0FO0?sid=1f47d1ad-2057-4ccc-8b95-f3fd9b87c1de@pdc-v-sessmgr04&vid=0&format=EB&rid=1.

Parsons, George R., Erik C. Helm, and Tim Bondelid. 2003. “Measuring the Economic Benefits of Water Quality Improvements to Recreational Users in Six Northeastern States: An Application of the Random Utility Maximization Model,” 26.
“Recreational Water Monitoring for Beach Operators.” 2019. *Alberta Health Services*, 3.
Russell, Clifford S, and William J Vaughan. 1982. “The National Recreational Fishing Benefits of Water Pollution Control.” *Journal of Environmental Economics and Management* 9 (4): 328–54. https://doi.org/10.1016/0095-0696(82)90026-2.
Segerson, Kathleen. 2014. “Valuing Environmental Goods and Services: An Economic Perspective.” In *A Primer on Nonmarket Valuation*, by Patricia A Champ, Thomas C Brown, and Kevin J Boyle. Vol. 1. Economics of Non-Market Goods and Resources 13. http://web.b.ebscohost.com.cyber.usask.ca/ehost/ebookviewer/ebook/bmxlYmtfXzEzNTM4ODNfX0FO0?sid=738854f9-a2dc-46f4-80e6-d6b6a176baec@pdc-v-sessmgr02&vid=0&format=EB&rid=1.
Shaw, Daigee. 1988. “ON-SITE SAMPLES’ REGRESSION Problems of Non-Negative Integers, Truncation, and Endogenous Stratification.” *Journal of Econometrics* 37 (2): 211–23.
Shaw, W. Douglass, and Michael T. Ozog. 1999. “Modeling Overnight Recreation Trip Choice: Application of a Repeated Nested Multinomial Logit Model.” *Environmental and Resource Economics* 13 (4): 397–414. https://doi.org/10.1023/A:1008218803875.
“Statistics Canada: Canada’s National Statistical Agency.” 2016. 2016. https://www.statcan.gc.ca/eng/start.
The Atlas of Canada - Lakes. 1973. “The Atlas of Canada - Lakes.” 1973. https://web.archive.org/web/20070410230512/http:/atlas.nrcan.gc.ca/site/english/learningresources/facts/lakes.html.
The Conference Board of Canada- Water quality index. 2018. “The Conference Board of Canada- Water Quality Index.” The Conference Board of Canada. December 2018. https://www.conferenceboard.ca/hcp/details/environment/water-quality-index.aspx.
Train, Kenneth. 2009a. “Discrete Choice Methods with Simulation.” In *Discrete Choice Methods with Simulation*, 2nd ed. Vol. 3. Cambridge University Press. https://eml.berkeley.edu/books/choice2.html.
———. 2009b. “Discrete Choice Methods with Simulation.” In *Discrete Choice Methods with Simulation*, 2nd ed. Vol. 2. Cambridge University Press. https://eml.berkeley.edu/books/choice2.html.
———. 2009c. “Discrete Choice Methods with Simulation.” In *Discrete Choice Methods with Simulation*, 2nd ed. Vol. 4. Cambridge University Press. https://eml.berkeley.edu/books/choice2.html.
US EPA, OW. 2015. “2000 National Water Quality Inventory Report to Congress.” Data and Tools. US EPA. September 21, 2015. https://www.epa.gov/waterdata/2000-national-water-quality-inventory-report-congress.
Van Houtven, George, Carol Mansfield, Daniel J. Phaneuf, Roger von Haefen, Bryan Milstead, Melissa A. Kenney, and Kenneth H. Reckhow. 2014. “Combining Expert Elicitation and Stated Preference Methods to Value Ecosystem Services from Improved Lake Water Quality.” *Ecological Economics* 99 (March): 40–52. https://doi.org/10.1016/j.ecolecon.2013.12.018.
White, Halbert. 1980. “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” *Econometrica* 48 (4): 817–38. https://doi.org/10.2307/1912934.

Wolf, David, Wei Chen, Sathya Gopalakrishnan, Timothy Haab, and H. Allen Klaiber. 2019. "The Impacts of Harmful Algal Blooms and E. Coli on Recreational Behavior in Lake Erie." *Land Economics* 95 (4): 455–72.

Wolf, David, Will Georgic, and H. Allen Klaiber. 2017. "Reeling in the Damages: Harmful Algal Blooms' Impact on Lake Erie's Recreational Fishing Industry." *Journal of Environmental Management* 199 (September): 148–57. https://doi.org/10.1016/j.jenvman.2017.05.031.

World Health Organization, ed. 2003. *Guidelines for Safe Recreational Water Environments*. Vol. Coastal and fresh waters. 1 vols. Geneva: World Health Organization.

Zhang, Wendong, and Brent Sohngen. 2018. "Do U.S. Anglers Care about Harmful Algal Blooms? A Discrete Choice Experiment of Lake Erie Recreational Anglers." *American Journal of Agricultural Economics* 100 (3): 868–88. https://doi.org/10.1093/ajae/aay006.

