UNIVERSITY OF WARSAW | FACULTY OF ECONOMIC SCIENCES

WORKING PAPERS
No. 5/2024 (441)

HOW STABLE AND PREDICTABLE ARE WELFARE ESTIMATES USING RECREATION DEMAND MODELS?

PATRICK LLOYD-SMITH
EWA ZAWOJSKA

WARSAW 2024

UNIVERSITY OF WARSAW
FACULTY OF ECONOMIC SCIENCES
WORKING PAPERS

How stable and predictable are welfare estimates using recreation demand models?

Patrick Lloyd-Smith[1], *Ewa Zawojska*[2]
[1] *University of Saskatchewan, Department of Agricultural and Resource Economics Saskatoon, Kanada;*
[2] *University of Warsaw, Faculty of Economic Sciences*
Corresponding author: ewa.zawojska@uw.edu.pl

Abstract: Economic analysis of environmental policy projects typically use pre-existing welfare estimates that are then transferred over time to the policy relevant periods. Understanding how stable and predictable these welfare estimates are over time is important for applying these estimates in policy. Yet, revealed preference models of recreation demand have received few temporal stability assessments compared to other non-market valuation methods. We use a large administrative dataset on campground reservations covering ten years to study temporal stability and predictability of recreation demand welfare estimates of lake water quality changes. Based on single-year models, our findings suggest welfare estimates are temporally stable across years in around 50% of the comparisons. Using an event study design, we find evidence that welfare estimates are stable within a year, that is, for weeks after a change in water quality. Our findings further reveal that having two years of data for predicting welfare estimates in subsequent years improves the prediction accuracy by 22% relative to using a single year of data, but further improvements in the prediction accuracy are modest when including additional years of data. Predictions of welfare estimates are not necessarily improved when using data closer in time to the prediction year. We discuss the implications of our results for using revealed preference studies in policy analysis.

Keywords: recreation demand, revealed preferences, temporal stability, prediction accuracy, water quality, welfare estimates

JEL codes: H41, Q26, Q51, Q53

Acknowledgments: We would like to thank Alberta Environment and Parks for the access to the data. We would like to thank participants of the 2022 Appalachian Experimental and Environmental Workshop, the 2022 Social Cost of Water Pollution Workshop, and the 2023 EAERE Annual Conference for valuable comments and suggestions to an earlier version of this paper. We gratefully acknowledge the support of Global Water Futures and the National Science Centre in Poland (Opus 19, grant no. 2020/37/B/HS4/01517).

1. Introduction

There is almost always a delay between the date of collecting preference and behavioral data and the application of the derived welfare measures in policy analysis. The temporal delay may only be months, often the delay is years, and sometimes it is decades. The delay is present in application of the data for policy assessments and benefit-cost analyses of projects with future streams of benefits. The underlying assumption made in these policy applications is that preferences, behaviors, and associated welfare measures are stable across the time of data collection, welfare estimation, and their application for policy purposes (Rolfe and Dyack 2019). Consequently, it is highly policy-relevant to understand whether welfare estimates are temporally stable and can be predicted based on existing welfare measures. In this study, we explore the question of whether recreation demand welfare estimates are stable over time and how the predictability of welfare measures across time depends on available data.

Revealed preference models of recreation demand have received relatively few stability assessments compared to other non-market valuation approaches, such as stated preference and hedonic price methods (Kling et al. 2012; Ji et al. 2020; Lupi et al. 2020).[1] Temporal stability is typically investigated through test-retest procedures by comparing welfare estimates from two or more points in time. This testing framework assumes that estimates of the same good obtained through the same research design from different points in time represent statistically indistinguishable values. While the literature is abundant with stated preference investigations of such test-retest type, the existing evidence about temporal stability of revealed preference recreation demand welfare estimates is scarce and predominantly based on survey data (e.g., Zandersen et al. 2007a, 2007b; Rolfe and Dyack 2019; Ji et al. 2020). With this study, we aim at contributing to empirical inquiries on temporal stability of recreation demand welfare estimates and their predictability based on existing estimates, by conducting the assessments with the use of administrative, not survey-based, revealed preference data.

The three purposes of this paper are: (i) to examine temporal stability across years; (ii) to assess temporal stability within a year; and (iii) to study the predictability of recreational demand

[1] Another term for temporal stability often used synonymously in non-market valuation literature is temporal reliability (e.g., Brouwer 2006; Liebe et al. 2016).

welfare estimates. The latter is assessed with respect to two dimensions: the numbers of years (data points) of existing data used to predict welfare measures and the time delay between the prediction year and the data points. To explore these research areas, we employ a regional recreation demand model using data for ten consecutive years. The empirical investigation is based on a large-scale set of administrative data on campground reservations in Alberta, Canada. The data contains daily information on overnight trips taken by recreationists over a season to provincial parks in Alberta for the years from 2013 to 2022. Each year of data includes 80,000 to 128,000 trips taken by 56,000 to 78,000 individuals to 58 campgrounds.

The welfare estimates considered in our study concern the societal value of lake water quality. The empirical data enables estimation of welfare effects of changes in lake water quality measured by the presence, or lack, of water quality advisories. Water quality advisories are issued by the local health authority and indicate the water body is unsafe for human or animal contact. Changes in water quality over time are captured by whether an advisory is present at a lake next to a campground. The advantage of this water quality measure, compared to a single annual average often used in existing valuation studies concerning water quality, is that it helps us capture the actual water quality condition experienced by individuals. Additionally, the advisories are salient for people compared to other scientific measures (indices) of water quality, which do not necessarily correspond to people's perceptions or may not be noticed by people.

Results of site choice models reveal that the marginal willingness to pay (MWTP) to remove a water quality advisory range from \$3 to \$26 per trip depending on the year considered.[2] Out of the total of 45 year-to-year pairwise comparisons, about half (23) reveal statistically indistinguishable MWTP values. We do not observe any systematic pattern that adjacent years generate more similar values than non-adjacent years. We use an event study design to assess the within-year stability of water quality welfare estimates based on week-to-week comparisons and find consistent values of lifting a water quality advisory across weeks of single years after the first week from the date of issuing an advisory. We find evidence that people respond to water advisories already in the first week of an advisory being in place, but the response is greater in (and consistent across) subsequent weeks. After the first week, the behavioral response and implied values of lifting an advisory range between \$13 and \$18 per trip, depending on the week, and these

[2] All value estimates in this study are in 2020 Canadian dollars.

estimates do not differ statistically from each other. We find no evidence for anticipatory effects as the placebo value of lifting an advisory is not statistically different from zero in the weeks preceding the advisory issuance.

We study the out-of-sample predictability of welfare estimates by using varying year subsets of the data to estimate and compare welfare measures. We find that having two years of data compared to one improves prediction accuracy by 22%, but the additional accuracy gains from using more than two years of data for predictions are more modest and are under 10%. Surprisingly, we do not observe that the time delay between the prediction year and the years of data used for making the predictions affect the accuracy of the prediction.

Most of the existing evidence suggests that recreation demand welfare estimates are unstable over time.[3] A summary of these studies along with a comparison with our study is presented in Appendix A.[4] Cooper and Loomis (1990) estimate consumer surplus per fishing trip in each year over a period of five consecutive years and observe the value of a five-year benefit stream to be underestimated by 17% when calculated as an estimate transferred from the base (i.e., first) year throughout subsequent years. Zandersen et al. (2007a, 2007b) use data from two points in time, twenty years apart, and find welfare estimates of recreation access to a forest site not to be temporally stable. Similarly, Rolfe and Dyack (2019) show that value estimates of a recreational trip to an estuarine region derived for two data points, seven years apart, are statistically different across the data years. Ji et al. (2020) use five points of recreation demand data over the span of eight years and observe welfare estimates of water quality changes not being temporally stable over the full range of years, but report similarities across selected pairs of adjacent years. The recreation demand literature also provides (some) evidence of temporal stability. Hellerstein (1993) reports mixed findings for a studied period of seven consecutive years depending on a model specification: own-year models reveal a recreation site value to be stable or slightly decreasing over time, while a more naive, as assessed by the authors, pooled model suggests that

[3] We do not refer here to recreation demand models based on contingent behavior data (e.g., Xie and Adamowicz 2022), which involves stated behaviors under hypothetical scenarios, in contrast to revealed preference recreation demand models relying on actual behaviors. We also do not refer to meta-analysis studies, which aggregate findings for various types and destinations of recreational trips and do not provide examination of temporal stability similar to test-retest procedures.

[4] The list has been prepared based on summaries of literature in other articles (e.g., Ji et al. 2020), thorough literature searches, and our best knowledge.

the value has been increasing over time. Yi and Herriges (2017) find estimates of the welfare loss from closure of two individual lakes to be fairly stable over a period of two consecutive years. All of the referred studies that are based on individual-level models rely on survey data.

This study brings several new contributions to the non-market valuation literature. First, it provides novel evidence on temporal stability of recreation demand welfare estimates. In contrast to former investigations in this area reviewed above, we use a large set of administrative data, which is not subject to recall issues or other behavioral biases observed in survey responses (Lades et al. forthcoming; Lupi et al. 2020). We employ a water quality measure easily understandable to people and measuring the actual experienced condition. Because our water quality measure is day-specific, instead of being a year or season average, we provide evidence on intra-annual temporal stability of welfare estimates which has not been the focus in the non-market valuation literature. The inclusion of the years 2020 and 2021 allows us to verify the stability of the welfare estimates in the face of a global pandemic and subsequent travel restrictions, which drastically changed recreation options. Second, building on the insights from temporal stability, we evaluate how well welfare estimates can be predicted based on two factors: the number of years of data used for forming the predictions and the time delay between the prediction year and the years of data used for making the predictions. This can be of practical importance for policy assessments requiring transferring welfare measures over time. Finally, we further argue that our day-specific trip and water quality data across ten years provides a unique opportunity to identify behavioral responses to changes in lake water quality. The data allows us to use an event study design and contributes to the need identified in recreational demand literature for expanded data sources allowing for validation of welfare estimates derived from single cross-section analyses (Lupi et al. 2020).

2. Data and methods

This section describes the data used in the analysis and the empirical approach. We start by discussing the data sources for camping trips and water quality advisories in Alberta, and the approach employed to calculate travel costs. We then describe the empirical strategy and modeling approach for the analysis.

2.1. Recreation data

We focus on trips by Albertans to basic campgrounds in provincial parks, using data from the Recreation Alberta Parks (RAP) database for the years 2013-2022. The online camping reservation system is the primary way people make camping reservations in the province and provides a single interface to compare campground availability, prices, and amenities.[5] The RAP database contains reservations made online through the RAP portal, by phone, and on-site walk-ins. The database includes the scheduled arrival and departure dates, group sizes, postal codes of the users making the reservations, and campground information. Each user of the online reservation portal is assigned a unique ID that stays with them over time, and thus we can connect multiple trips by the same person within a camping season and across years.

The raw data includes 2.19 million camping reservations from 0.95 million unique users and we apply a set of exclusion criteria for the data to ensure we have a well-defined commodity and to mitigate the potential issue of multi-purpose trips (Lupi et al., 2020). In specifying the choice set, we focus on basic campgrounds and exclude group camping, comfort camping, and equestrian campgrounds. We further restrict the sample to the set of 58 campgrounds that we have trip information for every year of the analysis, although we also consider models with all available sites each year as a sensitivity analysis and present the results in Appendix D.[6] We exclude individuals without a valid Alberta postal code as these out of province travelers are likely to have multiple purposes for their trip to Alberta besides camping and/or are more likely to visit multiple sites. We only include trips that are seven nights or less and involve five people or less for the purpose of studying a homogenous good. We focus the main results on trips taken during the summer season defined as between the third Monday in May and the first Monday in September (Victoria Day and Labor Day long weekends). Some campgrounds open later in the season and we only consider a campground as available to visit if at least one person made a reservation to that campground for a given day.

[5] The online reservation system of RAP: https://www.albertaparks.ca/albertaparksca/visit-our-parks/camping-in-alberta-parks/online-reservations/

[6] Some campgrounds were added to the RAP system throughout the study period and some campgrounds were closed in certain years due to fires and floods. The number of sites available in RAP over the years 2013-2022 ranged from 62 to 111.

With these exclusion criteria to the data applied, we use 977,834 camping trips made by 364,157 unique visitors over the ten-year period. Table 1 shows that each year includes approximately 56,000 to 78,000 unique individuals taking 80,000 to 128,000 trips.[7] The 2020 and 2021 years have around 30% more trips than previous years in large part due to out-of-province travel restrictions related to the COVID-19 pandemic.

2.2. Water quality advisory data

Water quality advisory data are obtained from Alberta Health Services (AHS). The AHS conducts a Routine Recreational Water Quality Monitoring Program involving a visual inspection of lakes and collecting water samples for lab testing. Advisories are issued when a blue-green algae (cynobacteria) bloom is identified and it poses a risk to human health. Advisories remain in place as long as there is a health risk, and this typically—but not always—lasts until the end of recreation season. We collect information on the dates of advisory issuance and lifting for each provincial park with lake access in the dataset. Advisories are published on the AHS website and shared with local media. Figure 1 presents an example of one of these advisory signs posted near a lake.

