Measuring the Value of Water Quality Improvements for Recreation Use on an Urban River in the USA: A Travel Cost Contingent Behavior Approach

Key Points

- The per trip value for contact use of the Brandywine Creek is $42 and for noncontact use is $29
- The per trip value for water quality improvements ranged from $4 to $63 for contact use and from $4 to $29 for non-contact use
- The larger number of non-contact users led to higher aggregate benefits for water quality improvements to this group
- Even though we measure sizable benefits, rough costs estimates appear to make it difficult for incremental improvements to exceed costs

Abstract

We estimate the value of water quality improvements for recreational activities on and near the Brandywine Creek in Delaware. We divide water-based recreational activities into two groups: water-contact and non-water-contact. We then consider the behavioral change of the recreationists in each group when faced with water quality improvements posed in a stated preference (contingent behavior) survey. We use five common physical attributes of water quality in the contingent behavior analysis: water clarity, catch rate of fish, safety of eating fish, safety for swimming, and ecological health. The survey was constructed in such a way that each of these attributes of water quality are valued. The per trip value for water quality improvements ranged from $4 to $63 for contact use and from $4 to $29 for non-contact use depending on which and how many attributes of water quality improve. Even though we measure sizable benefits, rough costs estimates appear to make it difficult for incremental improvements to exceed costs unless attribute improvement is widespread.

1 Background & Introduction

The Brandywine Creek is a prominent river in northern Delaware flowing through the city of Wilmington (see Figure 1). Before reaching the city, the Brandywine is in a mostly rural setting of woods, farms, and small towns. When it enters the city, it becomes distinctly urban but maintains a scenic stream-like quality. Given its natural beauty, proximity to people, and numerous access points, the Brandywine is an important recreational resource with water-contact activities including swimming, kayaking, tubing, fishing, wading, and non-water-contact activities like walking, running, picnicking, reading, and biking nearby where there are many walkways, paths and parks.

Like many urban rivers, the Brandywine Creek saw large declines in its water quality in the last century due to point and nonpoint source pollutants and legacy toxics (PCBs, dioxins, furans) from nearby industries. Regulatory actions and changes in some agricultural practices have resulted in improved conditions, but water quality issues especially for recreational use remain a concern. For example, signage along the Brandywine warns against swimming following storms due to run off and fish consumption advisories still exist. There is discussion now of a major revival of the Brandywine –

improving its clarity, making it safe for swimming, removing toxins, improving fishing quality, and the possible removal of dams to restore the ecological conditions unknown for decades.

Given this interest in the Brandywine and the likelihood that other rivers in similar urban settings may be targets for the same type of initiatives, we decided to measure the welfare effects of water quality improvements on the Brandywine. Any such efforts to make improvements will include costs and good policy must weigh the costs against the benefits. Our focus is on recreation since a portion of the benefits. Our method is a travel cost model with contingent behavior data along the lines first suggested by McConnell (1986) and Ward (1987). The contingent behavior data are responses to questions about how people would change their trip taking behavior in the presence of hypothetical improvements in water quality. The contingent behavior response data essentially "shifts" a single-site demand function and allows us to estimate the benefits of water quality improvements in terms of a change in consumer surplus.

We opted for a contingent-behavior-based approach for two reasons. First, the observed water quality on the Brandywine and its substitute recreation sites in the state showed insufficient variation to model a behavioral response using revealed preference data alone. Second, the changes in water quality we wanted to consider were not observable. We wanted to consider some large transformative-like changes in quality and the current conditions on other rivers and lakes in the state were not at these levels. Using contingent behavior data allows us to generate the variation we need for estimation and the sizes we are interested in considering.

Our paper is outlined as follows: study design, literature, sample & survey, data, econometric model, results, and conclusions.

2 Study Design

We use a single-site recreation demand model in our application. Our focus is on the benefits of water quality improvements to Delaware residents. We use a contingent behavior approach combining revealed and stated trip data using response data from an internet-based survey. The revealed trip data are a respondent's total number of trips to the Brandywine in the 2021. The contingent behavior or stated trip data are a respondent's reported trips in response to hypothetical questions about changes in water quality on the Brandywine.

The first contingent behavior question asks about trip change behavior for a large improvement in water quality over five measures of quality – clarity, safety for swimming, safety for eating fish, quality of fishing, and overall ecological health. The visual representation used to show the first water quality improvement in the survey is shown in Figure 2. Then, using the revealed and stated trip data, we estimate two demand functions – one without the improvement in water quality (revealed data) and one with the improvement (stated data)[1]. The welfare change for a given respondent due the improvement in water quality is measured as the area between the two estimated trip demand functions above the respondent's trip cost. This is the classic measure of a change in consumer surplus discussed in the earliest literature (McConnell, 1986) on measuring the benefits of water quality improvements in recreation demand models and is based on the assumption of weak complementarity in trips and site quality, which holds that a person receives use value related to the quality of a site only if he or she visits the site (Maler, 1974 ; Flores, 2017, 39; and Bockstael and McConnell, 2007, 47). Mathematically, the benefit for the water quality improvement for a given individual is

(1) $$\Delta W = \int_{tc^0}^{tc^{**}} x^1(tc, q^1, z) - \int_{tc^0}^{tc^*} x^0(tc, q^0, z),$$

[1] It is common in the literature (Whitehead et al., 2000) to consider revealed preference data in the current year and then stated preference data in a future year with and without an improvement, giving three data points for estimation. It is argued that this using the revealed (without improvement) and stated (with improvement) preference data could lead to overestimates of trip change. Our pretest and focus groups suggested otherwise. Asking for "trip change" due to the improvement from the existing trip level of trips seemed to simplify to task (only one response task was required) and add realism to the contingent choice. Also, using ranges for the change was easier to report, closer the degree of accuracy people could report, and was uniformly preferred in our pretest and focus group set up. The pretests and focus groups were made of graduate students, administrative staff, friends, and colleagues.

where $x^0(.)$ and $x^1(.)$ are the estimated demand functions without and with water quality improvements, q^0 and q^1 are the water quality levels without and with improvements, tc is the trip cost to the Brandywine, and z is a vector of other relevant shifters (like income and travel cost to substitute sites). tc^0 is the travel cost to the Brandywine for a person, and tc^* and tc^{**} the travel cost at which the person would no longer take trips. For an exact measure of surplus Hicksian demands in equation (1) are used (Bockstael and McConnell, 2007, 47). In our application, we use Marshallian demand as an approximation.

The second contingent behavior question asks about trip change behavior for a smaller improvement in water quality. The level of water quality change varies across respondents randomly but is always less than the improvement in the first contingent behavior question. With the smaller change each respondent sees one of 14 possible changes in water quality. These include cases where only one of the four water quality parameters change (e.g., only clarity is improved), where two parameters improve, and where three parameters change. In this way we hope to see the extent to which different features or water quality matter to respondents. Figure 3 is an example of how the visual appeared to a respondent for the second smaller contingent scenario. In our analysis, this gives us a second "with" water quality improvement demand function with a smaller area than the original shift. Our hypothesis is that the larger the small improvement is the larger the measured benefits, so the degree of shift is expected to vary. We also expected it to vary depending on the type of improvement in water quality (e.g., clarity versus safety for swimming).

The 14 possible combinations of changes in water quality features were constructed using a full factorial design. In this way orthogonality among the attributes is realized and we can isolate the effects of the individual attributes on trip taking behavior. We excluded ecological health (our fifth attribute of water quality) in the factorial design. Instead, we set an outcome for its level to be consistent with the changes realized over the other attributes. So, for example, if only one attribute was improved, the change in ecological health would be small. If all or most of the attributes were improved, the change would be large. The "forced" outcome insured realism.