Zimmer, Natalie M. P., Wiktor L. Adamowicz, and Peter C. Boxall. 2012. "The Impacts of Chronic Wasting Disease and Its Management on Recreational Hunters." *Canadian Journal of Agricultural Economics/Revue Canadienne d'agroeconomie* 60 (1): 71–92. https://doi.org/10.1111/j.1744-7976.2011.01232.x.

APPENDIX

	Which variables employed?	How variables used?	What did the results represent?
Recreationists RP in Lake Erie (Wolf et al. 2019)	• E. Coli • Harmful algal bloom (HAB) • Site attributes • Traveller's expenditures	• Maximum E. coli readings from the closest monitoring stations • A 10-day composite of HAB (Microcystins; July 1- October 30) • Summer-long mean from the nearest remote sensing locations.	Beachgoers and fishers would lose in aggregate 5.3 USD million and 59.2 USD million respectively each year if the lake were closed. 40% reduction in phosphorus leads to welfare gains. Beachgoers are more averse to E. coli. Fishers are more averse to algae.
SP in Lake Erie's shoreline (Murray, Sohngen, and Pendleton 2001)	• Annual number of trips • How information about beach advisories obtained • Sociodemographic and trip log information • Site attributes • E. Coli • Beach Advisory	• On-site Sample • Single-day trips • The average number of beach advisories over 3 years • Focused on the total annual number of beach advisories • Average E. Coli per season • Used Poisson Model	Beach advisory is used by 70% of the visitors. The seasonal value of removing one advisory is 28 USD per person. The welfare change associated with a one-unit reduction in beach advisories per year at all beaches is 1.85 USD per person per trip; 2.12 USD and 1.73 USD per trip, and they would gain 34 USD and 28 USD per season for people use signs and media, respectively.
Recreationists RP in Iowa Lakes (Ji and Keiser 2016)	• Secchi depth and Turbidity • Dissolved oxygen saturation • Phosphorus and Total nitrate • pH • Total solids • Sociodemographic characteristics • Site attributes	• Employed a water quality index for turbidity, pH, dissolved oxygen saturation, phosphorus, nitrate, and total solids • Used WQ as a site-specific attribute • Repeated Mixed Logit with and without ASC	With ASC: coefficients of turbidity, total solids and pH are significant in a few cases. Without ASC: positive and statistically significant correlation exists between water quality indices and the attractiveness of lake sites. There is no consensus on which water quality affects recreation demand and in which form they should be modelled.
Anglers' RP in Minnesota (Feather, Hellerstein, and Tomasi 1995)	• Lake area • Lake depth • Littoral area • average Secchi depth for open water season (Jun24-Sep11)	• Log (area) to capture the diminishing effect of lake size • Reported littoral in the percentage of lake size to avoid correlation • Log (Secchi) to account for the non-linear relationship between water clarity and trophic status	Welfare increases as a lake clarity improve. Participation and welfare will increase if water quality improves. Findings show different magnitudes for warm water and cold water regions.

	Which variables employed?	How variables used?	What did the results represent?
Fishing RP of 8 Ohio counties (Wolf, Georgic, and Klaiber 2017)	• Algae measures • Monthly mean surface water temperature • Monthly precipitation • Monthly fishing permit sales within a given zip code • Distance to Lake and ramp	• Advisory threshold • Water Temperature • Heavy Rain Heavy • Heavy Rain Squared • Dummy variables	• Sales drop by %10-13 if algae surpass a threshold of 20000 cells/ml. • 2.25 USD million to 5.58 USD million lost in fishing expenditures. • A discrete jump in reduced angling activity upon crossing 20,000 cells/ml threshold. • Heterogeneity exists in WQ across and during the years.
SP of recreational anglers in Ohio Lake Erie (Zhang and Sohngen 2018)	• Expected walleye catch rate • Algal bloom • Water clarity • Driving distance from angler's house to their preferred boat ramps • Boating time	• Variables were used in different models (conditional logit with and without interaction, two G-MNL, and latent class model) • Water quality indicated by the size of algal bloom the fishermen have to boat through	• A significant and substantial willingness to pay by anglers for a reduction in HABs. • Anglers are willing to pay 8 USD to 10 USD more per trip for one less mile of boating through HABs. • Anglers are willing to pay, on average, 40 USD to 60 USD per trip for a policy that cuts upstream phosphorus loadings by 40%.
Recreationists RP in Minnesota and Iowa (Keeler et al. 2015)	• Geotagged photographs • Water clarity (m) • Lake size(acres) • Geotagged photos and its latitude longitude	• The average annual number of photo-user-days per Lake • 30-m buffer zone around each lake for photographs taken along the shoreline • Used profile information for locations • Dummy variables	• Lake size, water clarity, near-lake population, presence of a boat ramp, and state were significant predictors of annual average per-lake trip. • Lake users were willing to travel 56 minutes farther (=22 USD in travel costs) for every 1m increase in water clarity. • Larger lakes received more visits than smaller ones, and lakes with a boat ramp attracted more visitors than lakes without one.
Recreationists RP in Iowa Lakes (Egan et al. 2009)	• Sociodemographic characteristics • Secchi transparency • Chlorophyll • Nutrients • Suspended Solids • Cyanobacteria and phytoplankton • Site attributes	• Different water quality measurements were taken 3times/year • Used average water quality values • Only single-day trips considered in travel cost	• Individuals are responsive to the full set of water quality measures. • Place a higher value on improving a subset of the lakes to superior water quality rather than providing adequate levels at all of the lakes. • Water clarity (Secchi depth) and high concentrations of nutrients are the measures that significantly affect people's behaviour.