Most recreation demand studies involving water quality use scientific measures of water quality, such as Secchi depth or total phosphorus, and there is an ongoing debate in the literature on which scientific measure matters for people's behavior and welfare (Egan et al. 2009; Ji and Keiser 2016). Most studies also use a single measurement or annual average of water quality for the whole year. However, water quality in a waterbody can change substantially within a year, for example, due to changing temperature and precipitation events and thus a single annual measure might not reflect the water quality people actually face during their recreation activity. We use the dates when water quality advisories are issued and lifted by the provincial health authority throughout the camping season in response to changing water quality conditions. We expect these advisories to be more salient for people than scientific metrics of water quality as they provide a simple indication of water quality and thus may better fit people's understanding of water quality.

[7] While our interest is in temporal stability of recreation demand welfare estimates, we do not restrict the sample of users to be the same each year due to data limitations. If we restrict the sample to include only the users taking trips each year, we are left with a sample size of 1,027 which is too small to identify the effect of advisories and with many campgrounds unvisited each year. Thus, we cannot rule out that the different composition of individuals across years is driving our finding of limited temporal stability.

Furthermore, the advisories are issued throughout the year, reflecting the time-varying nature of water quality.

Of the 58 campgrounds included in the analysis, 48 campgrounds are adjacent to a lake with the majority of the remaining campgrounds being alongside a river. A total of 22 campgrounds had at least one advisory between 2013 and 2022. Figure 2 shows the timing of the advisories for all campgrounds considered in the study over the ten years. Campgrounds without a water quality advisory in the studied period are omitted in the figure. The figure illustrates the variation in water quality across campgrounds and years that is used to identify the effect of water quality advisories on camping trips. The number of campgrounds with advisories varied across the years between 5 and 13, and the total number of days with advisories per year ranged from 89 to 533, as shown in Table 1.

2.3. Travel costs calculation

To calculate travel costs, we use the information on out of pocket travel expenses, the opportunity cost of time, and camping costs. Driving distance and time are measured from the centroid of each person's postal code to all available campgrounds using OpenStreetMap data. Driving costs are obtained from the Canadian Automobile Association's driving cost calculator. We use the average marginal costs of driving a compact car, sport utility vehicle, and pickup truck. The driving cost per km includes fuel use and maintenance costs, but excludes fixed costs such as age-related depreciation, registration costs, and insurance costs (Lupi et al., 2020). For fuel costs, we use average retail gas prices in Alberta between May and September for each year from Statistics Canada. Monetary driving costs are relatively flat over the ten years ranging from \$0.22 to \$0.32 per kilometer in nominal dollars.

For the value of travel time, we use postal code level median annual household income from the 2016 Canadian census, as year-specific income levels are unavailable for the ten-year study period. The annual income is converted to approximate the hourly wage using 2080 hours worked in a year. We use 2/3 of the imputed wage rate as the value of time. Based on previous modeling work, using a subset of this data, this assumption is observed to fit the data better than the more conventional 1/3 of the wage rate (Lloyd-Smith and Becker, 2020). We conduct a sensitivity analysis using 1/3 of the wage rate for the value of time and these results are presented in Appendix C.

The nightly camping fees for each campground and year are provided by Alberta Environment and Parks and range from \$18 to \$54 during the study period. Moreover, the recreationists are charged \$12 when making a reservation regardless of the reservation length and the party size. A typical camping reservation is for two to three people for a total of two nights. The total camping cost defined for individual *i* recreating at campground *j* in year *y* is represented in equation (1) and includes the campground fee per night in a given year (CF_{jy}), the number of nights in the reservation (NN_{ijy}), and the flat reservation fee of \$12 (R):

$$CC_{ijy} = CF_{jy} * NN_{ijy} + R. \tag{1}$$

Equation (2) shows the overall structure of the return-trip travel cost calculation for person *i* to travel to campground *j* during year *y*:

$$TC_{ijy} = 2 * \left[\left(\frac{DD_{ij} * DC_y + CC_{ijy}}{NC_{ij}}\right) + \left(\frac{2}{3} * \frac{Inc_i}{2080}\right) * DH_{ij}\right], \tag{2}$$

where DD_{ij} is the distance in kilometers from individual *i*'s postal code centroid to campground *j*, DC_y is the per kilometer driving cost for each year, CC_{ijy} is the camping cost, NC_{ij} is the number of people included in the reservation, $\frac{2}{3}$ is the opportunity of time cost ratio, Inc_i is the median annual household income in individual *i*'s postal code area, 2080 is the average number of worked hours in a year, and DH_{ij} is the one-way traveling time in hours. All travel costs have been adjusted for inflation to 2020 Canadian dollars.

In the summary statistics in Table 1, we report average travel costs for recreationists for the campgrounds actually selected by them and for all provincial campgrounds. As expected, people choose lower cost camping trips, among the options available for camping in Alberta.

2.4. Methods for assessing temporal stability across years

We model camping demand using a random utility framework (McFadden 1974; Parsons 2017). Specifically, we use a multinomial logit site choice model where the utility that person *i* receives from visiting campground *j* at year *y* and date *d* as:

$$U_{ijyd} = \beta_y^{ASC}\mathrm{ASC_{jy}} + \beta_y^{TC}TC_{ijy} + \beta_y^{WQ}WQ_{jyd} + \varepsilon_{ijyd}\ , \qquad (3)$$

where an alternative specific constant ($\mathrm{ASC_{jy}}$) is equal to one for campground j and zero otherwise; TC_{ijy} is the year-specific return-trip travel costs as defined in (2) and divided by 100 to help the convergence; WQ_{jyd} is a dummy variable equal to one if a water quality advisory is in place at campground j during year y and date d; and ε_{ijyd} is the error term. The parameters to be estimated are β_y^{ASC}, β_y^{TC}, and β_y^{WQ}. Robust standard errors are clustered at the individual level to account for individuals that take multiple trips in a year.

The site choice model used in the analysis here does not consider the participation decision of whether to go camping or stay at home and hence, in this sense, it is not a repeated discrete choice model that is often used to model seasonal recreation demand (Lupi et al. 2020). Preliminary results using a nested logit repeated discrete choice model on a subset of the data showed that the participation and site choice decisions are independent. This finding is perhaps not surprising as the vast majority of people take only a single camping trip each year (as seen from the comparison of the numbers of individuals and trips in the data summarized in Table 1). Adopting a repeated discrete choice modeling framework also raises the issue of how to divide the camping season into choice occasions. For the above mentioned preliminary analysis, we divided the camping season into seventeen 'weeks' from Wednesday to Tuesday, but we recognize the arbitrary nature of the choice occasion specification. Taking into account these considerations, we opt for the site choice model as a better-grounded modeling framework for our investigation.

We first estimate equation (3) separately for each of the 2013 to 2022 data years. We then calculate MWTP for removing one water quality advisory during a trip in year y, where $y = \{2013, \ldots, 2022\}$, using the negative of the ratio of the model coefficients as following:

$$MWTP_y = -\beta_y^{WQ}/\beta_y^{TC}. \qquad (4)$$

To assess the temporal stability of the MWTP estimates across the years, we compare the MWTP from each year to the MWTP estimates from every other year, which yields 90 pairwise tests (45 unique comparisons). We use a two-sided t-test for assessing statistical differences in these comparisons.

2.5. *Methods for assessing temporal stability within a year*

To assess temporal stability of welfare estimates within a year, we estimate a multi-year site choice model with the data from all study years pooled and include water quality advisory variables to capture time periods leading up to and since an advisory is issued. We bin our daily data into weeks which are defined from Wednesday to Tuesday to take into account long weekends. This setup is similar to an event study design where the treatment is the issuance of an advisory and we consider both the pre- and post-treatment periods (Marcus and Santa'Anna 2021).

We estimate the following specification for modeling the utility that person *i* receives from visiting campground *j* at year *y* and date (here, week) *d*:

$$U_{ijyd} = \beta_y^{ASC}\mathrm{ASC_{jy}} + \beta_y^{TC}TC_{ijy} + \sum_{k=T_0}^{-2}\beta_k^{WQ}WQ_{jyd} + \sum_{k=0}^{T_1}\beta_k^{WQ}WQ_{jyd} + \beta_y^{NO}NoWQ_{jy} + \varepsilon_{ijyd}\ , (5)$$

where ASC_{jy}, TC_{ij}, and WQ_{jyd} are defined as in (3), but travel costs are assumed not to be year-specific; $NoWQ_{jy}$ is equal to one for campgrounds that did not have an advisory in a specific year and zero otherwise; T_0 and T_1 are the lowest and highest number of weekly leads and lags, respectively, surrounding the treatment period (that is, the water quality advisory issuance date); ε_{ijyd} represents the error term; and β^{TC} and β_k^{WQ} are the main parameters of interest to be estimated. We set T_0 to be equal to -5 (five weeks or more than 28 days prior to a water quality advisory issuance date) and T_1 to be equal to 5 (five weeks or more than 27 days after a water quality advisory issuance date). The selection of the numbers of weeks is guided by insights from preliminary analysis and intention to have sufficient numbers of observations per parameter estimate. We further set the reference period to be one week before an advisory and fix the parameter equal to 0 (i.e., $\beta_{-1}^{WQ} = 0$), such that the parameters of interest (i.e., β_k^{WQ} for $k \neq -1$) can be interpreted relative to the week before the advisory. The comparison of the parameter estimates of these variables helps us assess the within-year temporal stability.

We consider two different model specifications of equation (5) due to some computational considerations. First, we specify a model with year-specific ASC variables (ASC_{jy}) for each campground to capture time varying changes in unobserved campground quality that might be correlated with water quality. This model specification includes 581 parameters, including 570

ASCs. We also estimate a model using a single ASC parameter for each site (ASC_j) and 68 parameters overall.

2.6. Methods for assessing predictability

To study predictability of recreation demand welfare estimates, we use the same modeling setup based on site choice models as the one applied for assessing across-years stability. However, here, we employ subsets of the ten-year database selected based on years of data. Specifically, for deriving the predicted welfare measures, we consider all unique multi- and single-year subsets of the database that allow for the prediction of a future welfare estimate within the study timeframe from 2013 through 2022. Separate models are estimated on each data subset, and the parameters estimated from each model are used to calculate MWTP for removing one water quality advisory and to compare this (predicted) welfare measure to the MWTP estimates from the left-out (i.e., not included in the subset) future single-year estimates. In this comparison, the latter MWTP values derived from the single-year models are treated as the actual values, which are compared to the predicted values obtained from the models estimated on subsets of past-years data. This helps us assess whether and how accurately we can predict welfare estimates in a given year based on empirical data from other years. The analysis is based on 499 multi-year site choice models and the ten year-specific models introduced in Section 2.4. We evaluate the accuracy of the predictions based on the absolute difference between the predicted MWTP based on the past-years models and the actual MWTP derived from models using that year of data.

The assessment of the welfare estimate predictability concerns two aspects. First, we assess the predictability of MWTP by the number of years of data included in the models used for deriving the predicted values. We hypothesize that including more years of data in the prediction models improves the prediction accuracy, and we quantitatively evaluate this hypothesis. Second, we assess the predictability of MWTP by the time delay between the data used for making the value predictions and the actual year for which the prediction is made. For example, if the 2016 and 2018 years of data are used to predict the 2020-year MWTP, then the average time lag in years is equal to |(2020-2016)+(2020-2018)|/2 = 3. We hypothesize that the accuracy of the predicted values improves when time lags decrease, and we also quantitatively evaluate this hypothesis.

3. Results

The description of the results is organized according to the three areas of our study. First, we examine the temporal stability of the welfare estimates across years. Second, we investigate the temporal stability of the welfare estimates within a year. And third, we study changes in the accuracy of welfare predictions when using more versus fewer years of data in the analysis and when using adjacent versus distant years of data for making the predictions.

3.1 Across-year temporal stability

Table 3 presents the results of the multinomial logit site choice model estimated separately for each of the ten years considered in the analysis following the model specification defined in equation (3). The table focuses on two main parameters of interest: one indicating if a water quality advisory for a lake next to a campground is in place during the visit and the other capturing the travel cost (in hundreds). The models also include ASCs for all 58 studied campgrounds to control for unobserved campground characteristics, and these ASC estimates are reported in Appendix B. For all models, the travel cost parameter is negative and precisely estimated as expected given the large number of observations. The parameter for the water quality advisory is also negative in all models and statistically different from zero for all years except the 2015-year model. Due to potential scale differences, the parameter estimates cannot be directly compared across the year-specific models so we focus the discussion on welfare measures.

A MWTP value for lifting a water quality advisory can be calculated as a ratio of the parameter estimate for water quality advisory and the negative of the parameter estimate for the travel cost (following equation (4)). The MTWP values derived from the year-specific models are presented graphically in Figure 3, which helps us illustrate the variation of the estimates across the study period.

The marginal willingness-to-pay values to lift a water quality advisory range from \$3 to \$26 per trip, depending on the year considered. There does not appear to be a clear temporal trend and the lowest value of \$3 is observed for the year 2015 and the highest value of \$26 is estimated for the year 2014. Six of these year-specific MWTP estimates are between \$10 and \$20 while the estimate for the year 2019 is \$6 and for the year 2022, it is \$22.