The survey is described in detail in the upcoming section. We sampled the entire state of Delaware. The data are off-site and include participants and non-participants. We use a count data structure to account for the integer and truncated nature of our data. We also consider two demand models throughout the analysis – contact and non-contact recreation. Contact recreation includes any activity where there is 'contact' with the water, like swimming, fishing, boating, wading, etc. Non-contact includes activities that take place near, but not in the water, like walking, picnicking, birdwatching, etc. Finally, the welfare estimates consider several scenarios with different degrees of water quality improvement.

3 Travel Cost Contingent Behavior Literature Applied to Water Quality

Economists have been supplementing travel cost data with contingent behavior data (TCM-CB) for over three decades. The primary reason is to create response data beyond the domain of observation in the market, i.e., extend observable site quality or trip cost beyond that realized with revealed preference data. This may be used to generate variation in a critical policy variable for estimation and/or to extend the measure of the variable into a policy-relevant range. Since contingent data are exogenous by construction, they are uncorrelated with observable covariates, which can also be an argument for introducing contingent behavior data where collinearity may be present in observable data.

The earliest TCM-CB applications were by McConnell (1986) and Ward (1987). Both have a theoretical foundation in Maler's (1974) concept of weak complementarity and the notion that the area between two demand curves differentiated by site quality is a measure of consumer surplus for that quality difference. McConnell (1986) considered the impact of PCB contamination on beach use in New Bedford Harbor Massachusetts, and Ward (1987) studied the effect of streamflow on sport fishing and boating (kayaking, rafting, and canoeing) on the Rio Chama in New Mexico.

These two articles along with an increase of interest generally in combining revealed and stated preference data beginning with Cameron (1992) and Adamowicz (1994)) gave rise to a large stream of articles. The applications cover fishing, beach use, hiking, biking, rock climbing, skiing, diving, and other forms of recreation, and consider a wide range of resource changes – improvements in water quality, changes in water levels on reservoirs, wider beaches, new skiing areas, changes in conditions due to the removal of dams, introduction of offshore wind turbines, lost access, changes in coral reef quality, and more. All but few are single-site models. The basic methodology has been more or less unchanged in the ensuing decades, although the data gathering approaches, survey techniques, and econometric analysis have all improved along with refinements in study designs and research questions asked. Nearly all of the modern applications use some form of a count data model with Negative Binomial and Random Parameters Poisson in their various forms being the most common. Many of the analyses in the TCM-CB literature are concerned with water quality improvements and use water quality measures similar to ours. We consider some of these below.

Bhat (2003) analyzed coral reef recreation (glass-bottom boating, snorkeling, diving) in the Florida Keys. Environmental quality in their contingent scenarios included fish abundance, underwater visibility, and improved coral quality. For each attribute, they used a simple binary variable to capture improved conditions. All three attributes were important in predicting contingent visitation – trips increased by 43% to 80% over the scenarios considered. Hanely et al. (2003) studied beach use in Scotland. Respondents were asked to rate water quality on a five-point scale from poor to very good at the beach they had visited. For the contingent scenario, they asked respondents to consider what they would do if water quality was instead at the highest rating on the 5-point scale. The improvements led to a modest 1.3% increase in total trips. Paccagnan (2007) applies a TCM-CB to a lake in northern Italy and uses a binary variable for improved water quality conditions on the lake in her contingent behavior scenario. It is not clear what the "improvement" included but respondents were shown photos of the lake in improved conditions. Trips increased by 41% with the quality improvement she considered.

Eiswerth et al. (2008) used contingent behavior data to value a loss in water clarity to anglers on a lake in Wisconsin. A single scenario was considered – a decline in visibility to 3 feet depth. The respondents were told that visibility levels averaged 10 feet during the summer. The decline was motivated by a narrative about increased algae. The response data showed a 38% decline in visitation with the decline in quality. Joen & Herriges (2010) considered water quality improvements on selected lakes across 131 possible lakes in Iowa. Water quality was measured using a traditional water quality ladder (0 to 10 rating) and each of the lakes was assigned a ranking in current and proposed conditions. The respondents were then asked to report three trip values – actual 2004, predicted 2005, and predicted 2005 if water quality was improved. In this way, respondents were allocating trips across 131 lakes with and without improvements at a subset of the lakes. The scenario considered moved all lakes below a 'swimmable' rating to at least that level. This affected 52 lakes. The five lakes with the worst water quality saw trips increase by 23%. Substitution from the currently clean sites is also evident. This is the most ambitious TCM-CB to date and shows promise for future multiple-site applications.

The three most recent water-quality applications of TCM-CB are all in Europe: Lankia et al. (2019), Bertram et al. (2019), and Borger et al. (2021). Lankia et al. (2019) studied swimming in Finland using a national inventory of recreation visitation data. All water-based recreation sites in the country are included and analyzed in a single-site framework – so data are stacked by sites. Two contingent scenarios were considered for water clarity – good and poor. The good scenario was for visibility to a depth of over 2 meters. The poor scenario was for visibility to a depth of less than 1 meter. The demand model shifts as expected with trips being 1.7 times the baseline (good conditions) and .67 times baseline (poor conditions). In both cases, the baseline was the water quality at the site visited by the respondent, which varied across sites.

Bertram et al. (2020) studied coastal recreation on the Baltic in Finland, Latvia, and Germany. They considered six attributes for water quality and varied how these changed across respondents. The attributes included clarity, algae blooms, onshore algae, number of plants and birds, number of fish, and facilities. Theirs is the most detailed attribute-based approach to date. Each respondent saw three scenarios, which varied the level of the attribute changes (improvements and deteriorations). The results showed clear sensitivity to changes in water quality as expected and considerable variation across the three countries and types of quality.

Borger et al. (2021) applied a TCM-CB to water-based recreation sites spread over 14 European Countries – 78% of the European Union's population. Like Lankia et al. (2019) and Bertram et al. (2020), they consider many sites but in a single-site framework with pooled data. Also like Lankia et al. (2019) and Bertram et al. (2020) respondents are asked to consider improvements and deteriorations in water quality from current perceived conditions. They use a 4-point water quality ladder that runs from poor to excellent. Respondents are shown a visual ladder and asked how their reported number of trips change if there was one-step improvement and one-step decline in water quality. Trips are 1.17 times baseline with the improvement and .93 with the deterioration – keeping in mind this combines moves of one-step from different baselines. Borger et al. (2021) provide detailed breakdown by country and recreation use as well as aggregation, which may be of use for policy.

4 Survey & Data Collection

We combine two data sets in our analysis: The first is an email-based sample from Dynata. We drew a stratified random sample with home addresses in Delaware from their national inventory of respondents. These are people who have agreed to participate in survey research, have provided demographic data, and have received incentives from Dynata to participate. Our strata are 60% from New Castle County, where the Brandywine Creek is located, and 40% from the rest of the state. The data were raked to match the population along the lines of income, age, and education.

The second data set is an addressed-based sample also from Dyanta. We were concerned that the sample of email addresses might not be representative of the population (even after raking), so we supplemented it with an addressed-based sample. Again, we drew a stratified random sample of addresses from people in Delaware. In this case we drew 25% from respondents in a zip code that contained or abutted the Brandywine, 25% in zip codes that did not contain or abut the Brandywine but were in New Castle County, and 50% from the rest of the state.