Table A.1 Summary of Related Studies

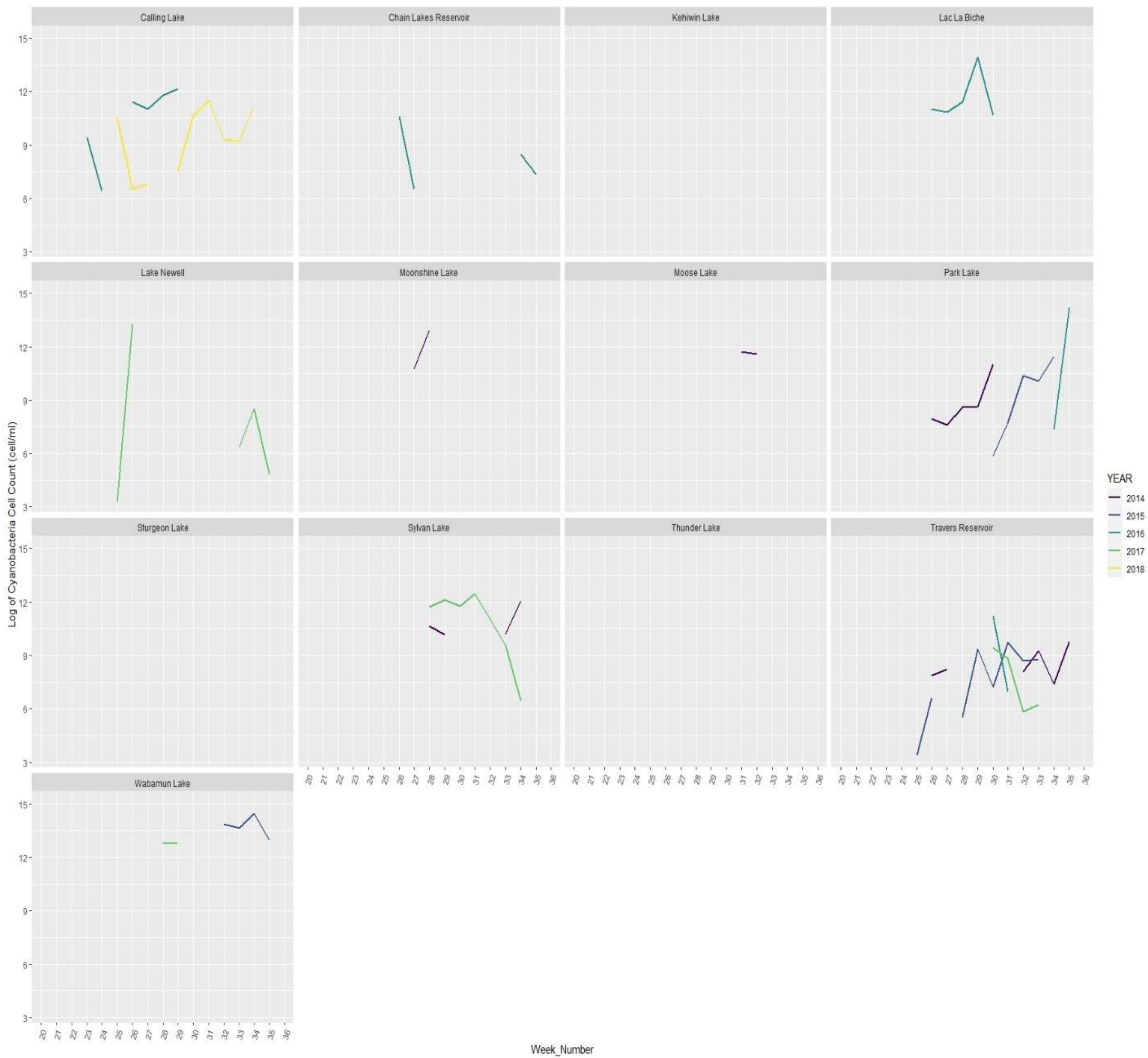

Figure A.1 Cyanobacteria Trend for the Lakes with Less than 9 Weekly Records in Alberta