We formally test for the statistical significance of the estimated differences in MWTP across the different years using a t-test of difference in means. In Figure 4, we illustrate the differences between the value estimates from each year to every other year, along with respective confidence intervals. The differences are ranked from smallest to largest in the figure. Of the 45 unique pairwise comparisons, 23 (51%) suggest statistically indistinguishable values. These results are robust to the alternative value of time assumption and inclusion of additional sites.[8] As expected from the results reported in Figure 3, the estimates for the years 2014 and 2015 are most substantially different from the estimates for the other years, which can be easily observed in Figure 5. For the years 2014 and 2015, the value estimates are statistically significantly different in 15 out of 18 (83%) of year-to-year comparisons). The value estimates from the remaining years are more similar and statistically different from each other in only 7 out of 27 (26%) year-to-year comparisons.

We do not observe any systematic pattern that adjacent years generate more similar values than non-adjacent years. Although it could be true, for example, for the year 2018, where the value estimate is not statistically different from the adjacent values and significantly differs from the estimates for 2014 and 2022, no similar relationship holds for other-year value estimates. For instance, the estimate for 2019 is statistically indistinguishable from the one for 2018, but is different from the one for 2020.

The data covers the time period during the global pandemic of Covid-19, which was related to a substantial change in lives of many people and thus could importantly affect recreation welfare estimates. The years 2020, 2021, and 2022 can be categorized as the beginning, middle, and end of the pandemic as by spring of 2022 most of Canada's pandemic restriction policies had been lifted using the Oxford Covid-19 Government Response Tracker Stringency Index (Hale 2021). We do not observe, however, that these years are related to any different results than other years in terms of the value estimates and the temporal stability comparisons. We thus conclude that the

[8] Figure C2 in Appendix C presents the comparisons of models that differ from those discussed here only in that 1/3 of the wage rate is used for the value of time, and the percentage of statistically indistinguishable values is also 51%. Figure D2 in Appendix D presents the comparisons of models that differ from those discussed here only in that the data for all available sites each year is used in the estimation, and the percentage of statistically indistinguishable values is the same at 51%

pandemic has not had a significant impact on the recreational demand welfare estimates of lake water quality in this case study.

3.2 Within-year temporal stability

Figure 6 reports the results from the event study specification to assess how MWTP varies in the weeks leading up to and after the issuance of the water quality advisory. Note that the parameter for the week directly before the advisory issuance date is fixed to be zero such that all MWTP estimates can be interpreted relative to this period. We present the results using a model with year-specific ASCs (solid line) and a single ASC for each campground (dashed line), and both models include all ten years of data.

The graphical analysis indicates that before an advisory is issued, the value of lifting an advisory is not statistically different from zero. This finding is in line with expectations and gives more confidence to the identification strategy used as we would not expect people to be willing to pay for lifting an advisory before it occurred. Once an advisory is issued, the MWTP is estimated to be $6 for the initial week (days 0 to 6) and then ranges from $13 to $19 for the following weeks. The lower value in the week of issuing an advisory might reflect that it takes time for people to learn about the advisory and adjust travel choices. The MWTP values after the initial week of issuing the advisory are not statistically different from each other, providing evidence for temporal stability. We note that in our data, once an advisory is issued for a given lake, it typically stays in effect till the end of the season (see Figure 2), which justifies the expectation and the finding of the within-year MWTP stability. In few cases in the studied period, the advisory was lifted before the end of the season. Related research shows that it does not necessarily affect the visitation behavior and, thus, may have limited effect on the welfare estimates, as people tend to avoid sites with earlier environmental warnings even after the event causing the warning has ended (Boudreaux et al. 2023). Based on our model estimates, we conclude that the within-year welfare estimates are temporally stable after the initial adjustment period.

3.3 Predictability of welfare estimates across years

In this section, we examine how predictable future welfare estimates are using existing data. We consider a total of 499 unique MWTP predictions. Each of these values serve as predicted values for the future years that are not included in estimation of a given model. At the same time, each of the single-year models provides a MWTP estimate that we consider as the actual MWTP for the given year of data on which the model is estimated. In that sense, the single-year models provide criterion values for assessing the accuracy of the predicted values. To measure the accuracy of the predictions, we use an intuitive measure of an absolute difference between the actual value and the predicted value. We evaluate the predictability of the MWTP values along two dimensions: the number of years of data used for forming the predictions and the lag between the data used for making the prediction and the year for which the prediction is made.

Figure 7 illustrates the results for the first dimension, namely it shows the magnitude of the accuracy improvement in the predictions upon increasing the number of years of data for forming the predictions. The figure presents only predictions with eight or less years of data as there is only one prediction possible using nine years of data (that is, estimating a model on nine years of data, only one year is out for which the value prediction can be made). The left panel shows the distribution of absolute differences in predicted versus actual MWTP values using different numbers of years of data for the predictions. The shaded area represents the interquartile ranges and the solid line is the average absolute difference. We see that the shaded area shrinks as the number of years of data increases suggesting that the range of differences in MWTP predictions is becoming smaller. The average absolute differences in MWTP also decrease as more years of data are taken into account for deriving predictions.

The right panel of Figure 7 presents the percentage reductions in absolute MWTP differences as one more year of data is added to the model used for making the predictions. Moving from making the predictions based on one year of data to two years of data, the accuracy of the prediction increases by 22%. Further moving from two years of data to three years of data improves the accuracy by around 9%. A similar improvement in the prediction accuracy is observed for considering four versus three years of data for forming the predictions and for considering five versus four years of data. Subsequent extensions of the number of years included in the prediction model improve the prediction accuracy to an even smaller degree. These results reveal that the largest benefit in terms of improving the prediction accuracy is obtained when extending the range of considered data for making the predictions from one to two years. Further extensions of the number of the considered years contribute to a much smaller degree to the prediction accuracy.

Figure 8 illustrates the results to the question of whether the time delay between the data used for deriving the predictions and the prediction year affects the prediction accuracy. The figure is a scatter plot representing the average absolute difference in years between the prediction year and the years of data used for making the predictions, and the absolute difference between the actual and predicted values. A solid line is added to the graph to illustrate the relationship. However, the line is nearly horizontal, as well as the scatter plot does not suggest any relationship. These findings show that the time delay between the prediction year and the years of data used for making the predictions does not affect the accuracy of the prediction in the studied dataset.

4. Conclusion

This study uses a large-scale administrative dataset on recreation behavior to shed new insights into the temporal stability and predictability of welfare estimates. Our main findings reveal that: (i) year-specific welfare estimates of lake water quality are stable across years in around half of the comparisons; (ii) the welfare estimates are temporally stable within a year; (iii) the accuracy of predicting future welfare estimates can be improved by using two years of data instead of one, but additional years of data bring only modest accuracy improvements; and (iv) the time delay between the prediction year and the years of data used for making the predictions do not affect the accuracy of predicted welfare estimates of lake water quality. The finding that across-years welfare estimates are statistically equal in about 50% of the comparisons point to the importance of having a large span of data for drawing conclusions on temporal stability of welfare estimates. For instance, if the analysis was conducted on only two years of data, the temporal stability assessment conclusion could go either way. The range of ten years included in this analysis gives a broader and more thorough picture of temporal stability of recreation demand welfare estimates.

Although a clear explanation of differences across years in the welfare estimates would bring an important insight to the scientific discussion on stability of preferences, visitation choices, and resulting welfare measures, we argue that a lack of any clear pattern in our year-specific findings may be a result of time dependent measurement noise. We have explored various possible explanations for the welfare estimate differences across years, but have not found the key drivers We have looked at correlations between the welfare estimates with the number of advisories in a given year (i.e., year-specific intensity of the water quality problem) and to the number of trips in a given year (i.e., demand for campground trips), however, this analysis does not reveal any correlations. We have further considered whether environmental conditions such as weather or economic conditions in the province of Alberta could explain the observed differences, however, we do not find significant variation in the environmental and economic measures (e.g., change in oil prices) that correlates with the welfare estimates differences. However, the lack of possible systematic relationships with other factors may be a signal of the differences being driven by time dependent measurement noise, where the noise can encompass any unobserved factors, such as some changes in the water quality awareness or unobservable variations in conditions in Alberta. This explanation reinforces the NOAA's panel recommendations for stated preference studies to average welfare estimates taken from data collected at different points in time to reduce time dependent measurement noise.[9] Future research should focus on identifying the factors that cause, or at least correlate with, temporal stability of recreation demand welfare estimates to improve predictions going forward. A broader research perspective would be to identify the conditions for preference stability over time and to develop measures of how much preferences change due to various factors.

Our finding that using adjacent years do not necessarily generate more accurate predictions of welfare is somewhat in contrast with the suggestions from the benefit transfer literature that value transfers generally become less reliable over time (Johnston et al. 2018). Nevertheless, this finding mirrors our conclusions from other parts of the analysis including lack of temporal trends in the welfare estimates and evidence of lack of significant differences in the welfare effects in 50% of the comparisons. Although this single set of data does not allow us to draw broader conclusions about the reasons underlying the divergence between our finding and the benefit transfer literature, we hypothesize that it could be linked to the specific nature of this type of data, based on administrative campground records, while the majority of empirical insights in the benefit transfer literature is derived from survey-based methods. Again, this provides a valuable avenue for future studies whether the degree of similarity of welfare measures from adjacent measures could be tied to the type of data used for deriving the welfare estimates.

We conclude by noting that while there are benefits of using administrative reservation data for recreation demand modeling more broadly, and temporal stability evaluations in

[9] The full text of the recommendation is "Time dependent measurement noise should be reduced by averaging across independently drawn samples taken at different points in time. A clear and substantial time trend in the responses would cast doubt on the 'reliability' of the finding" (Arrow et al. 1993, p. 33).

particular, there are also some potential issues. First, we do not have information on when a reservation is made so we cannot control for potential time lags between the reservation and visit date (Gellman et al. 2023). Water quality can change between the reservation and travel dates and people may stick with their original chosen campground due to commonly observed default effects in many human choices (Jachimowicz et al. 2019). A second issue is that popular campgrounds are often full on summer weekends and are not available to others making last minute camping plans. Since we do not know when people made their reservations, we cannot control for the capacity constraints and address issues of latent demand as we assume that all sites are available to everyone. Taken together, these two issues suggest that camping reservations may not be as responsive to changing water quality conditions as the underlying demand. Consequently, our empirical approach likely captures a lower-bound demand response of water advisories, and thus a lower-bound welfare estimate, of lifting water advisories. A third issue is the role of cancellations and 'no-shows'. People who cancel a reservation are not included in the database, but we also do not know how many people canceled their reservations. Given that the reservations are refundable except for the reservation fee, we expect that most people who do not expect to actually show up do cancel the trip. People might also not show up for their reservation and these no-shows will still appear in the database. To the extent that people do not show up for their reservation and this change in plans is driven by a water quality advisory, the estimated results provide a lower bound of the willingness-to-pay value for removing an advisory.

Tables

Table 1: Summary statistics for recreation and water quality data by year

	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Number of individuals	55,947	60,063	63,831	61,953	64,760	63,074	62,495	78,275	77,760	66,973
Number of trips	80,240	87,648	93,875	89,612	93,584	91,067	91,140	127,812	123,369	99,487
Travel costs for chosen sites	$256	$255	$251	$253	$256	$260	$255	$254	$250	$228
Travel costs for all sites	$512	$509	$495	$492	$494	$500	$488	$458	$451	$437
Number of campgrounds with water quality advisories	7	8	8	11	11	5	7	13	12	12
Number of days with water quality advisories (for all campgrounds jointly)	89	305	302	517	411	201	239	533	524	399

Table 2: Multinomial logit site choice model parameter estimates for each year

	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Water quality advisory	-0.104	-0.213	-0.027	-0.091	-0.128	-0.082	-0.049	-0.148	-0.129	-0.200
	(0.025)	(0.021)	(0.022)	(0.019)	(0.019)	(0.033)	(0.025)	(0.015)	(0.018)	(0.023)
Travel cost ($00s)	-0.797	-0.822	-0.851	-0.815	-0.801	-0.784	-0.796	-0.763	-0.784	-0.909
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
Campground ASCs	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
Log-likelihood value at convergence	-244,860	-267,640	-287,693	-278,507	-292,015	-284,236	-285,233	-420,456	-406,363	-314,673
Number of individuals	55,947	60,063	63,831	61,953	64,760	63,074	62,495	78,275	77,760	66,973
Number of observations	80,240	87,648	93,875	89,612	93,584	91,067	91,140	127,812	123,369	99,487