Since we intended for the second sample to also take our internet-based survey, we recruited them to participate by asking for an email address. Each person was sent a one-page letter (Supplementary Information) describing the project and asking if they might be interested in participating. If so, they were instructed to provide an email address by text, e-mail, or return letter. The letter included a $5 incentive. This was followed by similar second and third letters for those not responding (Supplementary Information). The third follow up included an additional incentive of $2. We started the recruitment process in early 2021 and launched a survey in fall 2021. Our recruitment response rate was 33.29% – those providing an email address. This gave us 1011 individuals to send surveys. Finally, we conducted a non-respondent survey to account for potential avidity bias. Individuals who had not provided an email address were sent a simple one-page questionnaire with two-questions about their water-based recreation trips in the past year. This question was identical to a question used in our survey, which allowed us to re-weight our sample for avidity bias. The non-respondent survey is shown in Supplementary Information.

The two samples received the same internet-based survey (see text version in Supplementary Information). The final sample size used for estimation is 880 respondents (413 email-based 467 address-based). The email-based sample was conducted by Dynata by oversampling the relevant population with a request to complete the survey and closing the request when the contracted number returns are reached. For this sample the response rates are unknown. For the address-based sample we had a response rate of 54.8%.

For the non-respondent survey, we have 354 completed surveys for a response rate of 11%. Table 1 shows the responses to the “correction question” for the respondent and non-respondent samples for contact and non-contact water-based recreation. There appears to be a modest amount of avidity bias for contact trips (i.e., higher trip taking behavior for respondents), and reverse for non-contact. These results were used to weight the data as described in the next section.

The survey itself is divided into four sections. The first describes the purpose, defines what was meant by contact and non-contact water-based recreation, and then asks respondents to report their total contact and non-contact trips to all sites in Delaware in 2021. Since nearly all the inland water-based recreation in Delaware is day-trip recreation, we ask respondents to report the number of days they spent in contact and non-contact water-based recreation. After reporting their total number trips, respondents were asked to report their number of trips specifically to the Brandywine. This helps to get people thinking about water-based recreation and then to distinguish their Brandywine trips from trips to other sites. The second section of the survey asks people to report their perception of the water quality on the waterbody nearest their home and on the Brandywine Creek. The third presents the questions used in our contingent behavior analysis. These are queries about how a hypothetical water quality improvement on the Brandywine might affect one's number of trips to the river. It begins with detailed information on how the quality outcomes would be achieved through various regulatory actions and public investments. Then, each respondent is asked two questions about changing trip behavior – first for a large change in water quality and second for a small change in quality. The large change is the same for all respondents and the small change varied over respondents and was discussed in detail in the previous section. Because the survey was taken during the COVID-19 epidemic, we ask respondents how the outbreak affected their trip taking behavior. In the final section, we asked people to report the distance they lived from the Brandywine, several demographic questions, and some follow-up questions on how the survey worked.

In developing our survey, we conducted several rounds of pre-test with friends and colleagues to be sure the flow, mechanics, and communication were working as intended. Then, using a random sample of 123 individuals from the same population we used in the full survey, we conducted another pretest, again to ensure the survey was functioning properly and that the response data was complete and reasonable. In both pretests we asked for feedback on the survey for clarity and believability, and for suggestions to improve communication. These tests were done in the summer of 2021. The sample of 123 was drawn using the same procedure for the full sample, so we also had a pretest for sampling. Based on the response to the test surveys we fine-tuned our survey and launched the final in the fall of 2021. The final version included graphics, maps, and through the usual piping and interactive features were tailored to the user. The survey took about 10 to 15 minutes to complete.

5 Response Data

Table 2 shows the distribution of our sample across the three Delaware counties in unweighted and weighted form. Weighting accounts for our stratified sampling and non-respondent bias. Since the share of New Castle County residents was oversampled, in the weighted data the count of respondents from New Castle is lower and from the other counties is higher versus their unweighted counterparts. The mean income of the respondents is $82,000 (2021 USD) and mean age is 53 years. Table 3 shows our trip frequency data. The mean number of contact trips (days recreating) is about 6 and non-contact trips is about 10. The data are heavily skewed. Non-participation is high as one would expect in an off-site sample. Over 60% of the respondents took no contact trips and over 45% of the sample took no non-contact trips.

Travel cost to and from the site includes out-of-pocket and time cost, where time cost is measured as one-third annual income divided by 2040 hours and out-of-pocket cost is 22 cent per mile traveled. We use respondent's self-reported distance to the river. We opted for this over a computed distance (e.g., using PC*Miler) given the highly local nature of the recreation use. About 30% of the respondents live within one mile of river, so using computed distances (zip-code-mid-point to zip-code-mid-point) would have introduced sizable measurement error. Indeed, a considerable amount of useful variation in travel cost is within a zip code (cases where the origin and destination are in the same zip). Also, the destination point was their mostly frequently visited spot, which was unknown to us, but this was implicit in the self-reported measure since respondents were asked to use that point on the river. The self-reported distance also has the appeal of being the 'perceived' distance on which people act. We excluded time on site, in effect assuming it is constant across trips. Yang (2012) calculated that the mean and median value of walking trips in the USA is .7 miles and .5 miles using NHTS 2009 data. We assumed that people who live within a half mile distance will walk while others will travel by car. Murtagh (2002) calculated average brisk walking speed. If we take his lower bound (assuming people are not in a hurry and relax) this is 3.5 miles per hour. People who live beyond walking distance but within 50 miles are assumed to use local roads to reach Brandywine Creek which has a speed limit of 35 mph. So, free flow (if there is no interruption in the road) travel speed is assumed to be 35 miles/hour. But if the distance is more than 50 miles, we expect that people might use highways with 55 to 65 mph speed limit, on average they will travel at 60 mph. However, they cannot use only highways on their way to the Brandywine Creek. We assume that these people travel half of the distance on the highway (at 60 mph) and half on the local road (35 miles/hour). In that case their average speed will be 47.5 mph. So, we use 45 mph to calculate free flow travel time for the people who live beyond 50 miles. However, this free flow time will not reflect actual time because of traffic interruptions, signals etc. According to the 2021 urban mobility report produced by University of Texas A&M, average actual travel time is a product of free flow time and time index (TTI). We used the TTI of 2019 which is 1.23. Travel cost distribution shows that over 50% or our respondent's travel cost is $25 or less – this reflects the large urban population near the Brandywine Creek in north Delaware.

Table 4 shows reported trips (days recreating) under contingent scenarios. This table reports actual trips from raw data. Trips change under the first contingent scenario where every attribute is improved is denoted as "Large Improvement" and the second contingent scenario where a subset of attributes improve is denoted as "Small Improvement". The table shows that contact trip demand is more responsive to water quality change than the non-contact trip demand. Mean contact trips increase by 30%

with a small improvement and almost 70% with the large improvement, while mean non-contact trips increase about 17% with a small improvement and roughly 37% with the large improvement. The same effect is observed for participation. Contact recreation increases by 11 percentage points (small improvement) and 26 percentage points (large improvement). For non-contact recreation the numbers are 5 and 15 percentage points.

Respondents are also asked to report their perception of existing water quality of the waterbody nearest their home and the Brandywine Creek in terms of a Water Quality Index (WQI). They were shown a rating scheme (see Figure 4) and asked to provide their best estimate. Using water quality monitoring stations on the Brandywine Creek reveals that actual WQI of Brandywine Creek is 70.5 (based on 2019 data). Only 30% of our samples perceived the water quality in the correct range – more than 50% of the respondents perceived quality to be below 70 and 15% perceived it above 70. On average, perceived water quality of the Brandywine Creek is better than the sample's nearest waterbody.

Most of the respondents in our data set indicated they spent most of their outdoor recreation time on the Brandywine near Wilmington (50%) and parts between Wilmington and Pennsylvania (46%). Very few reported that they spent most of their outdoor recreational time in Pennsylvania on the Brandywine portions of the river. People also reported the contact and noncontact activities they participated in most. For contact recreation these broke down as follows: swimming 27%, boating 21%, fishing 19%. For noncontact recreation the breakdown is walking 42%, relaxing 28%.