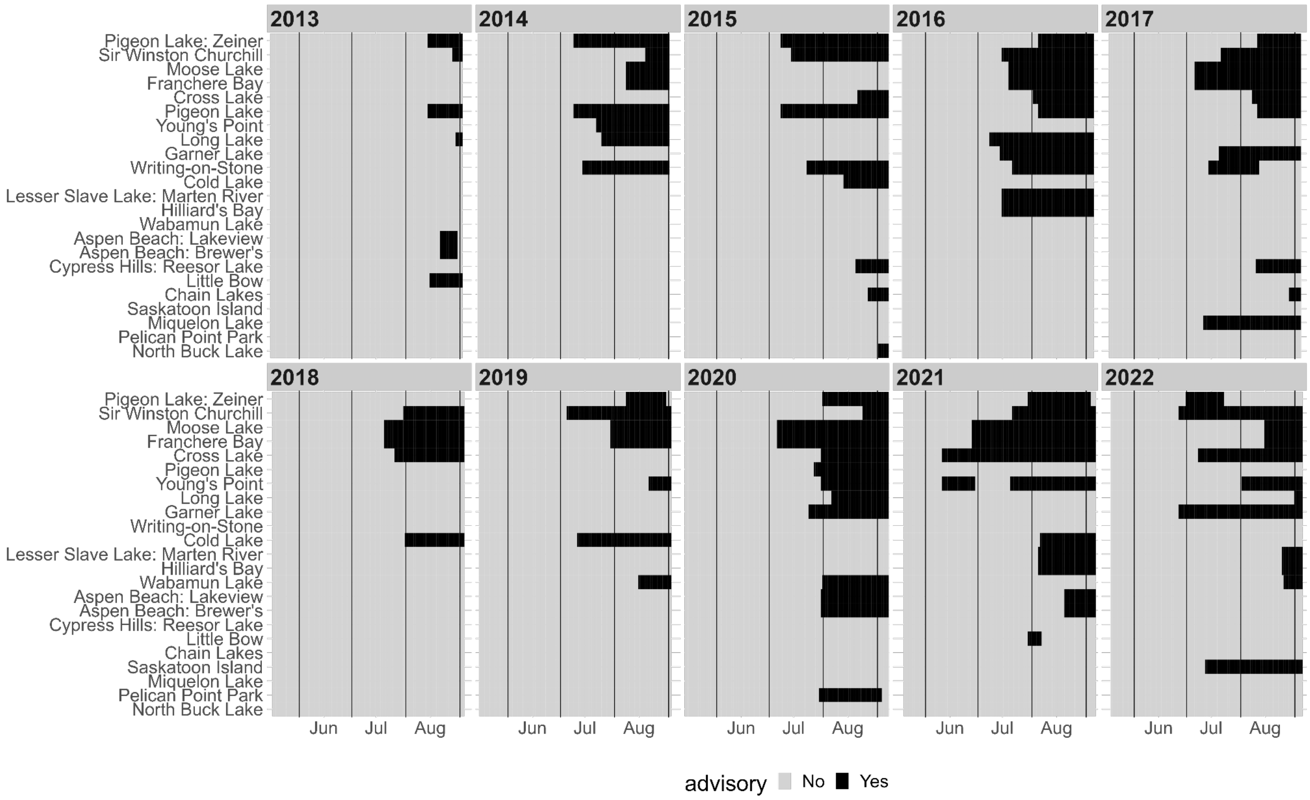
Notes: This table displays results from ten separate multinomial logit site choice models of recreational demand for the years 2013-2022 for a balanced panel of 58 campgrounds. All models include travel costs (in $00s), a zero-one-coded variable indicating if a water quality advisory was present at the campground for a specific date, and alternative specific constants (ASCs) for each campground to control for unobserved campground characteristics. Robust standard errors are clustered at the individual level and reported in parentheses.

Figures

Figure 1: Example of a water quality advisory posting

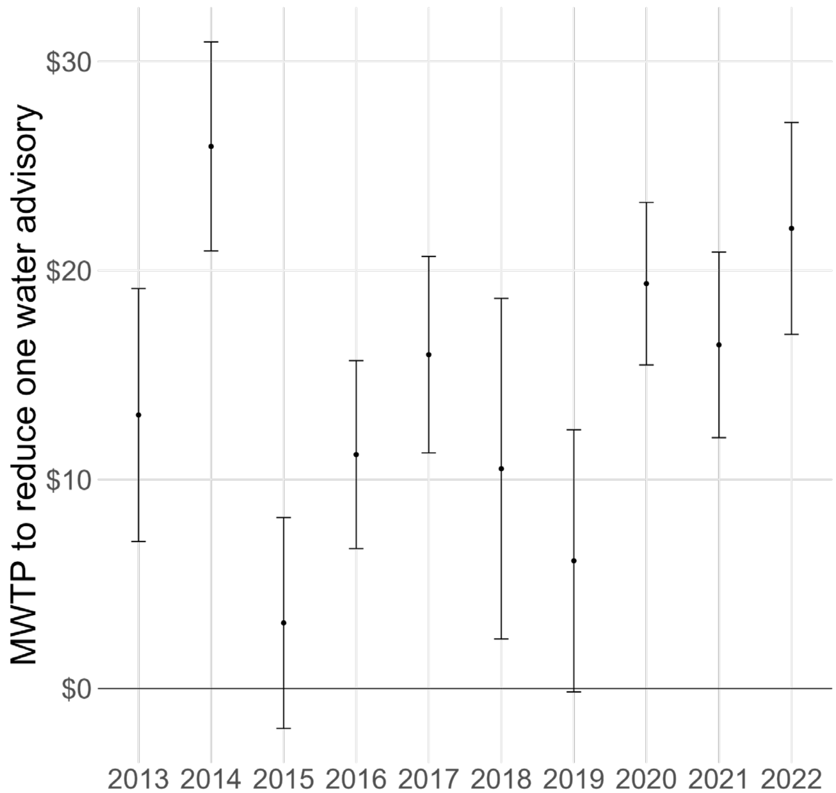
Source: File photo from CTV News.

Figure 2: Water quality advisory timing by campground and year



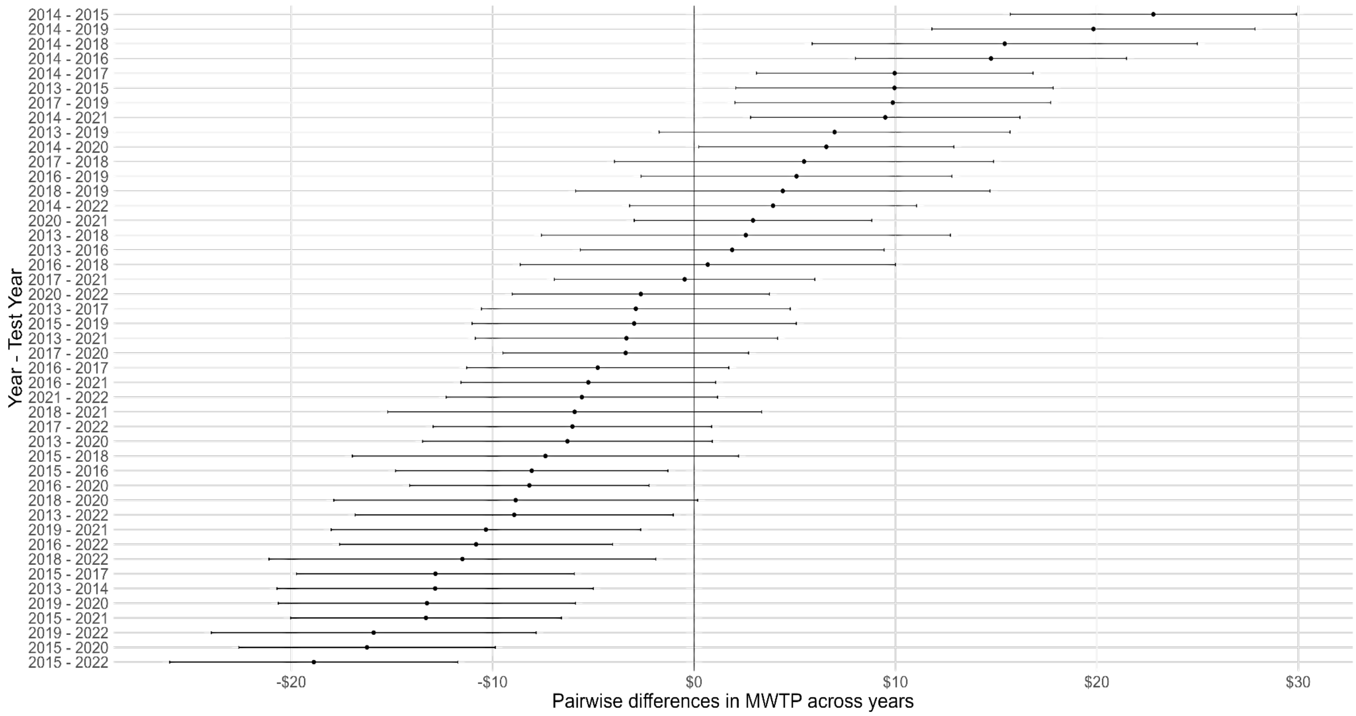
Notes: The figure presents information only for the studied period between the Victoria Day and Labor Day long weekends in Canada (the third Monday in May and the first Monday in September). Campgrounds without a water quality advisory in the studied period are omitted in the figure.

Figure 3: Marginal willingness to pay (MWTP) per trip for removing one water quality advisory



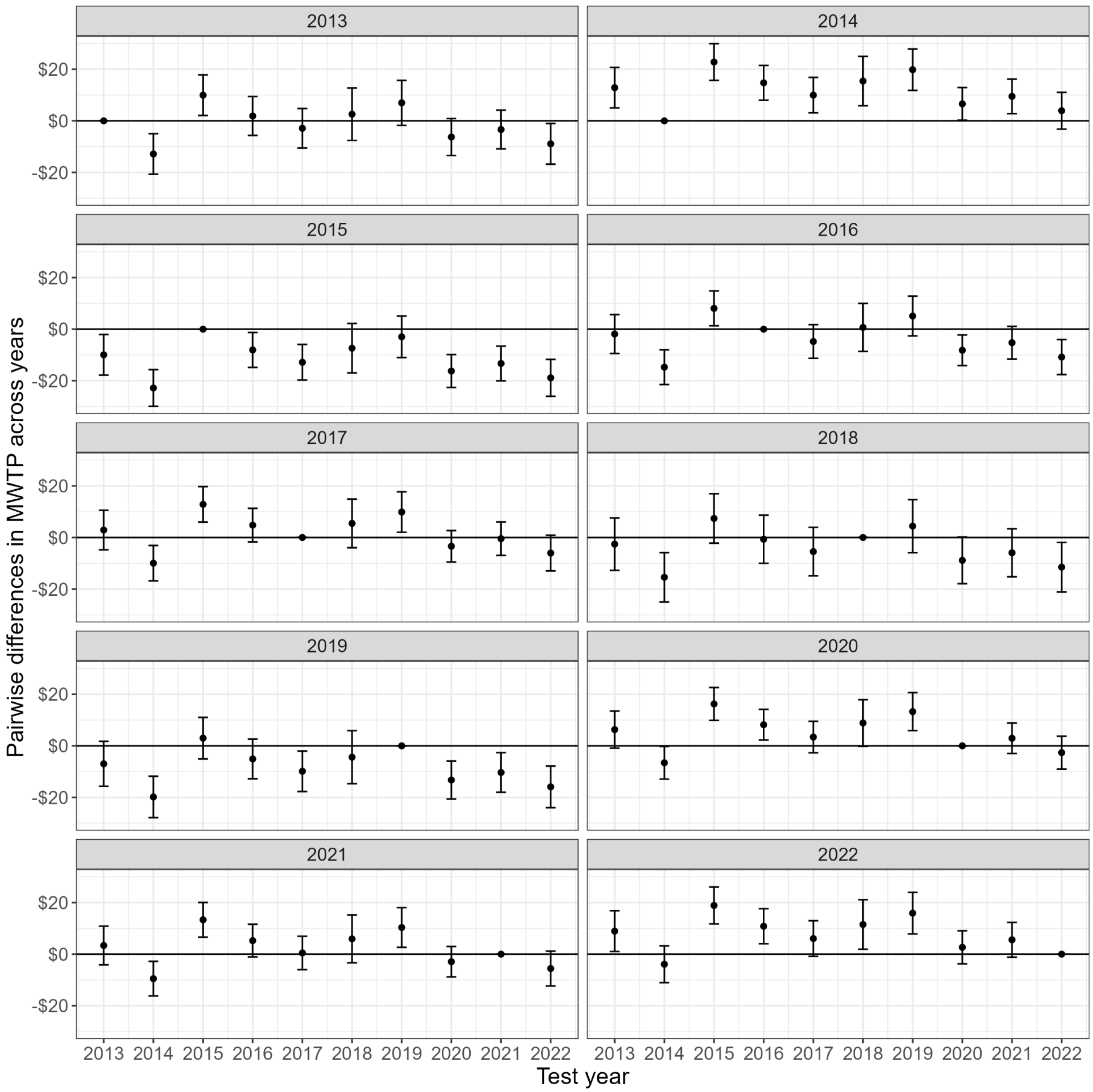
Notes: The figure reports the average MWTP per trip to reduce one water quality advisory at a campground for each of the ten year-specific multinomial logit site choice models. The dots represent average estimates and the capped vertical lines represent 95% confidence intervals of the MWTP values.

Figure 4: Difference in marginal willingness to pay (MWTP) to reduce one water quality advisory across years



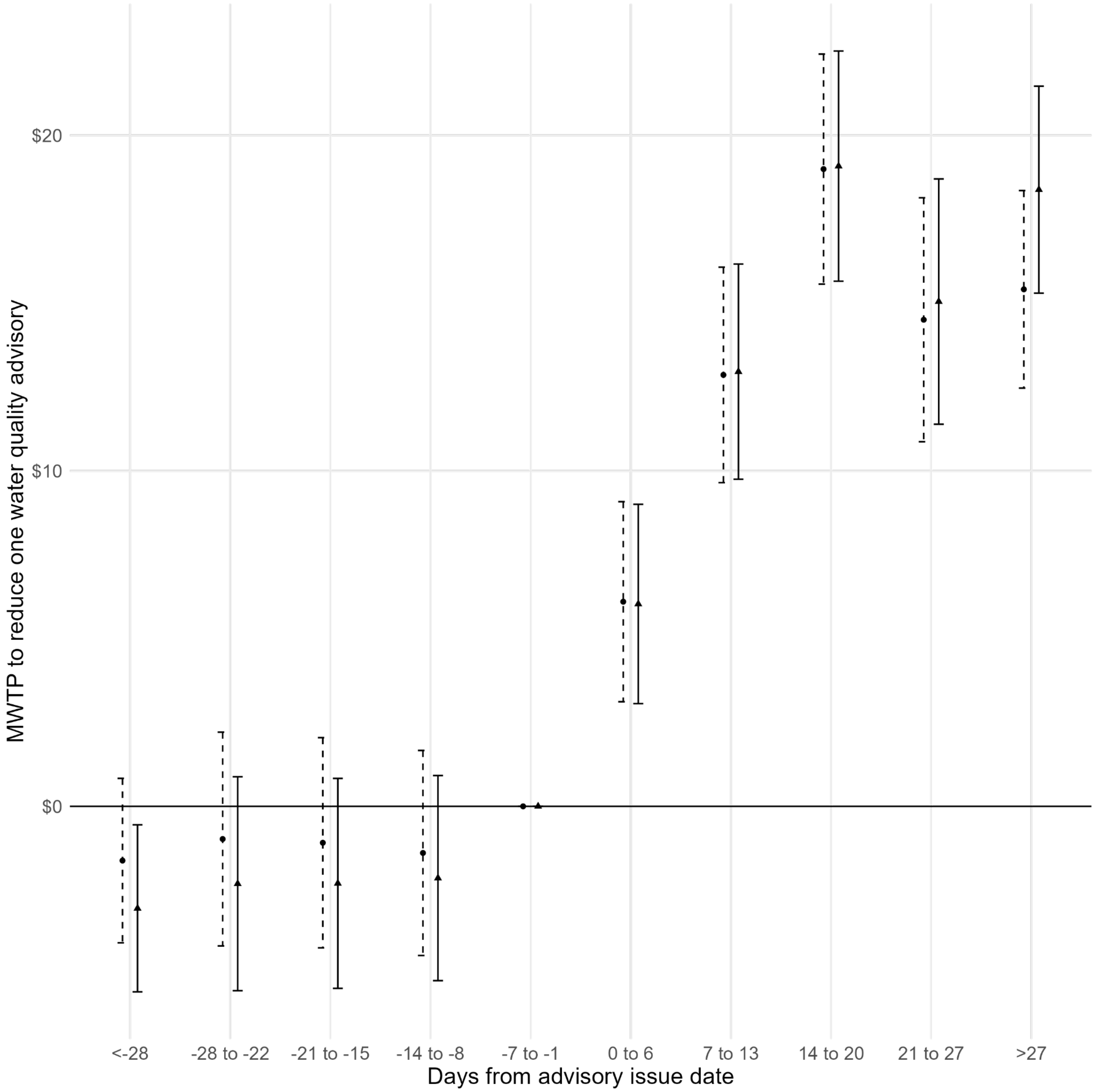
Notes: The figure shows pairwise MWTP differences in a given year to each of the other "Test Years" and ranks the differences from lowest to highest. The dots represent average estimates and the capped vertical lines represent 95% confidence intervals. The MWTP estimates come from the ten year-specific multinomial logit site choice models.

Figure 5: Comparison of marginal willingness to pay (MWTP) to reduce one water quality advisory across years



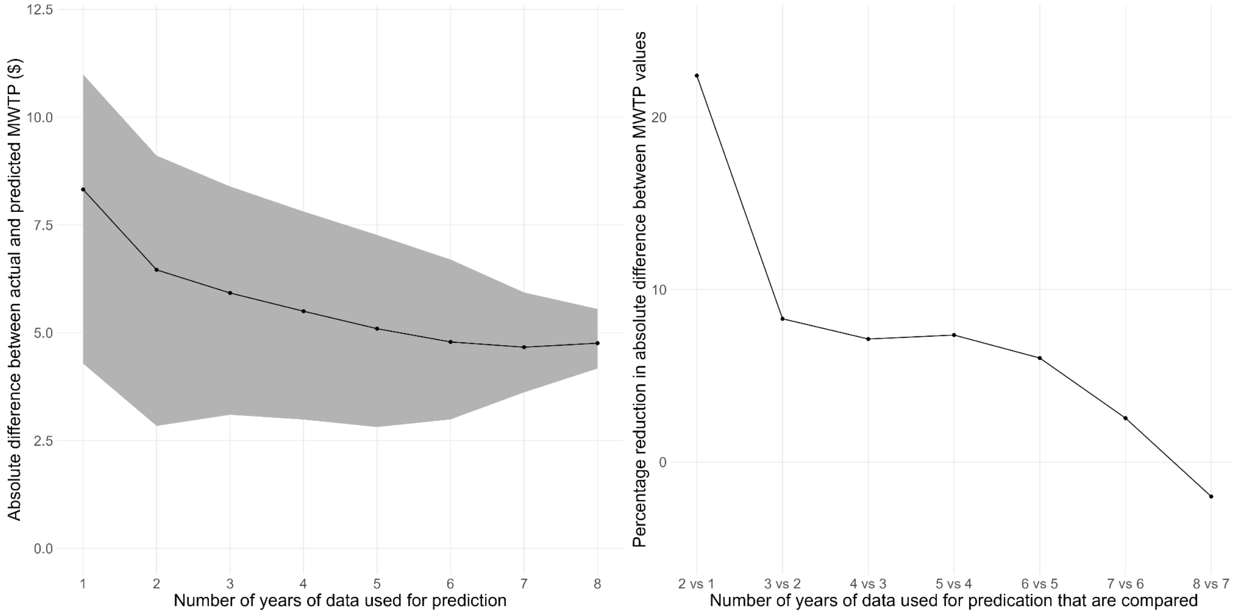
Notes: The figure shows pairwise MWTP differences in a given year to each of the other 'Test Years'. The dots represent average estimates and the capped vertical lines represent 95% confidence intervals. The MWTP estimates come from the ten year-specific multinomial logit site choice models.

Figure 6: Marginal willingness to pay (MWTP) to reduce one water quality advisory by week to and since the water quality advisory issuance date



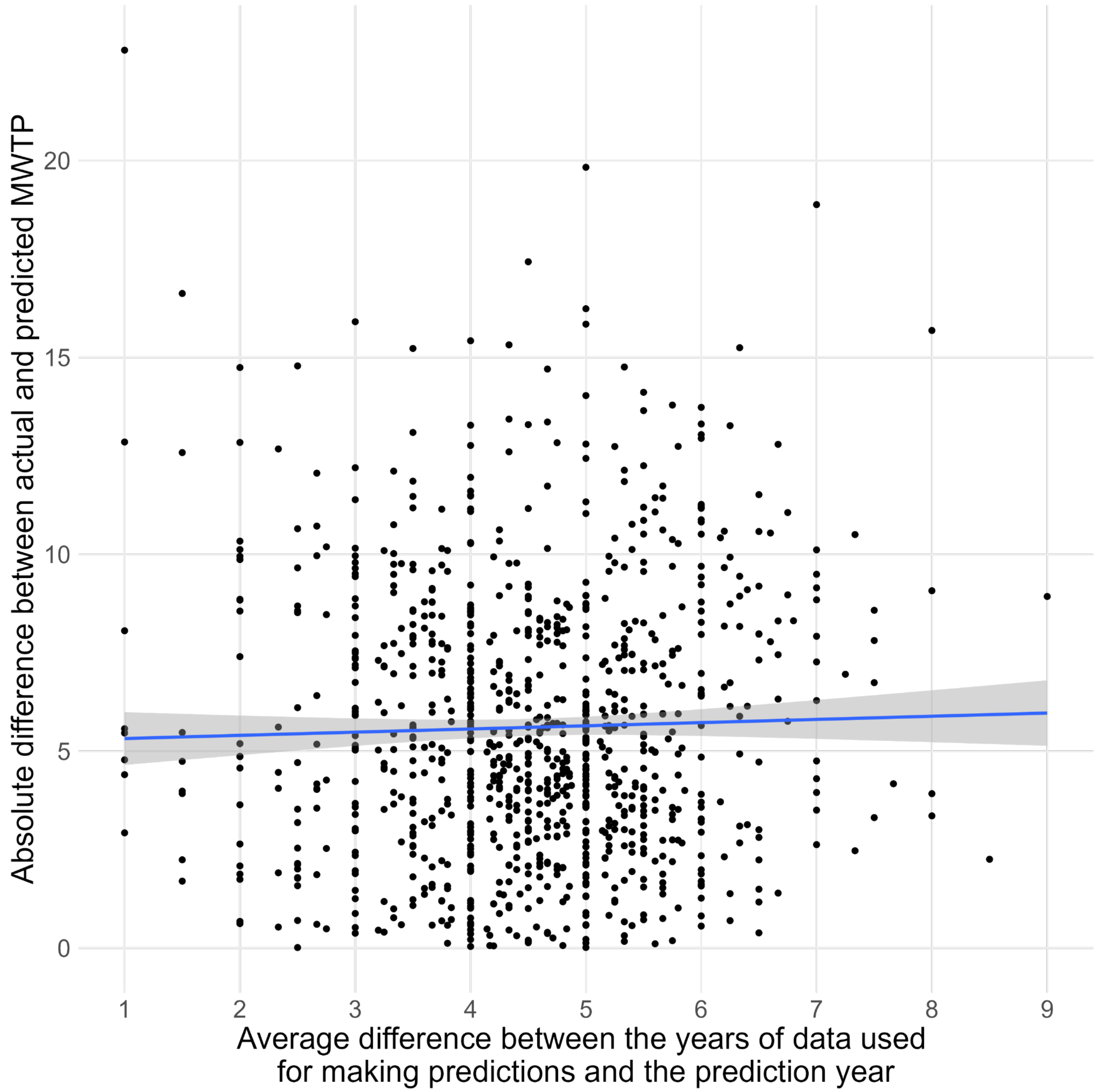
Notes: The MWTP estimates come from two multinomial logit site choice models with all years of data pooled. The model with year-specific ASCs is represented by the triangle points and solid lines, and the model with one set of ASCs for all years is represented by the circle points and dashed lines. The triangle/circle points indicate average estimates and the capped vertical lines are 95% confidence intervals. The reference level for the MWTP estimates is the MWTP value in the week (i.e. days -7 to -1) preceding the issuance of a water quality advisory.

Figure 7: Changes in the accuracy of predictions of marginal willingness to pay (MWTP) to reduce one water quality advisory upon using different numbers of years of data



Notes: The figure illustrates absolute differences between the actual and predicted MWTP values, where the latter are calculated based on 499 multi-year site choice models and 10 year-specific models. We omit the comparison of 9 years of data as that model is only used in a single prediction. In the left panel, the solid black lines represent the averages of the absolute differences across all MWTP comparisons for each number of years of data used for making predictions; the shaded area represents the interquartile ranges.

Figure 8: Changes in the accuracy of predictions of marginal willingness to pay (MWTP) to reduce one water quality advisory upon using years of data varying in the time distance to the year for which the value is predicted



Notes: The figure shows (i) the average absolute difference in years between the years of data used for making MWTP predictions and the prediction year on the x-axis, and (ii) the absolute difference between the predicted and actual MWTP values on the y-axis. For example, if the 2016 and 2018 years of data are used to predict the 2020 year MWTP, then the average difference in years is equal to |(2020-2016)+(2020-2018)|/2 = 3. The blue solid line plots the linear regression parameters of the difference in MWTP on the years difference and the shaded area is the 95% confidence interval.

References

Arrow, K., Solow, R., Portney, P. R., Leamer, E. E., Radner, R., Schuman, H. (1993). Report of the NOAA panel on contingent valuation. Federal register, 58(10), 4601-4614.

Bishop, R. C., Boyle, K. J. (2017). Reliability and validity in nonmarket valuation. In: Champ, P. A., Boyle, K. J., Brown, T. C. (Eds.) *A Primer on Nonmarket Valuatio*n (pp. 463-497). Dordrecht, the Netherlands: Springer.

Bishop, R. C., Boyle, K. J. (2019). Reliability and validity in nonmarket valuation. *Environmental and Resource Economics*, 72(2), 559-582.