Finally, since we conducted our survey in 2021, we had a concern that covid-19 might affect trip demand. To investigate its effect, we asked respondents how their trip behavior was changed due to covid-19. Covid both encouraged people to get outside (a substitute for indoor activities) but also discouraged people from interacting, so the effect could go either way. While most people reported not being affected (60%), those reporting that it caused a decrease in use (24%) was larger than an increase (13%).

6 Econometric Model

Single-site travel cost models use regression techniques to describe visit frequency over a season as a function of respondent characteristics and round-trip travel cost to the site – the function described in our Study Design section. In this realm our model, which may be derived from conventional theory of utility maximization, has the form

(2) $x_{ij} = f(tc_i, q_{ij}, z_i, y_i)$.

x_{ij} is the number of trips taken to the Brandywine Creek by respondent i for scenario j, tc_i is respondent $i's$ trip cost to the Brandywine, q_{ij} is the water quality of the Brandywine for respondent i for scenario j, z_i is a vector of individual characteristics believed to influence demand such as income and age, and y_i is a vector of substitute site travel cost. Equation (2) is formulated for our panel data – each respondent i ($i = 1, \ldots, 880$) faces 3 scenarios ($j = 1,2,3$) giving an 880 x 3 panel. The three scenarios are the current level of water quality and two contingent behavior scenarios – a large improvement in water quality and a small improvement in water quality.

The conventional econometric approach for estimating a single-site recreation demand model is a count data model. Count models accommodate the non-negative integer-only aspects of the dependent variable. A Negative Binomial form of the count model is preferred to a Poisson form since it avoids the restriction that the predicted mean and variance trips are equal (Deely et al., 2019; Englin et al., 2001; Kosenius & Horne, 2015; Lienhoop and Ansmann, 2011; Loomis, 2002; Rolfe and Dyack, 2011). Also, when using contingent behavior panel data, one encounters within person correlation of unobservables. This is usually handled by introducing some form of random effects into the model (Landry et al., 2012; Lankia et al., 2019; Parsons, 2013; Whitehead et al., 2000). In our application, we use a Negative Binomial form and include a mixing term on the constant to account for such correlation. This is a

Negative Binomial Mixed-Effects Model (NBME), which has been applied in various settings with longitudinal data (Yigra et al., 2020; Zhang et al., 2018) but not with contingent-behavior travel cost demand. It is a flexible approach and consistently provides a good fit to our data.

Our demand specification for equation (2) then is

$$(3) \qquad E(x_{ij}) = \mu_{ij} = \exp\left((\delta + u_i) + \alpha tc_i + \beta q_{ij} + \gamma z_i + \eta y_i\right)$$

where u_i is our mixing term. We assume u_i is normally distributed with $E(u_i) = 0$ and $Var(u_i) = \sigma^2$. The parameters in equation (4), including σ^2, then are estimated in a negative binomial form

$$(4) \quad x_{ij} \sim NB(x_{ij} \mid \mu_{ij}, \theta) = \frac{\Gamma(x_{ij} + \theta)}{\Gamma(\theta)\Gamma(x_{ij} + 1)} \cdot \left(\frac{\theta}{\mu_{ij} + \theta}\right)^{\theta} \cdot \left(\frac{\mu_{ij}}{\mu_{ij} + \theta}\right)^{\mu_{ij}}$$

where θ is the dispersion parameter controlling for overdispersion, μ_{ij} is the expected number of trips defined by our demand equation (3), and the variance is $\mu_{ij}(1 + a\mu_{ij})$, where $a = \frac{1}{\theta}$. This is Cameron and Trivedi's (1986) Negbin 2 version of the Negative Binomial. For our constant term, we estimate a mean and variance using a mixing procedure wherein each respondent's draw of u_i is the same across the three scenarios. In this way, unobserved effects on demand for each respondent are the same in each scenario. So, we have terms allowing for overdispersion (θ) and for controlling for within person correlation of unobservables (σ^2). Our model then is

$$(5) \qquad E(x_{ij}) = \mu_{ij} = exp((\delta + u_i) + \alpha tc_i + \beta_{2_C} qc_{ij} + \beta_{2_S} qs_{ij} + \beta_{2_E} qe_{ij} + \beta_{2_F} qf_{ij} + \gamma z_i + \eta y_i)$$

Here we have four policy-relevant variables in the contingent demands: $qc_{ij} = 1$ if clarity improves; $qs_{ij} = 1$ if swimming safety improves, and so forth. These correspond to our four water quality attributes: clarity, swimming safety, eating safety, and fish abundance. Again, the subscript $i = 1, \ldots, 880$

denotes a respondent, and $j = 1,2,3$ denotes a contingent behavior scenario (j =1 is the baseline or no contingent scenario). We considered several general forms of equation (5) such as interactions among the water quality variables and interactions with respondents' primary recreational activity as well as different estimates for the travel cost coefficients, but the simple form shown above fit best.

We also considered a simpler form our model, which we refer to as our “preliminary” model to test the basic properties of the response data. The preliminary model has the form

$$(6) \quad E(x_{ij}) = \mu_{ij} = exp((\delta + u_i) + \alpha tc_i + \beta_{sm} QS_{ij} + \beta_{\text{lg}} QL_{ij} + \gamma z_i + \eta y_i)$$

Here we have two policy variables, $QS_{ij} = 1$ for the second water quality scenario seen by each respondent, which is a *small improvement* and $QL_{ij} = 1$ for the first water quality improvement seen by each respondent, which is the *large improvement*. In this case, no account is taken of the type of water quality improvement. It is a simple test of a coarser version of model and serves as a scope test. We expect the larger changes in water quality will give higher visitation than small changes, so $\beta_{\text{lg}} > \beta_{\text{sm}}$.

Trips were reported in a response format as a range (eg., 0 days, 1-5 days, 6-10 days, etc.). So, we used a simple bootstrap to construct the dependent variable x_{ij}. For each person we drew a value from a uniform distribution within the range reported. So, for example, if a person reported taking 6 to 10 trips, we randomly drew an integer from 6 to 10 and used that draw in estimation. For each observation we conduct 1000 draws and give the results for 1000 estimated versions of the model to allow for any error the drawing procedure might introduce.

Trip cost tc_i includes travel time and out-of-pocket travel cost and is described in a previous section. The individual characteristics z_i include age, income, dependent children living at home, and race (nonwhite). Delaware has an abundance of lakes of streams distributed across the state nearly uniformly spatially and of similar quality, so substitute site variation is not particularly important. But there are two major water bodies that do vary in terms of access depending on location so we include two variables to capture this feature – a binary response variable = 1 if the respondent lived in a zip code touching the Delaware Bay and another for living in a zip code within 10 miles of the Delaware River.

For our welfare scenarios, we report per trip and population aggregate values. The per trip is derived from the seasonal measure by dividing by trips and useful for comparison across studies and for benefits transfer. Using the demand form in equation (3) we have a seasonal consumer surplus (equation (1)) for respondent i for water quality change scenario j of

(7) $$\Delta W_{ij} = -\{E(x_{ij}^{*}) - E(x_{ij})\}/\alpha$$

where x_{ij}^{*} is trip demand with the change in water quality. Dividing equation (6) by $E(x_{ij})$ gives a per trip value for a water quality improvement. For *population aggregate values* in annual terms then we use

(8) $$PAV_j = \{\sum_{i=1}^{880} \Delta W_{ij}\} \cdot (Pop/880)\ ,$$

where Pop is the 2020 population of Delaware residents 18 years old or older and 880 is our sample size.