Boudreaux, G., Lupi, F., Sohngen, B., & Xu, A. (2023). Measuring beachgoer preferences for avoiding harmful algal blooms and bacterial warnings. Ecological Economics, 204, 107653.

Brouwer, R. (2006). Do stated preference methods stand the test of time? A test of the stability of contingent values and models for health risks when facing an extreme event. Ecological Economics, 60(2), 399-406.

Brown, W.G., Nawas F. (1973). Impact of aggregation on the estimation of outdoor recreation demand functions. American Journal of Agricultural Economics, 55(2), 246-249.

Champ, P. A., Boyle, K. J., Brown, T. C. (2017). *A Primer on Nonmarket Valuation*. Dordrecht, the Netherlands: Springer.

Cooper, J., Loomis, J. (1990). Pooled time-series cross-section travel cost models: Testing whether recreation behavior is stable over time. *Leisure Sciences*, *12*(2), 161-171.

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

Gellman, J., Walls, M. A., & Wibbenmeyer, M. (2023). Welfare Losses from Wildfire Smoke: Evidence from Daily Outdoor Recreation Data (No. 23-31). Resources for the Future.

Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., & Tatlow, H. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour*, 5(4), 529-538.

Hellerstein, D. (1993). Intertemporal data and travel cost analysis. *Environmental and Resource Economics*, *3*, 193-207.

Kling, C. L., Phaneuf, D. J., Zhao, J. (2012). From Exxon to BP: has some number become better than no number? *Journal of Economic Perspectives*, 26(4), 3-26.

Jachimowicz, J.M., Duncan, S, Weber, E, Johnson. E (2019). When and Why Defaults Influence Decisions: A Meta-Analysis of Default Effects. *Behavioural Public Policy* 3 (2): 159–86.

Ji, Y., Keiser, D. A. (2016). Water-based Recreation and Water Quality Indices: A Revealed Preference Approach. *Economics Presentations, Posters and Proceedings*, 37.

Ji, Y., Keiser, D. A., Kling, C. L. (2020). Temporal reliability of welfare estimates from revealed preferences. *Journal of the Association of Environmental and Resource Economists*, 7(4), 659-686.

Johnston, R. J., Rolfe, J., Zawojska, E. (2018). Benefit transfer of environmental and resource values: progress, prospects and challenges. *International Review of Environmental and Resource Economics*, 12(2-3), 177-266.

Marcus, M., & Sant'Anna, P. H. (2021). The role of parallel trends in event study settings: An application to environmental economics. *Journal of the Association of Environmental and Resource Economists*, *8*(2), 235-275.

McFadden, D. (1974). The measurement of urban travel demand. Journal of Public Economics, 3, 303-328.

Lades, L., Zawojska, E., Johnston, R. J., Hanley, N., Delaney, L., and Czajkowski, M. (forthcoming). Anomalies or Expected Behaviors? Understanding Stated Preferences and Welfare Implications in Light of Contemporary Behavioral Economics. Review of Environmental Economics and Policy.

Liebe, U., Hendecagon, C., Beyer, H., & von Cramon-Taubadel, S. (2016). Context effects and the temporal stability of stated preferences. Social science research, 60, 135-147.

Lloyd-Smith, P, Becker, M. (2020). "The Economic Value of Camping using Administrative Data". *Journal of Agricultural and Resource Economics.* 45(3), 445-461.

Lupi, F., Phaneuf, D. J., von Haefen, R. H. (2020). Best practice in recreation demand analysis. *Review of Environmental Economics and Policy*, 14(2), 302-323.

Parsons, G. R. (2017). Travel cost models. In: Champ, P. A., Boyle, K. J., Brown, T. C. (Eds.) A Primer on Nonmarket Valuation (pp. 463-497). Dordrecht, the Netherlands: Springer.

Reiling, S. D., Boyle, K. J., Phillips, M. L., Anderson, M. W. (1990). Temporal reliability of contingent values. *Land Economics*, 66(2), 128-134.

Rolfe, J., Dyack, B. (2019). Testing temporal stability of recreation values. *Ecological Economics*, 159, 75-83.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34-55.

Stevens, J.B. (1969). Effects of nonprice variables upon participation in water-oriented outdoor recreation: Comment. American Journal of Agricultural Economics, 51(1), 192-193.

Trice, A. H., Wood, S. E. (1958). Measurement of recreation benefits. Land Economics, 34(3), 195-207.

Xie, L., Adamowicz, W. (2022). Temporal Reliability of Contingent Behavior Trip Data in Kuhn-Tucker Recreation Demand Models. *Land Economics*, 030521-0025R.

Yi, D., Herriges, J.A. (2017). Convergent validity and the time consistency of preferences: Evidence from the Iowa Lakes Recreation Demand Project. *Land Economics*, 93(2), 269-291.

Zandersen, M., Termansen, M., Jensen, F. S. (2007a). Evaluating approaches to predict recreation values of new forest sites. *Journal of Forest Economics*, 13(2-3), 103-128.

Zandersen, M., Termansen, M., Jensen, F. S. (2007b). Testing benefits transfer of forest recreation values over a twenty-year time horizon. *Land Economics*, 83(3), 412-440.

Appendix A. Summary of the existing empirical recreation demand studies assessing temporal stability of welfare estimates

	Trips to …	Years studied	Number of individuals/households/zones per year	Number of trips per year	Data source
Cooper and Loomis (1990)	Feather River in California, US	1981-1985	zonal; 342 observations (57 counties times 6 destination sites)	6,478-13,407	Survey
Hellerstein (1993)	Boundary Waters Canoe Area in Minnesota, US	1980-1986	zonal; 1,396 counties	Not reported	Administrative
Zandersen et al. (2007a, 2007b)	Forests in selected districts in Denmark	1977, 1997	6,580-6,987	Not reported	Survey
Yi and Herriges (2017)	Big Creek Lake and Saylorville Lake in Iowa, US	2004-2005	2,150	14,534-15,179[a]	Survey
Rolfe and Dyack (2019)	Estuarine region of the Coorong, Australia	2006, 2013	778-783	2,349-3,306[b]	Survey
Ji et al. (2020)	Lakes in Iowa, US	2002-2005, 2009	977	5,862[c]	Survey
Current study	Campgrounds in Alberta, Canada	2013-2022	55,947-78,275	80,240-127,812	Administrative

Notes: [a] These numbers are calculated based on information on the average numbers of trips reported in Table 1 in Yi and Herriges (2017). [b] These numbers are calculated based on information presented in Table 2 of Rolfe and Dyack (2019). Specifically, 783 people in 2006 took an average of 3 trips for 2,349 total and 778 people in 2013 took 4.25 trips per year for 3,306 total. [c] This number is calculated based on information in Ji et al. (2020, p. 665) that the studied households took on average about 6 trips per year.

Appendix B. Complete set of parameter estimates for each of the year-specific multinomial logit site choice models

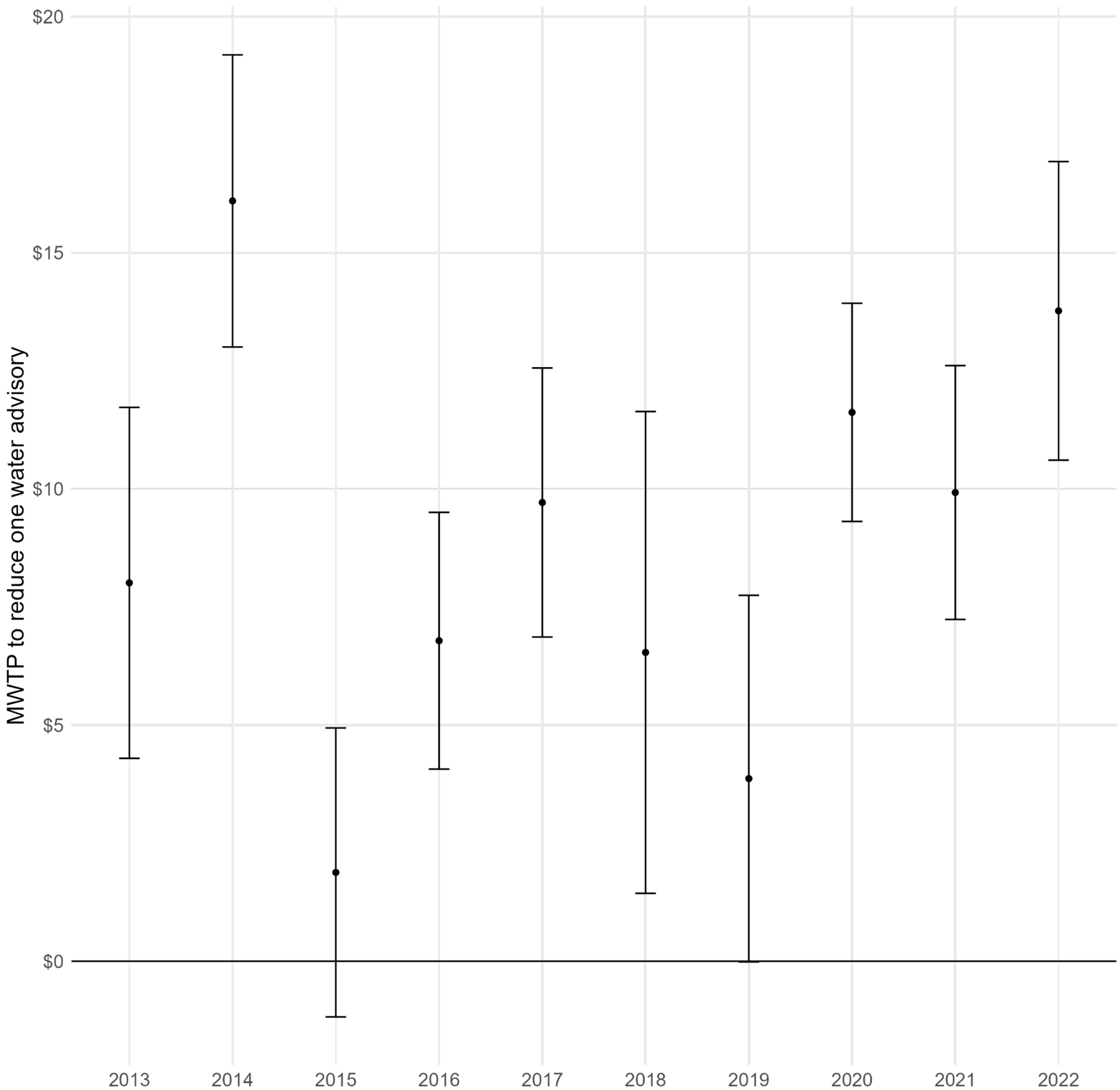
	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Aspen Beach × Brewer's					0 (fixed)					
Aspen Beach × Lakeview	-0.051	-0.043	-0.012	-0.043	-0.082	-0.106	-0.048	-0.097	-0.099	-0.169
	(0.026)	(0.026)	(0.025)	(0.026)	(0.025)	(0.026)	(0.026)	(0.026)	(0.024)	(0.026)
Beauvais Lake	-0.520	-0.502	-0.410	-0.428	-0.545	-0.455	-0.307	0.285	0.235	0.339
	(0.048)	(0.048)	(0.044)	(0.045)	(0.044)	(0.044)	(0.044)	(0.032)	(0.031)	(0.034)
Beaver Lake	-0.540	-0.670	-0.665	-0.902	-0.693	-0.703	-0.652	-0.481	-0.409	-0.673
	(0.045)	(0.044)	(0.044)	(0.047)	(0.043)	(0.045)	(0.046)	(0.039)	(0.038)	(0.045)
Boulton Creek	-0.272	0.032	0.060	0.076	0.087	0.147	0.406	0.472	0.287	0.392
	(0.035)	(0.031)	(0.030)	(0.031)	(0.028)	(0.029)	(0.028)	(0.026)	(0.026)	(0.027)
Bow Valley	-0.019	0.075	0.032	0.083	-0.036	0.067	0.162	0.178	-0.002	0.047
	(0.029)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)	(0.025)	(0.025)	(0.027)
Brazeau Res. West Canal	-2.455	-2.842	-2.394	-2.359	-2.383	-2.444	-2.420	-2.074	-1.964	-2.041
	(0.082)	(0.094)	(0.076)	(0.077)	(0.075)	(0.078)	(0.076)	(0.063)	(0.061)	(0.068)
Brazeau Reservoir	-1.673	-1.527	-1.550	-1.495	-1.577	-1.464	-1.494	-0.995	-1.025	-1.106
	(0.056)	(0.053)	(0.054)	(0.050)	(0.051)	(0.051)	(0.052)	(0.041)	(0.040)	(0.054)
Chain Lakes	-1.509	-1.472	-1.424	-1.384	-1.431	-1.356	-1.212	-0.536	-0.577	-0.703
	(0.048)	(0.045)	(0.044)	(0.044)	(0.044)	(0.044)	(0.045)	(0.032)	(0.030)	(0.035)
Cold Lake	0.182	0.369	0.359	0.393	0.378	0.367	0.412	0.570	0.688	0.629
	(0.042)	(0.038)	(0.037)	(0.037)	(0.036)	(0.038)	(0.040)	(0.033)	(0.032)	(0.036)
Crimson Lake × Twin Lakes	-3.033	-2.904	-2.893	-2.666	-2.485	-2.574	-5.495	-1.428	-1.418	-1.402
	(0.095)	(0.092)	(0.096)	(0.086)	(0.071)	(0.076)	(0.317)	(0.043)	(0.041)	(0.047)
Cross Lake	-0.165	-0.209	-0.167	-0.174	-0.313	-0.309	-0.278	-0.258	-0.243	-0.373
	(0.034)	(0.034)	(0.032)	(0.034)	(0.033)	(0.036)	(0.034)	(0.032)	(0.033)	(0.039)
Cypress Hills × Battle Creek	-2.024	-1.863	-1.909	-1.927	-2.155	-2.023	-1.909	-1.078	-1.250	-1.619
	(0.137)	(0.121)	(0.127)	(0.129)	(0.133)	(0.137)	(0.131)	(0.067)	(0.077)	(0.102)
Cypress Hills × Beaver Creek	-0.626	-0.895	-0.811	-0.698	-0.831	-0.765	-0.631	-0.857	-0.874	-0.752
	(0.081)	(0.091)	(0.073)	(0.065)	(0.073)	(0.070)	(0.069)	(0.071)	(0.066)	(0.070)
Cypress Hills × Elkwater	0.456	0.566	0.624	0.602	0.467	0.612	0.611	0.402	0.565	0.594
	(0.040)	(0.039)	(0.038)	(0.038)	(0.038)	(0.039)	(0.040)	(0.038)	(0.035)	(0.039)
Cypress Hills × Ferguson Hill	-0.548	-0.496	-0.595	-0.447	-0.548	-0.379	-0.405	-0.104	-0.249	-0.594
	(0.056)	(0.054)	(0.054)	(0.052)	(0.050)	(0.058)	(0.052)	(0.040)	(0.042)	(0.055)
Cypress Hills × Firerock	0.666	0.721	0.853	0.815	0.656	0.747	0.789	0.954	0.874	0.864
	(0.036)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.036)	(0.031)	(0.030)	(0.034)
Cypress Hills × Lakeview	-1.342	-1.265	-1.171	-1.227	-1.218	-0.975	-0.993	-1.186	-1.195	-1.017
	(0.085)	(0.083)	(0.079)	(0.083)	(0.082)	(0.071)	(0.076)	(0.075)	(0.072)	(0.079)
Cypress Hills × Lodgepole	-0.608	-0.609	-0.765	-0.538	-0.702	-0.527	-0.563	-0.121	-0.366	-0.713
	(0.058)	(0.058)	(0.060)	(0.053)	(0.055)	(0.063)	(0.057)	(0.041)	(0.045)	(0.059)
Cypress Hills × Old Baldy	-0.793	-0.790	-0.678	-0.730	-0.805	-0.662	-0.492	-0.916	-0.896	-0.698
	(0.068)	(0.073)	(0.066)	(0.071)	(0.072)	(0.065)	(0.064)	(0.067)	(0.063)	(0.067)
Cypress Hills × Reesor Lake	-0.678	-0.702	-0.571	-0.563	-0.714	-0.492	-0.506	-0.109	-0.265	-0.319
	(0.064)	(0.064)	(0.060)	(0.061)	(0.063)	(0.058)	(0.063)	(0.046)	(0.047)	(0.056)
Cypress Hills × Spruce Coulee	-1.980	-1.713	-1.686	-1.637	-1.805	-1.618	-1.450	-1.324	-1.466	-1.539
	(0.100)	(0.092)	(0.093)	(0.089)	(0.090)	(0.086)	(0.085)	(0.070)	(0.075)	(0.087)
Dillberry Lake	-1.161	-1.034	-1.022	-0.921	-0.976	-0.927	-0.850	-0.334	-0.302	-0.517