7 Estimation Results & Policy Simulations

We begin with a "preliminary" model (equation (6)) in which we consider two discrete shifts only in the demand function – one for each of the two contingent behavior questions. This allows for a shift for the large improvement and for the small improvement *without accounting for which of specific attributes of water quality that are changing*. The results are shown in Table 6 (see variable definitions in Table 5). Notice that the effects of the large improvement are greater than the small improvement in both the contact and non-contact models. This is as expected. In both cases the coefficients (large versus small change in quality) are statistically significantly different. This suggests at least at a coarse level that the respondents perceived the differences in water quality. Also, the results on most other covariates are similar to the complete model, which we will discuss next.

Table 7 shows the results for the complete model, our equation (5). Again, the variables used in the models are defined in Table 5. The results are again shown for contact and non-contact recreation separately. Since we estimated each version of the model 1000 times (iterating over the draws within the trip ranges), we report the mean, minimum, and maximum estimates for each coefficient. First, the models are quite stable over this range of draws, so any 'noise' associated with using ranges appears to be small. Second, the coefficients on the travel cost are negative and significant for contact and non-contact recreation. The estimate for contact trips (-.024) is less than for non-contact trips (-.035). So, there is less sensitivity to cost for contact recreation, which is what we expected. If a person is coming in direct contact with the water, we expect they may be somewhat more discerning in their choice. Non-contact trips are also more likely to be local (often walking) trips. The per trip values implied by these coefficients are \$42 for contact recreation and \$29 for non-contact recreation in 2021 dollars.

The model otherwise tells a reasonable story. For contact and noncontact recreation, age is negatively correlated with trips and income is positively correlated. Having a child in the household increases the likelihood of taking more trips. We had no prior on our non-white variable and find that is usually positively correlated with trips. Our proxies for substitute sites show mixed results. They work quite nicely in the non-contact models for the Delaware Bay but are otherwise not robust. As shown, the signs on these coefficients are not stable across the models and draws and are often not significant.

The variables of most interest in our application are the water quality variables – clarity, swimming safety, fish abundance, and eating safety. These are the q_i variables in equation (5). As shown, all have significant and positive coefficients and are robust over the draws. The coefficients for contact recreation are larger than for non-contact recreation suggesting contact users are more sensitive to water quality changes. Contact use also shows more sensitivity to the different types of improvements. So, they are more discerning about how water quality changes. Swimming safety for contact users is most important (β_{2_S}= .35) followed by safety of eating fish (.26), clarity (.23), and finally abundance (.08). Interestingly, the effect of eating fish safely is 3 times as important as approximately doubling the abundance of fish. Non-contact recreationists appear to be less concerned about how water quality changes than they are about knowing that it has changed in some way. The water quality coefficients are everywhere positive (and smaller than the contact coefficients), but their variation across the attributes is not large, ranging from .17 for swimming to .14 for fish abundance. Thinking of non-contact recreationists as "soft" users, their behavior has a nonuse aspect by not being sensitive to how the improvements are realized but wanting to see improvement.

Swimming safety is the most important attribute for both contact and non-contact recreation. This stands to reason for contact recreation, thinking about kayaking, tubing, canoeing, wading, and swimming on the Brandywine. The share of users who list swimming as their most important use is also the largest among the types of contact uses at 27%. Fishing is 19%. For non-contact recreation the greater importance to swimming may reflect a symbolic significance or familiarity – signage warning about swimming dangers on the river and newspaper articles mentioning the possibility of returning the river to swimmable. In this way swimming is probably the most salient for people who are not direct users, so it may stand out. Again, the difference across the four attributes for non-contact users is small and statistically insignificant so we do make too much about swimming having the largest coefficient.

Finally, as another test of scope (in addition to our preliminary model) we estimated a series of models wherein we considered a single-attribute change and then that same attribute paired with another – like swim only and then swim and eat. In these models we only used respondents who saw the relevant single attribute or pairing. In all cases, the models pass the scope test in that the paired-attribute models always show higher values than single-attribute models. Keep in mind the single-attribute respondents in all these cases are different from the paired-attribute respondents, so anchoring effects are obviated.

Table 8 shows several policy simulations using the model. We show the welfare effects of improving all attributes, each attribute individually, and then selected pairs. We also report the numbers in per trip and aggregated terms. The aggregated values are for the entire state for one year. We use the mean values for the coefficient estimates in Table 7 in the calculations. Consider the per trip values first.

The lowest and highest effects for a single attribute changing for contact recreation are $3.58 per trip for fish catch improving and $17.43 per trip for swimming safety improving. For non-contact the lowest and highest are $4.19 for fish catch improving and $6.13 for swimming safety improving. Again, we see the larger spread for contact versus non-contact. The table also shows some logical combinations of improvements like increased abundance of fish and improved safety of eating fish and improved view and swimming. These all give higher values and given the non-linearity of the demand model are greater than the simple sum of the individual attributes. For the improvement of all four attributes the per trip value is $63.10 for contact use and $27.70 for non-contact. Both are larger than the per trip values to the Brandywine (the loss incurred if the site were closed to recreation.) Again, these are improvements at very high levels but are imagined in discussions by policy makers.

Table 8 also shows the aggregate annual measures. Here we see a range of values for total annual benefits from $10 million (improving catch of fish only) to $89 million (improving all aspects of water quality). To put these welfare gains into perspective, Keiser and Shapiro (2019) estimate that it costs about $1.5 million per mile per year to keep a river suitable for fishing. This estimate is an 'average' marginal cost across all rivers in the United States to come into compliance with the Clean Water Act. No doubt the marginal cost varies by a river's baseline level of cleanliness and location, but the Brandywine is neither extremely dirty nor extremely clean, so the 'average' estimates are likely to be within reason. Since the Brandywine Creek is already fishable and assuming the marginal cost curve for cleaning a river mile is rising, $1.5 million per year is likely a lower bound estimate for the cost for water quality improvement for each of our scenarios. Also given the Brandywine's urban location, we can assume that it is probably on the higher (more costly) end of the distribution of marginal costs across rivers in the United States.

The Delaware portion of the Brandywine Creek is about 10 miles long. Let's assume the marginal cost is $2 million per mile to account to the rising marginal cost and urban aspect. This gives a cost of $20 million. If we include cleaning the Pennsylvania portion of the river, which would presumably have to happen simultaneously to realize the effects, the lower bound estimate of the costs is closer to $40 million per year. Comparing these costs to estimates in Table 8, we see that the costs are likely to exceed benefits if the gains in quality are limited to single attribute changes but do fall below benefits if the gains are over some pairs and triples of attributes and if all attributes improve. So, our results suggest that the potential for efficiency is there, but it does depend on some key assumptions: (1) regulatory and other changes yield *multi-level* improvements, (2) cost are not higher than what we have admittedly estimated only roughly, (3) no significant improvements elsewhere in the state, and (4) reasonably stable preferences over time. These numbers set rough boundaries on what is needed to pass a benefit-cost test. It also makes clear that efficiency is nearly clearly a winner or loser, which is sometimes assumed on both sides of the debate.

8 Conclusions

Many urban rivers in the United States have realized improvements in water quality in recent decades due to an increase in environmental regulatory actions. Further improvements may yield more benefits along the lines of visual and aquatic attributes of the rivers. At the same time, the costs for water quality improvement are rising and further incremental improvements will almost certainly be more costly than the improvements realized earlier in the evolution of regulatory control. This raises the question for many river systems, especially urban river systems, of how high to set the goals for improvement. In this paper we consider the benefits of one urban river and its tributaries, the Brandywine Creek in Delaware.

Using a contingent behavior travel cost model with survey-based data we show that there are indeed benefits for incremental improvements for commonly targeted attributes of water quality – clarity, swimming safely, fish abundance, and safety of eating fish. Benefits are realized for contact (swimming, fishing, etc.) and non-contact (walking, relaxing etc. nearby) recreation uses. The per trip values for contact uses are more than twice that of non-contact uses but the large numbers of trips for non-contact recreation make it larger in significance for aggregate benefits. We expect this is true for many urban rivers where people walk, relax, bike, and otherwise passively enjoy a river. Improvements in safety for swimming was the most important attribute of water quality for both contact and non-contact use. "Swimmable" appears to be a salient marker for individuals. The least important attribute was abundance of fish. This is interesting in that more studies measuring water quality benefits focus on fishing than any other attribute. Our results suggest that other aspects need more attention, especially non-contact uses. The lower fish abundance measures may also be due to satisfaction with current catch rates of fish. The safety of eating fish was far more important to contact users – again, something all urban rivers may see. Our analysis was done during COVID-19. Respondents indicated that their number of trips was modestly lower than it would otherwise have been. Finally, respondents' perception of water quality tended to understate true levels.

Methodologically we contribute to the contingent behavior literature in several ways. Only a few TCM-CB studies have used contingent behavior response data to value attributes of a resource instead of a full single resource change. This allows for a finer targeting of policy. We also reduce cognitive burden on the actual and contingent trips responses using ranges of trips (e.g., 6 to 10 trips) and then randomly drew a trip count in the range for each respondent. We believe that is the better side of trade off – lower cognitive burden versus accuracy in count. We also used a follow-up non respondent survey to correct for avidity bias (and find little bias). And finally, we offered a new method for recruiting an ABS sample to address bias from non-probabilistic internet-based samples, which may be transferable to other settings.

Acknowledgments

This publication was made possible by the National Science Foundation EPSCoR Grant No. 1757353 and the State of Delaware. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Data Availability

The data and supporting material is available at [Brandywine Data]

Reference List

Adamowicz, W ; Louviere, J; Williams, M. (1994). Combining revealed and stated preference methods for valuing environmental amenities. In *Journal of Environmental Economics and Management* (Vol. 26, pp. 271–292).

Bertram, C. (2020). Contingent Behavior and Asymmetric Preferences for Baltic Sea Coastal Recreation. *Environmental and Resource Economics*, *75*(1), 49–78. https://doi.org/10.1007/s10640-019-00388-x

Bhat, M. G. (2003). Application of non-market valuation to the Florida Keys marine reserve management. *Journal of Environmental Management*, *67*(4), 315–325. https://doi.org/10.1016/S0301-4797(02)00207-4

Bockstael, N & McConnell, K. (2007). *Environmental and Resource Valuation with Revealed Preferences: A Theoretical Guide to Empirical Methods.* Dordrecht: Springer.

Börger, T., Campbell, D., White, M. P., Elliott, L. R., Fleming, L. E., Garrett, J. K., Hattam, C., Hynes, S., Lankia, T., & Taylor, T. (2021). The value of blue-space recreation and perceived water quality across Europe: A contingent behaviour study. *Science of the Total Environment*, *771*. https://doi.org/10.1016/j.scitotenv.2021.145597

Cameron, T. A. (1992). Combining Contingent Valuation and Travel Cost Data for the Valuation of Nonmarket Good. *Land Economics*, *68*(3), 302–317. https://doi.org/10.4135/9781412969024.n17

Chase, L. C., Lee, D. R., Schulze, W. D., & Anderson, D. J. (1998). Ecotourism demand and differential pricing of national park access in Costa Rica. *Land Economics*, *74*(4), 466–483.

Deely, J., Hynes, S., & Curtis, J. (2019). Combining actual and contingent behaviour data to estimate the value of coarse fishing in Ireland. *Fisheries Research*, *215*(March), 53–61. https://doi.org/10.1016/j.fishres.2019.03.008

Eiswerth, M. E., Kashian, R. D., & Skidmore, M. (2008). Examining angler behavior using contingent behavior modeling: A case study of water quality change at a Wisconsin lake. *Water Resources Research*, *44*(11), 1–10. https://doi.org/10.1029/2006WR005828

Englin, J., Loomis, J., & González-Cabán, A. (2001). The dynamic path of recreational values following a forest fire: a comparative analysis of states in the Intermountain West. *Canadian Journal of Forest Research*, *31*(10), 1837–1844. https://doi.org/10.1139/cjfr-31-10-1837

Flores, N. E. (2017). Conceptual Framework for Nonmarket Valuation. In *A primer on nonmarket valuation* (2nd ed., p. 39). Springer. https://doi.org/10.1007/978-94-007-7104-8

Kosenius, A. K., & Horne, P. (2016). Welfare effects of mining externalities: a combined travel cost and contingent behaviour study. *Journal of Environmental Economics and Policy*, *5*(3), 265–282. https://doi.org/10.1080/21606544.2015.1107511

Kragt, M. E., Roebeling, P. C., & Ruijs, A. (2009). Effects of Great Barrier Reef degradation on recreational reef-trip demand: A contingent behaviour approach. *Australian Journal of Agricultural and Resource Economics*, *53*(2), 213–229. https://doi.org/10.1111/j.1467-8489.2007.00444.x

Landry, C. E., Allen, T., Cherry, T., & Whitehead, J. C. (2012). Wind turbines and coastal recreation demand. *Resource and Energy Economics*, *34*(1), 93–111. https://doi.org/10.1016/j.reseneeco.2011.10.001

Lankia, T., Neuvonen, M., & Pouta, E. (2019). Effects of water quality changes on the recreation benefits of swimming in Finland: Combined travel cost and contingent behavior model. *Water Resources and Economics*, *25*(September 2017), 2–12. https://doi.org/10.1016/j.wre.2017.10.002

Layman, R. C., Boyce, J. R., & Criddle, K. R. (1996). Economic valuation of the chinook salmon sport fishery of the Gulkana River, Alaska, under current and alternate management plans. *Land Economics*, *72*(1), 113–128. https://doi.org/10.2307/3147161

Lienhoop, N., & Ansmann, T. (2011). Valuing water level changes in reservoirs using two stated preference approaches: An exploration of validity. *Ecological Economics*, *70*(7), 1250–1258. https://doi.org/10.1016/j.ecolecon.2011.02.014

Loomis, J. (2002). Qualtifying recreational use values from removing dams and restoring free-flowing rivers: A contingent behaviour teavel cost demand model for Lower Snake River. *Water Resource Research*, *38*(6), 1–8. https://doi.org/10.1029/2000WR000136

Maler, K.-G. (1974). *Environmental Economics: A Theoretical Inquiry*. Johns Hopkins Press for Resources for the Future.

Mcconnell, K. E. (1986). *The Damage to Recreational Activities from PCB's in the New Bedford Harbor*.