	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
	(0.070)	(0.065)	(0.062)	(0.059)	(0.059)	(0.057)	(0.059)	(0.044)	(0.042)	(0.052)
Dinosaur	0.441	0.495	0.615	0.638	0.496	0.537	0.674	0.703	0.573	0.483
	(0.032)	(0.031)	(0.029)	(0.029)	(0.029)	(0.030)	(0.029)	(0.027)	(0.027)	(0.030)
Dunvegan	-2.196	-1.958	-2.096	-1.984	-1.887	-1.958	-2.009	-0.143	-0.102	-0.468
	(0.096)	(0.088)	(0.090)	(0.095)	(0.094)	(0.095)	(0.107)	(0.046)	(0.046)	(0.054)
Elkwood	-0.926	-0.003	0.085	0.088	0.016	0.070	0.161	0.136	-0.029	0.124
	(0.044)	(0.032)	(0.031)	(0.030)	(0.030)	(0.030)	(0.030)	(0.028)	(0.027)	(0.030)
Etherington Creek	-1.709	-1.374	-1.239	-1.138	-1.270	-1.034	-1.043	-0.768	-1.033	-1.040
	(0.116)	(0.058)	(0.049)	(0.049)	(0.046)	(0.044)	(0.047)	(0.036)	(0.038)	(0.044)
Fish Lake	-0.746	-0.658	-0.641	-0.622	-0.664	-0.505	-0.414	0.525	0.450	0.474
	(0.051)	(0.050)	(0.049)	(0.048)	(0.046)	(0.046)	(0.049)	(0.029)	(0.029)	(0.031)
Franchere Bay	-0.536	-0.544	-0.673	-0.795	-0.786	-0.835	-0.880	-0.358	-0.454	-1.117
	(0.042)	(0.043)	(0.043)	(0.048)	(0.047)	(0.050)	(0.049)	(0.040)	(0.039)	(0.055)
Garner Lake	-0.937	-0.901	-1.003	-1.012	-0.995	-1.037	-1.031	-0.792	-0.829	-0.987
	(0.045)	(0.043)	(0.043)	(0.046)	(0.045)	(0.045)	(0.046)	(0.041)	(0.038)	(0.048)
Hilliard's Bay	0.114	0.665	0.672	0.520	0.608	0.499	0.450	0.851	0.889	0.422
	(0.091)	(0.040)	(0.039)	(0.042)	(0.039)	(0.039)	(0.039)	(0.033)	(0.033)	(0.042)
Jarvis Bay	-0.318	-0.258	-0.253	-0.285	-0.393	-0.320	-0.171	-0.079	-0.026	0.009
	(0.027)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.025)	(0.024)	(0.025)
Kinbrook Island	0.433	0.454	0.539	0.557	0.618	0.656	0.761	0.687	0.763	0.842
	(0.031)	(0.030)	(0.029)	(0.029)	(0.028)	(0.028)	(0.028)	(0.027)	(0.025)	(0.027)
Lac Des Arcs	-2.203	-2.038	-1.994	-1.930	-1.856	-1.903	-1.755	-1.607	-1.896	-1.847
	(0.055)	(0.050)	(0.049)	(0.051)	(0.050)	(0.048)	(0.047)	(0.039)	(0.042)	(0.048)
Lesser Slave Lake × Marten River	0.567	0.501	0.577	0.539	0.594	0.500	0.262	0.637	0.699	0.610
	(0.037)	(0.037)	(0.035)	(0.038)	(0.034)	(0.036)	(0.038)	(0.033)	(0.031)	(0.036)
Little Bow	0.409	0.444	0.443	0.466	0.324	0.347	0.432	0.463	0.289	0.359
	(0.029)	(0.029)	(0.028)	(0.028)	(0.027)	(0.028)	(0.028)	(0.026)	(0.026)	(0.028)
Little Elbow	-2.580	-1.902	-1.813	-1.772	-1.786	-1.690	-1.543	-0.520	-0.801	-0.763
	(0.062)	(0.049)	(0.045)	(0.046)	(0.044)	(0.044)	(0.043)	(0.028)	(0.029)	(0.032)
Long Lake	0.246	0.258	0.202	0.156	0.131	0.191	0.196	0.245	0.228	0.172
	(0.028)	(0.030)	(0.027)	(0.031)	(0.027)	(0.028)	(0.028)	(0.026)	(0.026)	(0.028)
McLean Creek	-0.792	-0.684	-0.677	-0.654	-0.771	-0.638	-0.458	-0.428	-0.775	-0.673
	(0.036)	(0.035)	(0.035)	(0.034)	(0.033)	(0.033)	(0.032)	(0.029)	(0.032)	(0.033)
Miquelon Lake	-0.238	-0.285	-0.356	-0.384	-0.307	-0.390	-0.293	-0.367	-0.277	-0.400
	(0.025)	(0.025)	(0.025)	(0.025)	(0.027)	(0.026)	(0.026)	(0.026)	(0.024)	(0.026)
Moose Lake	-1.502	-1.627	-1.735	-1.722	-1.861	-2.155	-1.963	-0.920	-0.900	-1.584
	(0.068)	(0.068)	(0.069)	(0.076)	(0.079)	(0.087)	(0.078)	(0.051)	(0.048)	(0.069)
Park Lake	-0.401	-0.352	-0.373	-0.315	-0.441	-0.343	-0.233	-0.233	-0.313	-0.204
	(0.039)	(0.037)	(0.037)	(0.037)	(0.037)	(0.036)	(0.037)	(0.034)	(0.034)	(0.036)
Pelican Point Park	-1.453	-1.495	-1.632	-1.535	-1.599	-1.568	-1.492	-1.367	-1.228	-1.344
	(0.050)	(0.049)	(0.051)	(0.051)	(0.049)	(0.049)	(0.051)	(0.043)	(0.040)	(0.044)
Pembina River	-0.390	-0.384	-0.440	-0.407	-0.435	-0.410	-0.491	-0.367	-0.298	-0.388
	(0.029)	(0.028)	(0.027)	(0.028)	(0.028)	(0.028)	(0.029)	(0.027)	(0.026)	(0.028)
Pigeon Lake	-0.274	-0.091	-0.179	-0.126	-0.131	-0.106	-0.033	0.013	-0.029	-0.131
	(0.027)	(0.029)	(0.029)	(0.027)	(0.025)	(0.025)	(0.026)	(0.025)	(0.023)	(0.025)
Pigeon Lake × Zeiner	-1.176	-1.134	-1.223	-1.201	-1.160	-1.150	-1.058	-1.024	-1.044	-1.284
	(0.038)	(0.037)	(0.037)	(0.037)	(0.035)	(0.035)	(0.035)	(0.033)	(0.031)	(0.037)
Red Lodge	-0.863	-0.948	-0.933	-0.924	-1.009	-0.934	-0.808	-0.985	-0.967	-1.063
	(0.031)	(0.032)	(0.031)	(0.033)	(0.032)	(0.031)	(0.032)	(0.032)	(0.030)	(0.032)

	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Saskatoon Island	-1.567	-1.600	-1.661	-0.450	-0.308	-0.430	-0.492	-0.132	-0.116	-0.288
	(0.067)	(0.062)	(0.065)	(0.051)	(0.049)	(0.051)	(0.051)	(0.043)	(0.044)	(0.052)
Sir Winston Churchill	0.015	-0.089	-0.002	-0.044	0.030	-0.073	0.004	-0.015	0.200	0.226
	(0.038)	(0.038)	(0.038)	(0.039)	(0.036)	(0.038)	(0.040)	(0.036)	(0.033)	(0.040)
Tillebrook	-0.609	-0.574	-0.572	-0.559	-0.625	-0.666	-0.556	-0.558	-0.656	-0.836
	(0.039)	(0.040)	(0.039)	(0.039)	(0.039)	(0.041)	(0.040)	(0.037)	(0.037)	(0.044)
Vermilion	-0.686	-0.638	-0.746	-0.705	-0.627	-0.700	-0.647	-0.594	-0.548	-0.819
	(0.043)	(0.043)	(0.041)	(0.040)	(0.039)	(0.039)	(0.038)	(0.036)	(0.035)	(0.042)
Wabamun Lake	-0.120	-0.107	-0.137	-0.143	-0.122	-0.051	-0.001	0.140	0.117	0.072
	(0.026)	(0.025)	(0.025)	(0.025)	(0.024)	(0.024)	(0.025)	(0.023)	(0.022)	(0.024)
Whitney Lakes × Ross Lake	-0.002	0.200	-0.068	-0.106	-0.197	-0.086	0.037	0.337	0.398	0.238
	(0.043)	(0.041)	(0.041)	(0.041)	(0.040)	(0.039)	(0.048)	(0.034)	(0.032)	(0.037)
Whitney Lakes × Whitney Lake	-1.767	-1.236	-1.658	-1.636	-1.576	-1.752	-1.749	-1.063	-1.029	-1.404
	(0.082)	(0.070)	(0.081)	(0.076)	(0.080)	(0.091)	(0.111)	(0.054)	(0.053)	(0.072)
William A. Switzer × Gregg Lake	0.780	0.783	0.920	0.925	1.009	1.019	1.108	1.368	1.221	1.218
	(0.037)	(0.037)	(0.034)	(0.033)	(0.031)	(0.032)	(0.032)	(0.027)	(0.028)	(0.032)
William A. Switzer × Jarvis Lake	-0.815	-0.929	-0.823	-0.842	-0.784	-0.660	-0.596	-0.362	-0.444	-0.424
	(0.065)	(0.065)	(0.062)	(0.060)	(0.058)	(0.054)	(0.058)	(0.046)	(0.048)	(0.054)
Writing-on-Stone	0.960	1.096	0.906	0.944	0.746	0.760	0.886	0.790	0.831	0.847
	(0.036)	(0.037)	(0.036)	(0.036)	(0.035)	(0.035)	(0.035)	(0.032)	(0.031)	(0.034)
Young's Point	-0.178	-0.242	-0.192	-0.144	0.047	-0.034	-0.095	0.214	0.104	-0.189
	(0.046)	(0.046)	(0.043)	(0.039)	(0.037)	(0.039)	(0.040)	(0.036)	(0.038)	(0.044)
Water quality advisory	-0.104	-0.213	-0.027	-0.091	-0.128	-0.082	-0.049	-0.148	-0.129	-0.200
	(0.025)	(0.021)	(0.022)	(0.019)	(0.019)	(0.033)	(0.025)	(0.015)	(0.018)	(0.023)
Travel cost ($00s)	-0.797	-0.822	-0.851	-0.815	-0.801	-0.784	-0.796	-0.763	-0.784	-0.909
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
Log-likelihood value at convergence	-244,860	-267,640	-287,693	-278,507	-292,015	-284,236	-285,233	-420,456	-406,363	-314,673
Number of individuals	55,947	60,063	63,831	61,953	64,760	63,074	62,495	78,275	77,760	66,973
Number of observations	80,240	87,648	93,875	89,612	93,584	91,067	91,140	127,812	123,369	99,487