Murtagh, E. M., Boreham, C. A. G., & Murphy, M. H. (2002). Speed and exercise intensity of recreational walkers. *Preventive Medicine*, *35*(4), 397–400. https://doi.org/10.1006/pmed.2002.1090

Paccagnan, V. (2007). On combining stated preferences and revealed preferences approaches to evaluate environmental resources having a recreational Use. https://mpra.ub.uni-muenchen.de/5867/

Parsons, G. R., Chen, Z., Hidrue, M. K., Standing, N., & Lilley, J. (2013). Valuing beach width for recreational use: Combining revealed and stated preference data. *Marine Resource Economics*, *28*(3), 221–241. https://doi.org/10.5950/0738-1360-28.3.221

Rolfe, J., & Dyack, B. (2011). Valuing Recreation in the Coorong, Australia, with Travel Cost and Contingent Behaviour Models. *Land Economics*, *87*(277), 282–293. https://doi.org/10.1111/j.1475-4932.2010.00683.x

Ward, F. A. (1987). Economics of water allocation to instream uses in a fully appropriated river basin: Evidence from a New Mexico wild river. *Water Resources Research*, *23*(3), 381–392. https://doi.org/10.1029/WR023i003p00381

Whitehead, J. C., Haab, T. C., & Huang, J. C. (2000). Measuring recreation benefits of quality improvements with revealed and stated behavior data. *Resource and Energy Economics*, *22*(4), 339–354. https://doi.org/10.1016/S0928-7655(00)00023-3

Yang, Y., & Diez-Roux, A. V. (2012). Walking distance by trip purpose and population subgroups. *American Journal of Preventive Medicine*, *43*(1), 11–19. https://doi.org/10.1016/j.amepre.2012.03.015

Yirga, A. A., Melesse, S. F., Mwambi, H. G., & Ayele, D. G. (2020). Negative binomial mixed models for analyzing longitudinal CD4 count data. *Scientific Reports*, *10*(1), 1–15. https://doi.org/10.1038/s41598-020-73883-7

Zhang, X., Pei, Y. F., Zhang, L., Guo, B., Pendegraft, A. H., Zhuang, W., & Yi, N. (2018). Negative binomial mixed models for analyzing longitudinal microbiome data. *Frontiers in Microbiology*, *9*(JUL), 1–10. https://doi.org/10.3389/fmicb.2018.01683

Tables

Table 1: Respondent and non-respondent comparison (unweighted data)

	Contact days		Non-Contact days	
	Respondent (%)	Non-respondent (%)	Respondent (%)	Non-respondent (%)
0 days	57.2	53.9	39.2	34.1
1 to 5 days	22.3	19.8	24.2	14.4
6 to 10 days	9.6	9.6	12.3	13.2
11 to 25 days	5.1	9.0	10.8	12.6
26 to 50 days	3.2	5.4	6.3	11.4
51 to 100 days	1.7	1.8	3.8	4.8
101 to 200 days	1.0	0.0	3.5	6.0
> 200 days	0.0	0.6	0.0	3.6
Total	100	100	100	100
Mean Trips	6.6	5.3	15.4	21.0

Table 2: Sample distribution by county (N = 880)

County	Raw (%)	Weighted (%)	Population (%)
Kent County	13.5	18.4	18.4
New Castle County	67.3	57.7	57.0
Sussex County	19.2	24.0	24.7
Total	100	100	100

Table 3: Trip Distribution to Brandywine Creek (N = 880)

	Contact days		**Non-contact days**	
Number of days	Raw (%)	Weighted (%)	Raw (%)	Weighted (%)
0	57.2	61.3	39.2	46.9
1 to 5	22.3	19.3	24.2	24.7
6 to 10	9.6	9.3	12.3	10.7
11 to 25	5.1	4.5	10.8	9.8
26 to 50	3.2	2.9	6.3	4.2
51 to 100	1.7	1.6	3.8	1.8
101 to 200	1.0	1.1	3.5	2.0
Total	100	100	100	100
Mean Trips	6.6	6.3	15.4	10.0

Table 4: Raw trip distribution in contingent scenarios (N = 880)

	Contact Trips			**Non-contact Trips**		
Trips	Current Condition (%)	Small Improvement (%)	Large Improvement (%)	Current Condition (%)	Small Improvement (%)	Large Improvement (%)
0	57.2	46.4	31.6	39.2	34.0	24.7
1 to 5	22.3	23.8	26.0	24.2	19.9	18.3
6 to 10	9.6	11.7	16.1	12.3	13.8	15.5
11 to 25	5.1	10.8	15.5	10.8	15.2	19.1
26 to 50	3.2	4.2	7.3	6.3	8.3	12.3
51 to 100	1.7	1.7	1.6	3.8	5.1	5.7
101 to 200	1.0	1.4	1.7	3.5	3.0	3.1
>200	0.00	0.1	0.2	0.0	0.80	1.5
Total	100	100	100	100	100	100
Mean trip	6.6	8.6	11.2	15.4	18.0	21.1

Table 5: Description of variables

Variables	Description
Trip Cost	Travel plus time cost
ln(age)	Natural logarithm of individuals age
Income	Individuals' income in 1000 USD
Child	= 1 if individual lives with dependent children
Nonwhite	= 1 if individual in non-white
Delaware River	= 1 if individual live in a ZIP that touches Delaware River
Delaware Bay	= 1 if individual live in a ZIP within 10 miles of Delaware Bay
Clarity	= 1 if view depth improved from 1.5 ft. to 4 ft.
Swimming safety	= 1 if swimming safety improved from 7% to 0% chance of getting sick
Fish abundance	= 1 if catch rate increase 100% & stocking increase 50%
Safety of eating fish	= 1 if safety of eating fish increase to unlimited consumption
Large Improvement	= 1 if improvement is realized in all four attributes (clarity, swimming safety, fish abundance, safety of eating fish)(this is the first contingent behavior question for each respondent)
Small Improvement	= 1 if improvement is realized less than all four attributes (this is the second contingent behavior question for each respondent)

Table 6: Preliminary mixed effects negative binomial model with improvement-level distinction only

Variables	Contact Trips	Non-contact Trips
Constant	7.65*	5.05*
	(0.84)	(0.85)
Trip Cost	-0.025*	-0.036*
	(0.003)	(0.003)

Large Improvement	0.95*	0.70*
	(0.08)	(0.05)
Small Improvement	0.42*	0.30*
	(0.04)	(0.03)
ln(age)	-2.09*	-1.13*
	(0.21)	(0.21)
Income (in 1000)	0.01*	0.01*
	(0.00)	(0.00)
Child	0.42*	0.53*
	(0.16)	(0.15)
Nonwhite	0.26	0.11
	(0.17)	(0.16)
Delaware River	0.14	-0.08
	(0.21)	(0.19)
Delaware Bay	0.05	-0.13
	(0.23)	(0.24)
ln(alpha)	-1.91*	-2.40*
	(0.23)	(0.24)
Var(constant)	4.11*	4.02*
	(0.30)	(0.29)
Observations	2,640	2,640
Sample Size	880	880

Robust standard errors in parentheses
* p<0.01

Table 7: Mixed effects negative binomial model with 1000 draws

	Contact Trips			Non-contact Trips		
	Mean	Min	Max	Mean	Min	Max
Constant	7.81*	7.34*	8.32*	4.96*	4.49*	5.38*

	(0.83)	(0.81)	(0.86)	(0.84)	(0.80)	(0.87)
Trip Cost	-0.024*	-0.027*	-0.022*	-0.035*	-0.038*	-0.032*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Clarity	0.23*	0.17*	0.31*	0.17*	0.13*	0.22*
	(0.05)	(0.05)	(0.06)	(0.04)	(0.03)	(0.04)
Swimming safety	0.35*	0.29*	0.44*	0.19*	0.14*	0.24*
	(0.06)	(0.05)	(0.06)	(0.04)	(0.03)	(0.05)
Fish abundance	0.08*	0.03*	0.16*	0.14*	0.10*	0.19*
	(0.04)	(0.04)	(0.05)	(0.03)	(0.03)	(0.03)
Safty of eating fish	0.26*	0.21*	0.31*	0.17*	0.13*	0.21*
	(0.05)	(0.04)	(0.05)	(0.03)	(0.03)	(0.04)
ln(age)	-2.12*	-2.25*	-2.01*	-1.11*	-1.22*	-0.99*
	(0.21)	(0.20)	(0.21)	(0.21)	(0.20)	(0.22)
Income (in $1000)	0.01*	0.01*	0.01*	0.01*	0.01*	0.01*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Child	0.38*	0.31*	0.47*	0.47*	0.38*	0.54*
	(0.16)	(0.15)	(0.16)	(0.15)	(0.15)	(0.16)
Nonwhite	0.28	0.18	0.37	0.11	-0.01	0.20
	(0.17)	(0.17)	(0.17)	(0.16)	(0.15)	(0.16)
Delaware River	0.17	0.05	0.27	-0.08	-0.20	0.05
	(0.21)	(0.20)	(0.21)	(0.19)	(0.19)	(0.20)
Delaware Bay	-0.01	-0.17	0.12	-0.19	-0.38	-0.04
	(0.22)	(0.22)	(0.23)	(0.24)	(0.23)	(0.24)
ln(alpha)	-2.08*			-3.57*		
	(0.19)			(0.30)		
Var(constant)	3.97*			3.99*		
	(0.29)			(0.29)		
Sample Size		880			880	

Robust standard errors in parentheses
* p<0.01

Note: The mixed effects model was estimated 1000 times. This table reports the mean, minimum, and maximum parameter estimates across those 1000 models.

Table 8: Annual aggregate welfare estimation (in 2021 USD)

	Contact use	Non-contact use	Total Benefit
View depth 1.5 ft. to 4 ft.			
Per trip	$10.55	$5.36	
Aggregate	6.42M	9.75M	16.17M
Swimming safety: 7% to 0% chance of getting sick			
Per trip	$17.43	$6.13	
Aggregate	10.61M	11.15M	21.77M
Catch rate: double & 50% stocking increase			
Per trip	$3.58	$4.19	
Aggregate	2.18M	7.61M	9.79M
Safety of eating fish: 3 fish/year to unlimited consumption			
Per trip	$12.47	$5.45	
Aggregate	7.59M	9.91M	17.50M
Improving catch rate and safety of eating fish together			
Per trip	$17.13	$10.44	
Aggregate	10.43M	18.98M	29.41M
Improving swimming safety and view depth together			
Per trip	$32.45	$12.64	
Aggregate	19.76M	22.99M	42.75M
Improving catch rate, safety of eating fish and view depth together			
Per trip	$32.07	$17.76	
Aggregate	19.53M	32.28M	51.81M
Improving all attributes			
Per trip	$63.1	$27.7	
Aggregate	38.42M	50.36M	88.77M

Figures

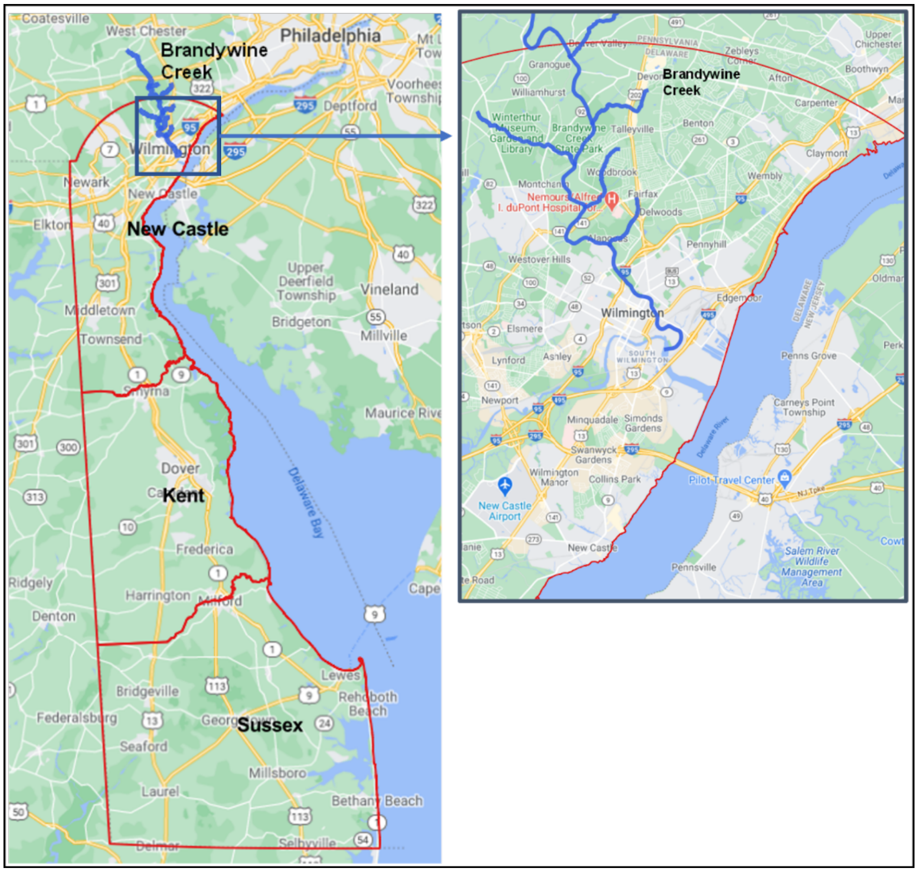

Figure 1: The Brandywine Creek and its tributaries

Category of Water Quality		Current Conditions	Conditions with Improvement
Clarity of Water		Can see to 1.5 feet deep	Can see to 4 feet deep
Safety for Swimming		7% chance of getting sick	0% chance of getting sick
Abundance & Type of Fish		Fish like Bass, Perch, Catfish occur naturally Fish supported by stocking like Trout, Salmon	Increased abundance of naturally occurring game fish Catch rates expected to double Increased stocking by 50%
Safety of Eating Fish		Eat no more than 3 fish per year caught from these waters.	Safe for unlimited consumption
Ecological Health		Fair	Very Good

Figure 2: First contingent behavior scenario: improvement in all attributes

Category of Water Quality	Current Conditions		Conditions with Improvement
Clarity of Water	Can see to 1.5 feet deep	No Change	Can see to 1.5 feet deep
Safety for Swimming	7% chance of getting sick	Improves	0% chance of getting sick
Abundance & Type of Fish	Fish like Bass, Perch, Catfish occur naturally Fish supported by stocking like Trout, Salmon	No Change	Fish like Bass, Perch, Catfish occur naturally Fish supported by stocking like Trout, Salmon
Safety of Eating Fish	Eat no more than 3 fish per year caught from these waters.	Improves	Safe for unlimited consumption
Ecological Health	Fair	Improves	Very Good

Figure 3: Example of second contingent behavior scenario: improvement in a subset of attributes

WATER QUALITY RATING SYSTEM	
Score Range	**Water Quality Characteristics**
Very Good 81 - 100	Clear to 6 feet deep Safe for swimming and drinking (without treatment) Naturally occurring game fish (like bass) and sport fish (like trout) No fish consumption advisories Abundant bank-side vegetation
Good 61 - 80	Clear to 4 feet deep Safe for swimming but not drinking (without treatment) Naturally occurring game fish (like bass) but no sport fish (like trout) Minor fish consumption advisories Good bank-side vegetation
Fair 41 - 60	Clear to 2 feet deep Unsafe for swimming or drinking Naturally occurring game fish (like bass) but no sport fish (like trout) Some fish consumption advisories Moderate bank-side vegetation
Poor 21 - 40	Clear to 1 foot deep Unsafe for swimming or drinking Naturally occurring rough fish (like carp), no game fish (like bass), and no Sport fish (like trout) Major fish consumption advisories Limited bank-side vegetation
Very Poor 1 - 20	Clear to a few inches deep Unsafe for swimming or drinking No fish No bank-side vegetation

Figure 4: Water Quality Rating System