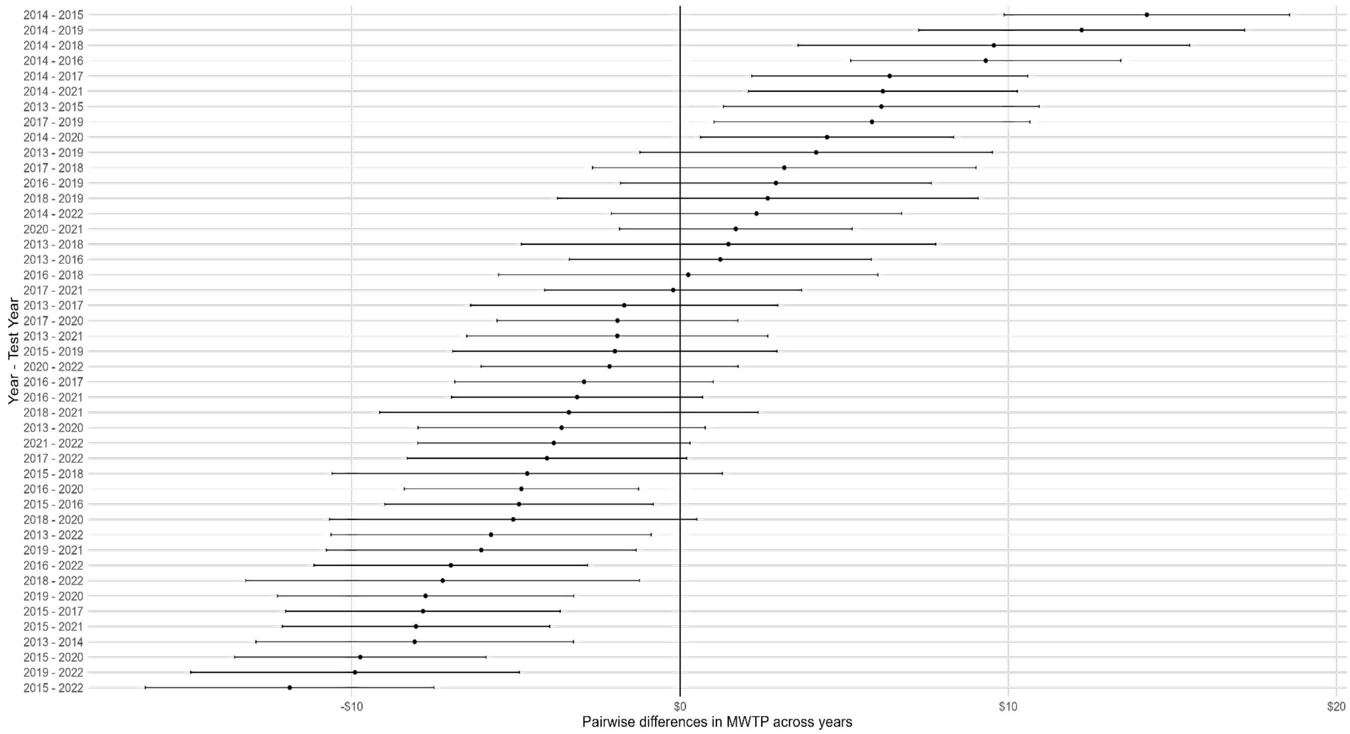
Appendix C. Sensitivity analysis based on 1/3 of the wage rate for the value of time

Figure C1: Marginal willingness to pay (MWTP) per trip for removing one water quality advisory using alternative value of time assumption.



Notes: The figure reports the average MWTP per trip to reduce one water quality advisory at a campground for each of the ten year-specific multinomial logit site choice models using 1/3 of the wage rate as the value of time. The dots represent average estimates and the capped vertical lines represent 95% confidence intervals of the MWTP values.

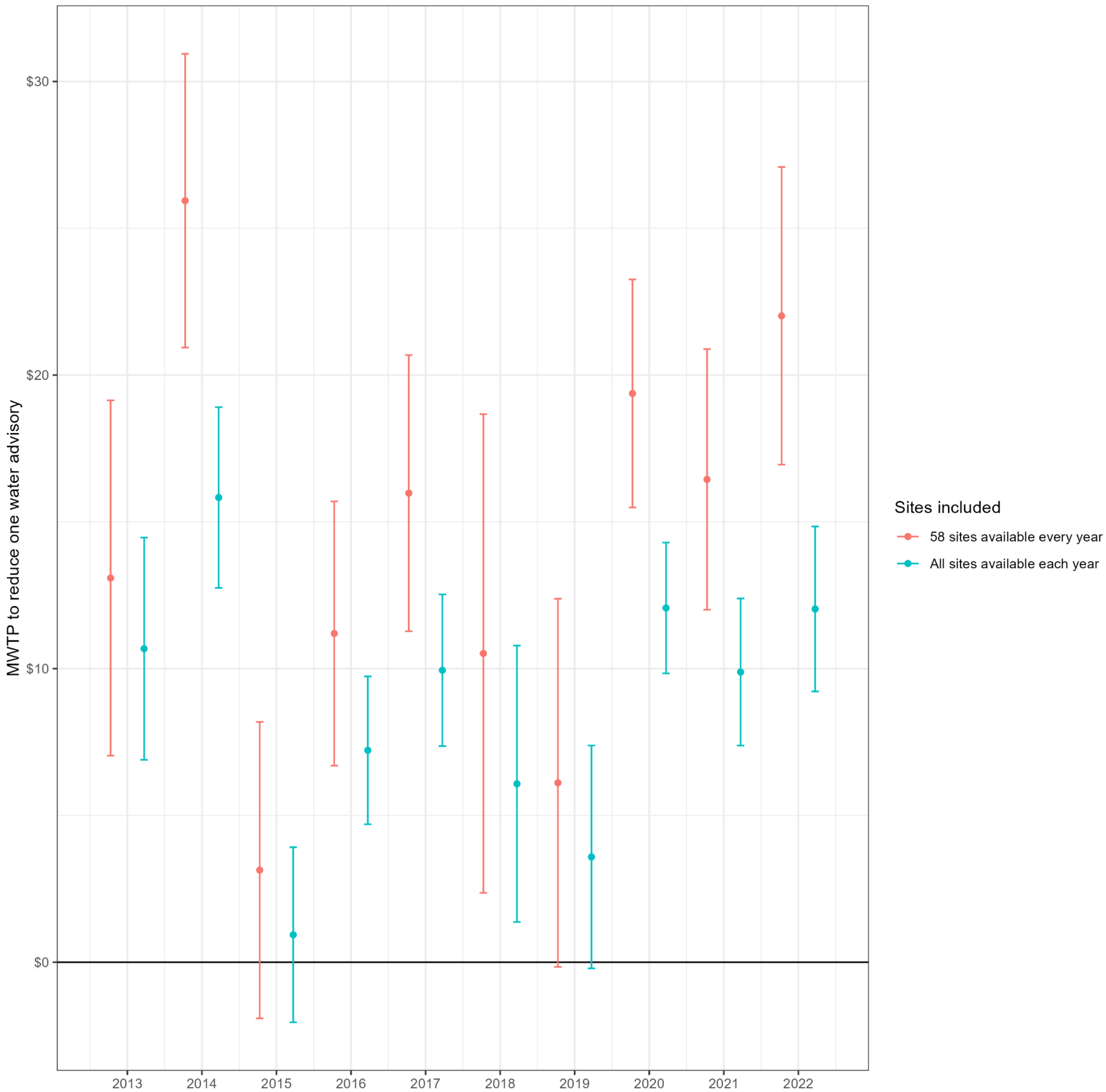
Figure C2: Difference in marginal willingness to pay (MWTP) to reduce one water quality advisory across years using alternative value of time assumption.



Notes: The figure shows pairwise MWTP differences in a given year to each of the other "Test Years" and ranks the differences from lowest to highest. The dots represent average estimates and the capped vertical lines represent 95% confidence intervals. The MWTP estimates come from the ten year-specific multinomial logit site choice models using 1/3 of the wage rate as the value of time.

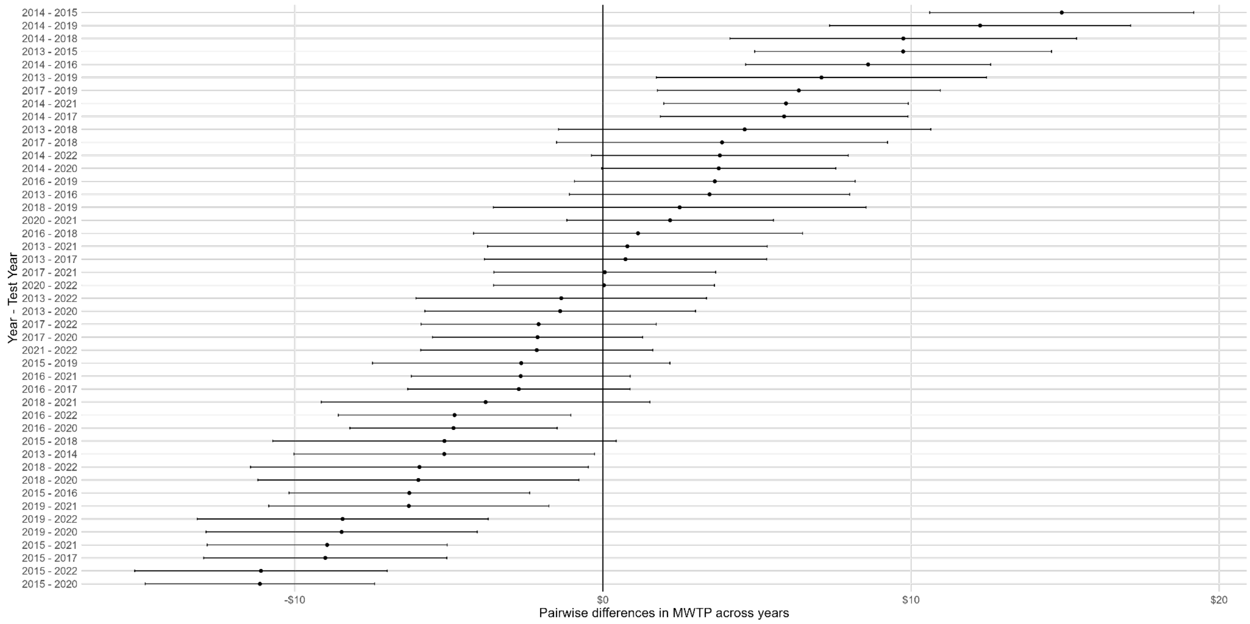
Appendix D. Sensitivity analysis based on all available campgrounds

Figure D1: Marginal willingness to pay (MWTP) per trip for removing one water quality advisory using all available sites each year.



Notes: The figure reports the average MWTP per trip to reduce one water quality advisory at a campground for each of the ten year-specific multinomial logit site choice models estimated with the 58 sites available in all 10 years (base model) and models estimated with all available sites each year. The number of sites for the all sites model ranges from 62 to 111 between the years. The dots represent average estimates and the capped vertical lines represent 95% confidence intervals of the MWTP values.

Figure D2: Difference in marginal willingness to pay (MWTP) to reduce one water quality advisory across years all available sites each year.



Notes: The figure shows pairwise MWTP differences in a given year to each of the other “Test Years” and ranks the differences from lowest to highest. The dots represent average estimates and the capped vertical lines represent 95% confidence intervals. The MWTP estimates come from the ten year-specific multinomial logit site choice models using all available sites each year.

University of Warsaw
Faculty of Economic Sciences
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl