Panel data causal inference for travel cost models: An application to beach closures

Andrew Earle and Hyunjung Kim*

September 11, 2024

Abstract

Panel data causal inference techniques are fundamental econometric tools, but they have not been fully leveraged to estimate structural demand models. We introduce a method that allows researchers to apply panel data causal inference techniques within travel cost random utility models. The method is particularly useful for valuing changes in the quality of choice alternatives. We illustrate our approach by valuing the recreational welfare losses caused by water-quality-induced beach closures. Our analysis uses administrative visitation logs documenting more than 740,000 visits and finds that welfare losses can persist weeks after a beach reopens.

*Earle: East Carolina University, earlea23@ecu.edu. Kim: Korea University, hyunjungkim@korea.ac.kr. We thank Soren Anderson, Frank Lupi, Roger von Haefen, and seminar participants at East Carolina University, NC State University, the Social Cost of Water Pollution Workshop, Heartland Environmental and Resource Economics Workshop, the W5133 Workshop, and the 2024 AERE Summer Conference for their comments. We thank the Huron-Clinton Metroparks Staff for sharing their data and supporting this work.

1 Introduction

Panel data causal inference techniques are fundamental econometric tools, but they have had a limited impact on certain economic fields (Angrist & Pischke, 2010; Currie et al., 2020; Goldsmith-Pinkham, 2024). For example, studies estimating structural demand models typically favor instrumental variables, rather than difference-in-differences or synthetic control methods. While instrumental variables have proven useful for addressing price endogeneity in industrial organization settings, some structural demand models may benefit from incorporating panel data causal inference techniques. Travel cost random utility models offer one such case, as travel cost analyses often seek to value changes in some amenity or site attribute (Kremer et al., 2011; English et al., 2018; Kuwayama et al., 2022). Researchers frequently apply panel data causal inference techniques to analyze such changes in other empirical contexts, suggesting that panel data approaches could also be useful in travel cost models.

Nevertheless, travel cost analyses have faced two barriers to exploiting panel variation. The first barrier is data-related. Estimating a travel cost model requires information on visitors' travel costs and trip-taking behavior. Traditionally, this information has been collected with surveys, which often struggle to capture temporal variation. Modern mobility datasets, often derived from cellphones or collected for administrative purposes, now provide similar information at larger scales and relatively high frequencies.[1] Consequently, the remaining barrier to leveraging panel variation in travel cost models is methodological. Travel cost analyses often prioritize modeling inter-site substitution and conducting counterfactual wel-

[1]Kremer et al.'s analysis valuing drinking water well improvements illustrates how data limitations can impact travel cost analyses. The authors' survey data only include trips in the post-improvement period, so they cannot exploit within-well variation in water quality. See Newbold et al. (2022), Christensen and Osman (2023), Gellman et al. (2023), Knittel et al. (2023), and Kreindler and Miyauchi (2023) for examples of modern mobility datasets.

fare analysis, making the travel cost random utility model (RUM) a popular tool. Estimating RUMs, in general, can be computationally burdensome, as it often involves calculating and optimizing complex, non-linear functions. This complexity encourages parsimonious specifications that average over temporal variation. When RUMs do preserve temporal variation, their non-linearity can complicate identification. Up to this point, the panel data causal inference techniques so popular in reduced-form, linear models have not been adapted to estimate RUMs.

As a result of these challenges, many travel cost analyses rely on cross-sectional variation—e.g., comparing visitation at sites with good and poor water quality—which can leave estimates susceptible to omitted variables bias (Ji et al., 2020). In the recreation demand context, where travel cost RUMs are the workhorse model, leaders have called for researchers to implement more rigorous identification strategies, including panel data approaches (Lupi et al., 2020).

We propose a method that allows researchers to apply panel data causal inference techniques, including difference-in-differences and synthetic control methods, within a travel cost RUM. The method is particularly useful for valuing the welfare impacts of shocks to the quality choice alternatives, or sites. Our approach adapts the standard travel cost RUM by including a panel of site mean utilities that vary at relatively high-frequencies—e.g., a site-by-date or site-by-week panel. These mean utilities capture cross-sectional and temporal variation in the quality of each site, and we identify them using similar panel variation in site visitation. After estimating the mean utilities, we unpack them in a second estimation stage. In its simplest form, this second stage is a linear regression of the mean utilities on site attributes, and it can be estimated with OLS. Thus, if at least one site experiences a quasi-experimental change in one of its attributes, researchers can apply panel data causal inference techniques, including difference-in-differences and synthetic control methods, to estimate the causal effect of the attribute change on the mean utilities. With estimates in hand, we calculate the welfare impact of the attribute change by comparing individuals' expected utility given observed mean utilities to their expected utility given counterfactual mean utilities—i.e., the mean utilities the treated sites would have provided if the attribute change had not occurred. In short, our method allows researchers to value quasi-experimental changes in site attributes using “structural difference-in-differences” or “structural synthetic controls”.

We illustrate our method by valuing the recreational welfare losses from beach closures caused by poor water quality. We focus on a series of beach closures at Lake St. Clair Metropark between July 21 and September 15, 2022, and our analysis combines administrative visitation logs with daily beach closure records. The visitation logs contain visitors’ home ZIP code and the minute of entry for over 1 million visits to the Huron-Clinton Metroparks in southeast Michigan. We restrict our focus to visits made by annual passholders, because a unique household identifier allows us to observe each household’s full visitation history for the season. This restriction leaves us with 740,000 visits (roughly 70 percent of all visits) made by 115,000 annual passholders.

In the first estimation stage, we recover a full panel of site-by-date mean utilities. These estimates describe the mean utility a visitor receives from visiting a given park on a given day under observed conditions in the summer of 2022—with the beach closures. In the second estimation stage, we use the synthetic control method to estimate the causal effect of the beach closures on Lake St. Clair Metropark’s mean utilities. The estimated causal effect allows us to calculate Lake St. Clair Metropark’s counterfactual mean utilities—i.e., the mean utilities the park would have provided in absence of the beach closures. We calculate the welfare impact of the closures by comparing expected utility under observed conditions (with closures) to expected utility under counterfactual conditions (no closures).

We estimate that the 2022 beach closures caused aggregate welfare losses of $150,000 for annual passholders, about $14 per lost trip. Applying an event study design in the second stage produces similar estimates. We also find evidence that the closures cause a stigma effect, as welfare losses can continue to accumulate weeks after a closure ends. Losses tend to be larger on weekends and hotter days, when the swimming beach is likely most popular.

We compare our estimates to two more traditional models with a cross-section of alternative-specific constants and an indicator variable equal to one on days when Lake St. Clair is impacted by the closures. Estimates from this traditional approach are sensitive to assumptions on the duration of stigma effects. If we assume Lake St. Clair is impacted every day after the first closure, then the aggregate welfare loss estimate is similar to the estimate from our proposed method, only about 9 percent smaller. If we assume Lake St. Clair is only impacted on days when the beach is closed—i.e., there is no stigma effect—then the aggregate welfare loss is only one-third as large as our proposed method. Unlike these traditional approaches, our method requires minimal assumptions regarding the presence and duration of stigma effects. Instead, we specify the initial closure date and allow the data to reveal when Lake St. Clair is impacted.

Our work makes three main contributions. First, we contribute to the literature developing structural demand modeling and estimation methods. Much of this research originated in industrial organization and consumer demand contexts where unobserved product attributes correlated with prices pose an identification challenge. Berry (1994), Berry et al. (1995), Nevo (2001), and others avoid bias from these unobservable attributes by linearizing the RUM estimation problem using a contraction mapping, which allows them to instrument for product prices. Murdock (2006) solves the analogous identification challenge in travel cost models without applying instrumental variables. She leverages individual-level variation in travel costs to separately identify the travel cost coefficient and a cross-section of site mean utilities. Including the mean utilities controls for all observed and unobserved site attributes and prevents them from biasing the travel cost coefficient estimate.[2] Although the mean utilities subsume preferences for time-invariant site attributes, these preferences

[2]Note that travel costs effectively serve as prices. In traditional industrial organization settings, prices do not vary across individuals, and thus, preferences for prices are subsumed by mean utilities.

can be recovered by regressing mean utilities on site attributes in a second estimation stage. The linear nature of the second-stage regression has proven useful in other travel cost analyses. For example, Timmins and Murdock (2007) use the linear second-stage to estimate preferences for site congestion using instrumental variables.

In sum, existing RUM estimation methods tend to focus on overcoming price (or travel cost) endogeneity and rely on instrumental variables identification strategies. Intuitively, our method is similar to the classic, industrial organization papers that linearize the RUM estimation problem to apply instrumental variables, a technique originally used in reduced-form settings. Our method differs in that we do not linearize the estimation problem to instrument for prices. Rather, we linearize the estimation problem to estimate the causal effect of some attribute change on site mean utilities. Methodologically, our two-stage approach is similar to Murdock and Timmins'. Our approach differs by preserving panel variation for the second estimation stage, which permits a broader array of identification strategies. The popularity of panel data identification strategies in reduced-form settings suggests that the ability to apply difference-in-differences or synthetic control methods within travel cost RUMs could prove useful. Indeed, the ability to leverage within-site variation facilitates our remaining contributions.

Second, we contribute to the recreation demand, nonmarket valuation literature by providing a method optimized to value shocks to environmental quality. Much of the recreation demand literature estimates preferences for marginal changes in average environmental quality measures (Phaneuf et al., 2000; Egan et al., 2009; Keiser, 2019). Yet, many natural resource changes are nonmarginal and difficult to capture in average measures—e.g., harmful algal blooms, wildfires, and infrastructure improvements. Fewer recreation demand papers use revealed preference methods to value shocks like these. One exception is English et al.'s (2018) analysis valuing the recreational welfare impacts of the *Deepwater Horizon* oil spill. Their model allows mean utilities to vary between non-spill and spill conditions and estimation occurs in a single stage. Our approach has two main advantages relative to English et al.'s. First, our method can capture substantial temporal heterogeneity in the magnitude and timing of welfare impacts. English et al. average over this temporal heterogeneity, because it is computationally burdensome to include a large number of parameters their one-stage estimation procedure. Second, our second estimation stage allows researchers to leverage any panel data causal inference technique. Researchers can articulate identification strategies in conventional causal inference terminology, base their identification strategies on common assumptions, and assess the plausibility of these assumptions—e.g., by checking for pre-trends.

Third, our application valuing the impacts of beach closures contributes a policy-relevant welfare estimate. Since the passage of the Clean Water Act (CWA) in 1972, the United States has spent over $1 trillion in an effort to make all waters fishable and swimmable, and these investments have led to widespread water quality improvements (Keiser & Shapiro, 2018). Nonetheless, roughly 30 percent of U.S. beaches experience at least one closure or water quality advisory each year, rendering them temporarily "unswimmable" (EPA 2023 Beach Report). Cost-benefit analyses of water quality policy require reliable estimates of the welfare losses caused by these beach closures. Some research valuing the welfare losses of beach closures uses stated preference methods to value hypothetical closures (Roberts et al., 2008; Boudreaux et al., 2023). Other papers use revealed preference, recreation demand methods, but these revealed preference papers typically describe beach closures approximately, using full site closures or summary statistics—e.g., the average number of beach closures over the last three years (Murray et al., 2001; Parsons et al., 2009; Wolf et al., 2019).[3] An exception, Lew and Larson (2005) exploit within-season variation to estimate the welfare losses caused by observed beach closures. Counterintuitively, their estimate of the beach closure coefficient is not statistically significant. Like Lew and Larson, we value the impacts of observed beach

[3]Approximating the welfare impacts of a beach closure by treating it as a full site closure may be misleading, because beach closures can prohibit swimming while still allowing beach access. This approach may also omit welfare impacts caused by stigma effects.

closures without using approximations based on full site closures or summary statistics. However, we document significant welfare losses and lingering stigma effects.

The remainder of the paper proceeds as follows. Section 2 and Section 3 describe our model and two-stage panel estimation procedure. Section 4 applies our approach to estimate the welfare impacts of the 2022 beach closures at Lake St. Clair Metropark, and Section 5 concludes.

2 Model

We introduce our model in the context of recreation demand. We adopt a repeated travel cost RUM framework, in which individuals repeatedly decide where to recreate, selecting the alternative from their choice set that provides them with the highest utility at each choice occasion. The choice set consists of recreation sites, indexed $j = \{1, ...J\}$, as well as an outside option, indexed $j = 0$, and we assume that each individual faces daily choice occasions—i.e., each individual makes a recreation choice every day.

We specify the utility individual i receives from visiting site j at choice occasion t as

$$U_{ijt} = \delta_{jt} + \alpha^{TC} TC_{ij} + \alpha^X X_{ijt} + \epsilon_{ijt}, \tag{1}$$

where TC_{ij} is the travel cost, X_{ijt} contains demographic characteristics, and ϵ_{ijt} is an idiosyncratic error term. We call the δ_{jt} parameters "mean utilities". They are site-by-choice occasion fixed effects that represent the mean utility an individual receives from accessing site j at choice occasion t after controlling for travel costs and the X_{ijt} variables. Many recreation demand papers include time-invariant mean utilities and refer to them as alternative- or site-specific constants.

Some existing studies decompose mean utilities using a vector of site attributes, such as water quality measures and fish catch rates. Decomposing mean utilities using a variety of attributes is also feasible in our model, but it may not fully leverage the mean utilities' panel variation. Some attributes will not vary meaningfully within a season and including many attributes may obscure sources of identifying variation.

Rather, our model's relative strength is that it can isolate quasi-experimental variation in site attributes. For example, we may seek to understand how some resource shock, like the beach closures from our application, impacts a site's mean utility. We could model the impact of a resource shock, indicated by D_{jt}, on site j's mean utility using a two-way fixed effects specification:

$$\delta_{jt} = \phi_j + \xi_t + \beta D_{jt} + \nu_{jt}, \tag{2}$$

where ϕ_j is a site fixed effect, ξ_t is a choice occasion fixed effect, D_{jt} equals one if site j is impacted at choice occasion t and zero otherwise, and ν_{jt} is an error term. The coefficient β captures the average effect of the attribute change on the impacted sites' mean utilities, an average treatment effect on the treated sites.[4]

After estimating the causal effect of the site attribute change on the mean utilities, we can then define counterfactual mean utilities—i.e., the mean utilities that each site would have provided if there were no resource shock:

$$\delta^0_{jt} = \begin{cases} \delta_{jt} - \beta & \text{if } j \in \mathcal{D} \ \& \ t \geq \tau \\ \delta_{jt} & \text{otherwise} \end{cases} \tag{3}$$

where $\mathcal{D}$ denotes the set of impacted sites and τ denotes the first post-shock choice occasion.

We assume the error term, ϵ_{ijt}, follows a Generalized Extreme Value distribution that implies a two-nest structure, where one nest contains only the outside option and the second nest contains all the recreation sites included in the choice set. This assumption yields the

[4]For simplicity, we abstract from concerns regarding staggered treatment timing and treatment effect heterogeneity. In practice, researchers can specify equation 2 to leverage any causal inference technique appropriate for their empirical setting. Our two-way fixed effects model is simply one potential mean utility decomposition.

closed-form choice probabilities:

$$P_{ijt} = \begin{cases} \dfrac{\exp(V_{i0t})}{\exp(V_{i0t}) + \left[\sum_{j=1}^{J} \exp\left(\frac{V_{ijt}}{\lambda}\right)\right]^{\lambda}}, & \text{if} \;\; j = 0 \\ \dfrac{\exp\left(\frac{V_{ijt}}{\lambda}\right)}{\sum_{j=1}^{J} \exp\left(\frac{V_{ijt}}{\lambda}\right)} \dfrac{\left[\sum_{j=1}^{J} \exp(\frac{V_{ijt}}{\lambda})\right]^{\lambda}}{\exp(V_{i0t}) + \left[\sum_{j=1}^{J} \exp\left(\frac{V_{ijt}}{\lambda}\right)\right]^{\lambda}}, & \text{if} \;\; j \in \{1, ..., J\} \end{cases} \tag{4}$$

where $V_{ijt} \equiv \delta_{jt} + \alpha^{TC} TC_{ij} + \alpha^{X} X_{ijt}$ is the deterministic portion of the utility from equation 1. The dissimilarity coefficient, λ, captures the similarity of alternatives in the "visit" nest. The model is consistent with utility maximizing behavior when the dissimilarity coefficient lies between 0 and 1, and values closer to 1 indicate alternatives in the "visit" nest are more dissimilar (McFadden, 1978).

We calculate the welfare impact of the attribute change using the log-sum formula for compensating variation—the monetary compensation needed to equate an individual's expected maximum utility under observed and counterfactual resource conditions. The compensating variation is given by

$$CV_{it} = -\frac{1}{\beta^{TC}} \left[\ln \left(\exp(V_{i0}) + \left[\sum_{J=1}^{J} \exp(V_{ijt}^{1}/\lambda) \right]^{\lambda} \right) - \ln \left(\exp(V_{i0}) + \left[\sum_{j=1}^{J} \exp(V_{ijt}^{0}/\lambda) \right]^{\lambda} \right) \right] \tag{5}$$

where V_{ijt}^{1} represents the deterministic portion of utility under observed conditions, and V_{jt}^{0} represents the deterministic portion of utility under counterfactual conditions:

$$V_{jt}^{0} = \delta_{jt}^{0} + \alpha^{TC} TC_{ij} + \alpha^{X} X_{ijt}. \tag{6}$$

Our model assumes that the Generalized Extreme Value error terms are independent across individuals and choice occasions. This simplification abstracts from several plausible features of recreation behavior. For instance, it assumes that recreators do not have consistently high (or low) unobservable affinity for certain sites and that a recreator's unobservable affinity for a site yesterday is uncorrelated with their unobservable affinity for the site today. The precise temporal variation captured by many emerging datasets makes relaxing these assumptions and estimating dynamic models increasingly feasible. Yet, fully dynamic travel cost models are rare (see Provencher and Bishop, 1997, for one exception), and developing and estimating a fully dynamic model is beyond the scope of this paper.

Nonetheless, our method can be adapted to incorporate some aspects of dynamic behavior. Swait et al. (2004) describe discrete choice models capturing temporal dependence as a continuum. Models assuming independent choices over time and across individuals are at one extreme and fully dynamic models are at the other. Swait et al. and Moeltner and Englin (2004) both adopt "middle ground" approaches. They allow individuals' past decisions to affect their current choice, while maintaining the assumption that error terms are independent across individuals and choice occasions. Our model can be adapted to incorporate similar elements of dynamic behavior. Past decisions can be included as covariates in X_{ijt}, and the mean utility provided by a site can be written as a function of past conditions.

3 Estimation

We estimate the model in two stages. First, we estimate the parameters in equation 1 and the dissimilarity coefficient using maximum likelihood estimation. The log-likelihood of observing individual i's visitation history is

$$\mathcal{L}_i\left(\delta, \alpha, \lambda\right) = \sum_{t=1}^{T} \sum_{j=0}^{J} y_{ijt} \ln\left(P_{ijt}\right), \tag{7}$$

where P_{ijt} is defined by equation 4 and y_{ijt} equals 1 if individual i chooses alternative j at occasion t and 0 otherwise. We define α as a vector containing α^{TC} and α^{X} and δ as a vector of all the mean utilities. Summing $\mathcal{L}_i$ across individuals yields the objective function for the optimization routine.

Rather than estimating the mean utilities directly, we recommend applying the Berry (1994) contraction mapping to solve for the unique mean utilities that match the daily visitation shares predicted by the model to the daily visitation shares observed in the data. Estimation with the contraction mapping produces same estimates as maximum likelihood estimation with respect to δ, α, and λ, but it allows the optimization routine to search along fewer dimensions. Even with a relatively small choice set, the model in our application contains over 700 mean utility parameters. Therefore, the contraction mapping substantially reduces the computational burden of the maximum likelihood estimation. After the first stage, we possess estimates of the α coefficients from equation 1, the dissimilarity coefficient, λ, and the mean utilities, δ.

The estimated mean utilities capture all observable and unobservable site attributes, which means they prevent unobservable site attributes from biasing the other first stage coefficient estimates. Applying the contraction mapping requires normalizing the mean utility provided by the outside option to equal zero on each date. Thus, raw mean utility estimates should be interpreted relative to that day's outside option. On days when visiting recreation sites becomes more appealing relative to the outside option, such as weekends or holidays, mean utilities will increase.

In the second stage, we estimate the causal effect of some event, or site attribute change, on the mean utilities. Again, researchers can apply any panel data econometric technique to achieve this goal, including difference-in-differences, event study designs, or the synthetic control method. Continuing with the difference-in-differences example from equation 2, we would estimate mean utilities simply by regressing the first stage mean utility estimates on site indicator variables, choice occasion indicator variables, and a treatment indicator variable. The identification assumptions required to consistently estimate second-stage parameters are analogous to those required in reduced-form settings. Difference-in-differences and event study designs require some form of the parallel trends assumption, and just like in reduced-form settings, researchers can assess the plausibility of a parallel trends assumption by checking for pre-trends.

Identification also requires a stable unit treatment value assumption (SUTVA), which can be violated if the treatment spills over and impacts control sites. To understand SUTVA in the context of our second-stage regression, it is helpful to note that our second-stage regression differs from a reduced-form regression where visitation is the outcome. If we regress visitation on some variable indicating the site attribute change, then SUTVA is likely to be violated, because the attribute change will cause visitors to substitute to or from other sites, which serve as controls. By embedding causal inference techniques within a RUM, we explicitly model inter-site substitution. Mean utilities, not visitation, serve as the outcome variable in the second stage. Thus, in the context of our second-stage regression, SUTVA requires that an attribute change at one site does not impact mean utilities at control sites. This version of SUTVA is more likely to hold than the reduced-form analog, because mean utilities at control sites are determined by their own attributes, rather than attributes at other sites. Nonetheless, SUTVA should not be ignored completely. For example, it is possible that a change at one site could affect congestion at control sites or lead visitors to visit new sites, learn, and update their preferences. The importance of these concerns will depend on the empirical setting.

4 Application

4.1 Background & Data

We demonstrate our method by estimating the recreational welfare losses caused by beach closures at Lake St. Clair Metropark during the summer of 2022. Lake St. Clair Metropark lies on the shores of Lake St. Clair, about 20 miles northeast of Detroit. It is part of the Huron-Clinton Metroparks system, a collection of thirteen parks throughout southeast Michigan. Like many of these metroparks, Lake St. Clair features a variety of recreational amenities, including walking trails, kayak rentals, a swimming pool, and a mini-golf course. If the swimming beach is closed, visitors are free to enjoy these other amenities.

Our recreation data are administrative visitation records collected by Huron-Clinton Metroparks staff at park entrance booths. The data log the minute-of-entry for over 1 million visits between May 15 and October 15, 2022. They also contain each visitor's residential ZIP code, allowing us to estimate a travel cost model. To enter a park, visitors must display an annual pass ($40) or purchase a daily pass ($10). When visitors purchase a daily pass, the staff ask for their home ZIP code and the payment system records the time of entry. The annual pass is a windshield sticker. Park staff scan a barcode on the sticker, which documents both the time of entry and the pass identification number. Thus, we observe a detailed visitation history for annual passholders—i.e., diary-style data. We restrict our focus to annual passholders and drop all daily pass visits to simplify the exposition of our method. This restriction preserves around 740,000 visits made by 115,000 annual passholders (about 70 percent of all visits).

The accuracy of these visitation records depends on reliable entry booth staffing. To check this reliability, we identify all "reported zeros"—days where we do not observe any visitors to a park. These days most likely result from unstaffed entry booths rather than zero attendance. Appendix figure B.1 shows that three parks regularly report zero visitation and that reported zeros become more common from late-September through mid-October. Therefore, we focus our analysis from May 15 to September 19, 2022, and we keep the six parks with no reported zeros during this time period as distinct alternatives in the choice set. We group the other seven parks with the outside option.

The metroparks are popular recreation sites in southeast Michigan. Stony Creek, Kensington, and Lake St. Clair Metroparks each received more than 100,000 visits from annual passholders between May 15 and October 15, 2022 (table B.1). Figure 1a shows metropark locations, as well as the frequency of visitation by ZIP code. ZIP codes with the highest visitation tend to be near the metroparks, often in the Detroit suburbs. Figure 1b shows how many times annual passholders visited any of the six parks in our choice set. The distribution of these visitation frequencies is skewed to the right. Passholders visited these parks 4.5 times on average, and 90 percent of passholders visited fewer than ten times. A subset of passholders visited much more frequently; 2,251 passholders (2 percent) made at least 30 visits.

One limitation of this dataset is that the visitation records do not contain any demographic information. In the absence of individual-level demographic variables, we assign each passholder the demographic information associated with their home ZIP code. Table 1 contains a full list of demographic variables.

Table 1 also shows statistics describing passholders' travel costs. Our travel cost calculations follow standard practices in the recreation demand literature. Specifically, we calculate passholder i's round-trip travel costs of accessing park j as

$$TC_{ij} = 2 \cdot \left(p^M \ Miles_{ij} + p_i^T \ Time_{ij}\right) \tag{8}$$

where $Miles$ represents the one-way driving distance in miles and $Time$ represents the one-way driving time in hours. The p^M term represents the per-mile, "out-of-pocket" cost of driving, which includes fuel, maintenance, and depreciation costs. The p_i^T term represents the opportunity cost of time. We calculate driving mileages and times from residential ZIP codes to metroparks using the Open Source Routing Machine, and we calculate a per-mile, "out-of-pocket" driving cost of 34.8 cents per mile using the 2022 AAA Your Driving Costs report. It is standard practice to assume an individual's opportunity cost of time is one-third of their hourly wage rate. Following this precedent, we assume a passholder's opportunity cost of time equals one-third of the median hourly wage rate in the their residential ZIP code. We calculate each ZIP code's median hourly wage rate by dividing its annual median household income by 2080 (= 52 weeks × 40 hours per week).

We augment these visitation data with water quality test results and beach closure records from the Michigan Department of Environment, Great Lakes, and Energy's (EGLE) Beach-

Guard monitoring program. EGLE coordinates water quality tests that track *E. coli* bacteria levels at 450 swimming beaches throughout Michigan. State law sets one-day and 30-day mean standards for bacteria levels, and a beach is closed to swimming if bacteria levels exceed either the one-day or 30-day mean threshold. County health officials test the water at Lake St. Clair on Mondays and Wednesdays from mid-April through the end of September. If bacteria levels exceed the daily threshold, the water is retested the following day. If retested levels still exceed the daily threshold, then testing returns to its normal Monday and Wednesday schedule. In the summer of 2022, elevated bacteria levels closed the Lake St. Clair Metropark beach for 33 days in all, and the first closure occurred July 21.

Potential visitors may learn of beach closures through several channels. EGLE posts water quality test results and announces beach closures and contamination advisories on the BeachGuard website, and local news outlets regularly publish lists of closed beaches. Closures are not rare—23 percent of monitored beaches experience a closure each year—so many visitors are aware of a beach's status before taking their trip (EGLE 2019).

The raw visitation data reveal some evidence of the Lake St. Clair beach closures (figure 2). Visitation rates from low travel cost deciles are slightly lower the two weekends just after the first closure than they are the two weekends just before the first closure. In the lowest travel cost decile, the difference in visitation rates amounts to about 75 fewer visits each day. There is no discernible difference between pre- and post-closure visitation rates at high travel cost deciles, at least in part because Lake St. Clair receives fewer visitors from these distant locations.

4.2 Applying the method

We model recreation demand for the metroparks using the repeated RUM with daily choice occasions presented in section 2. We apply our estimation procedure by first estimating daily mean utilities for each of the six parks in our choice set, then using the synthetic control method to estimate the causal effect of the beach closures on Lake St. Clair Metropark's mean utilities. The synthetic control method is frequently applied in settings with few treated and control units, making it well-suited for our setting (Abadie & Gardeazabal, 2003; Abadie, 2021).

We construct synthetic mean utilities for Lake St. Clair Metropark as a weighted average of other parks' mean utilities. We choose weights to minimize the error between the synthetic mean utilities and the observed mean utilities prior to the first beach closure. More formally, we select the weights by solving

$$\begin{aligned} &\arg\min_{w} \sum_{t<\tau} \left(\hat{\delta}_{1t} - \sum_{j=2}^{J} w_j \hat{\delta}_{jt} \right)^2 \\ &\text{subject to} \quad \sum_{j=2}^{J} w_j = 1 \end{aligned} \tag{9}$$

where Lake St. Clair Metropark is indexed by $j = 1$ and τ represents the date of the first beach closure. We assume these synthetic mean utilities represent Lake St. Clair's counterfactual mean utilities—that is,

$$\hat{\delta}^0_{1t} = \sum_{j=2}^{J} \hat{w}_j \hat{\delta}_{jt}$$

where $\hat{w}$ is estimated in equation 9. We calculate the aggregate welfare loss by summing equation 5 across all passholders and all dates after the first closure.

In addition to estimating the aggregate welfare loss, we also report the welfare loss per lost trip to Lake St. Clair. We calculate the number of lost trips to Lake St. Clair as

$$Lost\ Trips = \sum_{t \geq \tau} \left[\sum_{i} P_{i1t} \left(\hat{\delta}^0_{1t} \right) - Trips_{1t} \right]$$

where $Trips_{1t}$ indicates the number of observed trips to Lake St. Clair on date t.

4.3 Results

Table 2 reports first stage parameter estimates for all variables except the panel of mean utilities, which we omit for brevity. The travel cost coefficient is negative and statistically significant, as expected. The dissimilarity coefficient is between zero and one, implying the model is consistent with utility-maximizing behavior. We interact several demographic variables with the outside option, and the coefficient estimates on these interaction terms produce multiple findings. First, the income coefficients indicate that passholders from high income ZIP codes are more likely to visit a metropark than passholders from lower income ZIP codes, all else equal. Similarly, the race coefficients indicate that passholders from ZIP codes with a higher share of Black and other non-white residents are more likely to visit metroparks than passholders from ZIP codes with a higher share of white residents. On the other hand, passholders from ZIP codes with higher shares of Hispanic residents and higher shares of residents with bachelor's degrees, graduate degrees, and children are more likely to choose the outside option. These coefficients suggest the metroparks may be attractive outdoor recreation sites to underrepresented racial minorities and less educated people, which is potentially noteworthy for research focused on increasing outdoor recreation participation rates from these groups. However, we recommend interpreting these demographic coefficients with caution, because we lack individual-level demographic data and focus only on annual passholders.

Now, we explore how the beach closures impacted Lake St. Clair's mean utilities. Figure 3 compares Lake St. Clair's observed and synthetic mean utilities. Peaks tend to correspond to weekends and holidays, where visiting a park is more appealing relative to the outside option. The vertical line marks July 21, 2022, the date of the year's first beach closure. Synthetic mean utilities match observed mean utilities closely before the first closure, suggesting that synthetic Lake St. Clair provides reasonable estimates of counterfactual mean utilities. After the first closure, observed mean utilities are consistently lower than the synthetic mean utilities. Figure 4 shows this more clearly by plotting the gap between the observed and synthetic mean utilities. This gap represents the causal effect of the beach closures on Lake St. Clair's mean utilities.

The beach closures decrease Lake St. Clair's mean utility on many days after the first closure, even days when the beach is not closed. This persistent impact is particularly apparent after the August 4th beach closure. The mean utility provided by Lake St. Clair drops sharply during the closure and takes several weeks to return to the counterfactual level. This delayed recovery suggests that the beach closures cause a stigma effect. There could be several factors contributing to the stigma. It is possible that some recreators are not be aware that the beach has reopened, in which case more complete information on beach reopenings might help limit welfare losses. On the other hand, even recreators that know the beach has reopened may prefer to avoid a site that recently experienced elevated *E. coli* levels. Indeed, Boudreaux et al. (2023) find that recreators are hesitant to return to beaches even when they know the water is swimmable again. Using a discrete choice experiment, they estimate that recreators are willing to drive 77 miles, on average, to avoid a site that experienced elevated bacteria levels six days ago. Given their result, it seems likely that the persistent decline in mean utilities is not completely driven by imperfect information, but rather that imperfect information could exacerbate stigma effects. Investigating how water quality testing and communication can help mitigate stigma effects may be a fruitful path for future research.

We execute a placebo test to gauge the significance of the drop in Lake St. Clair's mean utilities. We estimate the impact of the Lake St. Clair beach closures on each control park's mean utilities, repeatedly creating synthetic control parks and plotting the gap between each control park's synthetic and observed mean utilities. Figure 5 displays the gaps for control parks in grey and the gap for Lake St. Clair (the same as in figure 4) in black. Although several control parks have short spells with mean utility gaps similar to Lake St. Clair's, no park exhibits such a sustained change in mean utilities. We confirm this visual observation by computing the root-mean-squared error (RMSE) between observed and synthetic mean utilities for each park before and after the first closure. Lake St. Clair's ratio of post-closure RMSE to pre-closure RMSE is 2.45, while control parks' RMSE ratios range between 0.71 and 1.96. These ratios suggest that the decrease in mean utilities at Lake St. Clair is more dramatic than the change in mean utilities at any other park, and thus, our results pass this placebo test. Abadie (2021) provides a test of statistical significance based on these RMSE ratios. With only five control parks, we cannot obtain a statistical significant p-value according to any common threshold, but we do obtain the smallest possible p-value for our empirical context, 1/6 or 0.17.

Figure 5 also helps to assess potential SUTVA violations. As discussed in section 3, spillovers that impact control sites' mean utilities may violate SUTVA and pose a threat to identification. If control parks' mean utilities were contaminated by spillover effects of the beach closure, then the park(s) experiencing the greatest spillovers would likely appear treated in figure 5. Yet, no park exhibits a consistent change in its mean utility after the beach closure other than Lake St. Clair. While this is not conclusive evidence that SUTVA holds, it does provide additional support for our identification strategy.

Lake St. Clair's decreased mean utilities translate to recreational welfare losses. Our two-stage synthetic controls approach estimates the aggregate welfare loss (summed across all days after the initial closure and all annual passholders) at \$152,500 and the loss per lost trip at \$14.40. Dividing the aggregate loss by the number of days when the beach is closed, 33, yields a welfare loss per closure day of \$4,600. We compare these estimates to those produced by several alternative models. The first uses our two-stage estimation approach but estimates counterfactual mean utilities using an event study design with weekly coefficients. Figure B.2 shows the coefficient plot from this second-stage regression, which reveals negative beach closure impacts consistent with our synthetic controls results. The two-stage event study approach produces an aggregate welfare loss estimate of \$144,700, about 5 percent smaller than the two-stage synthetic control estimate.

We also estimate the welfare impact of the closures using two more traditional, "baseline" models. These baseline models include a travel cost variable, a cross-section of park mean utilities, and an indicator variable equal to one when Lake St. Clair is impacted by a beach closure (the indicator variable equals zero at all dates for other parks). We use two definitions to indicate when Lake St. Clair is "impacted by a beach closure". The first, labeled "Baseline 1," sets the indicator variable equal to one only on days when the beach is closed. The second, labeled "Baseline 2," sets the indicator variable equal to one on all days after the first closure. Appendix A describes the baseline models in more detail.

Welfare loss estimates from the baseline models are sensitive to how we define the beach closure indicator variable. The aggregate welfare loss estimate from the Baseline 1 model is \$48,800, while the aggregate welfare loss estimate from the Baseline 2 model is \$139,700, nearly three times larger. Again, one benefit of our two-stage approach is that we must only specify the initial closure date. We allow the data to reveal when Lake St. Clair is impacted.

The aggregate welfare loss estimate from the Baseline 2 model is only 9 percent smaller than the two-stage synthetic control estimate. The fact that these approaches produce similar estimates is somewhat surprising, given that the Baseline 2 model averages over temporal variation in the mean utilities and the magnitude of beach closure impacts. Although this is just one comparison, the similarity of the Baseline 2 and two-stage estimates provides some degree of convergent validity for previous studies using the baseline model, at least when the timing of welfare impacts has been properly defined.

While total welfare loss estimates from the Baseline 2 and two-stage panel models are similar, the two-stage panel approach uncovers rich temporal variation in welfare impacts. This variation may be of interest to park managers and policy-makers, seeking to understand how a site's recreational value varies over time. For example, managers may find it useful to know the benefits of rushing to complete a project before a holiday weekend. Figure 6 shows that welfare losses grow quickly in late July and early August, corresponding to the time when the beach closures cause the largest decrease in mean utilities. Losses continue accumulating steadily until late August, they then grow more quickly and finally flatten off. The baseline models capture no heterogeneity in the timing of welfare impacts.

To check the plausibility of our results, we regress daily welfare loss estimates on the temperature, a rainy day indicator, and a weekend/holiday indicator (table 4). This regression is descriptive; we are no longer estimating causal effects or structural parameters. Welfare losses tend to increase with temperature. Daily welfare losses are around $1,000 larger when the temperature increases ten degrees, holding precipitation and the weekend indicator fixed. The sign of the rainy day coefficient depends on the second-stage model. Welfare losses from the two-stage event study model tend to be lower on rainy days, matching our prior expectation, but welfare losses from the two-stage synthetic control model are actually larger on rainy days. Although, the coefficient estimate has a large standard error. The uncertainty in this estimate could be caused by our simplistic rainy day indicator variable, which does not capture variation in the time of day, duration, or intensity of rain showers. Finally, both models find that welfare losses are larger on weekends and holidays than on weekdays, holding weather conditions fixed. Overall, these results are fairly intuitive. A swimming beach likely provides more surplus for more recreators on a warm, sunny Saturday than a cool Tuesday, and so it is reasonable that the beach closures cause larger welfare losses on warmer days and weekends.

5 Conclusion

This paper provides a method that combines the desirable properties of both panel data causal inference techniques and structural demand modeling. Our approach allows researchers to value changes in site attributes by embedding difference-in-differences and synthetic control methods within a travel cost RUM. It complements existing demand modeling and estimation methods that tend to rely on cross-sectional variation or instrumental variables.

We illustrate our method by valuing the recreational welfare losses caused by beach closures at Lake St. Clair Metropark. Our method produces similar aggregate welfare loss estimates to a more traditional approach, but it requires fewer assumptions regarding the timing of welfare impacts and uncovers rich temporal variation. This temporal variation reveals that stigma effects cause welfare losses even after the beach reopens for swimming and that welfare losses tend to be larger on weekends and hotter days.

Our method may be appealing in many empirical contexts. In recreation demand settings, environmental conditions at a recreation site can change quickly and vary dramatically over time. Our method allows researchers to value the welfare impacts of environmental shocks, such as wildfires, infrastructure improvements, or the beach closures we study in our application. More broadly, travel cost models offer an appealing tool for leveraging the temporal variation in modern mobility data. Our method allows researchers to exploit this temporal variation without sacrificing the ability to model individual decision-making, control for the quality of substitute sites, and conduct welfare analysis.

References

Abadie, A. (2021). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, *59*, 391–425. https://doi.org/10.1257/jel.20191450

Abadie, A., & Gardeazabal, J. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review*, *93*(1), 113–132. https://doi.org/10.1257/000282803321455188

Angrist, J. D., & Pischke, J.-S. (2010). The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics. *Journal of Economic Perspectives*, *24*(2), 3–30. https://doi.org/10.1257/jep.24.2.3

Berry, S. (1994). Estimating Discrete-Choice Models of Product Differentiation. *The RAND Journal of Economics*, *25*, 242–262. https://www.jstor.org/stable/2555829

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, *63*, 841–890. https://www.jstor.org/stable/2171802

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

Christensen, P., & Osman, A. (2023). *The Demand for Mobility: Evidence from an Experiment with Uber Riders* (Working Paper No. 31330). National Bureau of Economic Research. https://doi.org/10.3386/w31330

Currie, J., Kleven, H., & Zwiers, E. (2020). Technology and Big Data Are Changing Economics: Mining Text to Track Methods. *AEA Papers and Proceedings*, *110*, 42–48. https://doi.org/10.1257/pandp.20201058

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. https://doi.org/10.1111/j.1467-8276.2008.01182.x

English, E., von Haefen, R. H., Herriges, J., Leggett, C., Lupi, F., McConnell, K., Welsh, M., Domanski, A., & Meade, N. (2018). Estimating the value of lost recreation days from the Deepwater Horizon oil spill. *Journal of Environmental Economics and Management*, *91*, 26–45. https://doi.org/10.1016/j.jeem.2018.06.010

Gellman, J., Walls, M., & Wibbenmeyer, M. (2023). Welfare losses from wildfire smoke: Evidence from daily outdoor recreation data. *RFF Working Paper 23-31*.

Goldsmith-Pinkham, P. (2024). Tracking the Credibility Revolution across Fields. https://arxiv.org/abs/2405.20604

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. https://doi.org/10.1086/708662

Keiser, D. A. (2019). The Missing Benefits of Clean Water and the Role of Mismeasured Pollution. *Journal of the Association of Environmental and Resource Economists*, *6*(4), 669–707. https://doi.org/10.1086/703254

Keiser, D. A., & Shapiro, J. S. (2018). Consequences of the Clean Water Act and the Demand for Water Quality. *The Quarterly Journal of Economics*, *134*(1), 349–396. https://doi.org/10.1093/qje/qjy019

Knittel, C. R., Li, J., & Wan, X. (2023). *I love that dirty water? Value of water quality in recreation sites* (Working Paper).

Kreindler, G. E., & Miyauchi, Y. (2023). Measuring Commuting and Economic Activity inside Cities with Cell Phone Records. *The Review of Economics and Statistics*, *105*(4), 899–909. https://doi.org/10.1162/rest_a_01085

Kremer, M., Leino, J., Miguel, E., & Zwane, A. P. (2011). Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions. *The Quarterly Journal of Economics*, *126*(1), 145–205. https://doi.org/10.1093/qje/qjq010

Kuwayama, Y., Olmstead, S., & Zheng, J. (2022). A more comprehensive estimate of the value of water quality. *Journal of Public Economics*, *207*, 104600. https://doi.org/10.1016/j.jpubeco.2022.104600

Lew, D. K., & Larson, D. M. (2005). Valuing Recreation and Amenities at San Diego County Beaches. *Coastal Management*, *33*(1), 71–86. https://doi.org/10.1080/08920750590883079

Lupi, F., Phaneuf, D., & von Haefen, R. (2020). Best Practices for Implementing Recreation Demand Models. *Review of Environmental Economics and Policy*, *14*, 302–323. https://doi.org/10.1093/reep/reaa007

McFadden, D. (1978). Behavioural Travel Modelling. In D. A. Hensher & P. R. Stopher (Eds.). Croom Helm London.

Michigan Department of Environment, Great Lakes, and Energy. (2024a). *Beach Water Monitoring*. https://www.michigan.gov/egle/about/organization/water-resources/beaches

Michigan Department of Environment, Great Lakes, and Energy. (2024b). *Michigan Beach-Guard System*. https://www.egle.state.mi.us/beach/Default.aspx

Michigan Department of Environment, Great Lakes, and Energy (EGLE) – Water Resources Division. (2019). Michigan beach monitoring year 2018 annual report. https://www.michigan.gov/egle/about/organization/water-resources/beaches

Moeltner, K., & Englin, J. (2004). Choice Behavior Under Time-Variant Quality. *Journal of Business & Economic Statistics*, *22*(2), 214–224. https://doi.org/10.1198/073500104000000109

Murdock, J. (2006). Handling unobserved site characteristics in random utility models of recreation demand. *Journal of Environmental Economics and Management*, *51*, 1–25. https://doi.org/10.1016,j-jeem.2005.04.003

Murray, C., Sohngen, B., & Pendleton, L. (2001). Valuing water quality advisories and beach amenities in the Great Lakes. *Water Resources Research*, *37*(10), 2583–2590. https://doi.org/https://doi.org/10.1029/2001WR000409

Nevo, A. (2001). Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica*, *69*(2), 307–342.

Newbold, S. C., Lindley, S., Albeke, S., Viers, J., Parsons, G., & Johnston, R. (2022). Valuing Satellite Data for Harmful Algal Bloom Early Warning Systems. *RFF Working Papers*.

Parsons, G. R., Kang, A. K., Leggett, C. G., & Boyle, K. J. (2009). Valuing Beach Closures on the Padre Island National Seashore. *Marine Resource Economics*, *24*(3), 213–235. https://doi.org/10.1086/mre.24.3.42629652

Phaneuf, D. J., Kling, C. L., & Herriges, J. A. (2000). Estimation and Welfare Calculations in a Generalized Corner Solution Model with an Application to Recreation Demand.

The Review of Economics and Statistics, *82*(1), 83–92. https://doi.org/10.1162/003465300558650

Provencher, B., & Bishop, R. C. (1997). An Estimable Dynamic Model of Recreation Behavior with an Application to Great Lakes Angling. *Journal of Environmental Economics and Management*, *33*(2), 107–127. https://doi.org/10.1006/jeem.1996.0978

Roberts, D. C., Boyer, T. A., & Lusk, J. L. (2008). Preferences for environmental quality under uncertainty. *Ecological Economics*, *66*(4), 584–593. https://doi.org/https://doi.org/10.1016/j.ecolecon.2008.05.010

Swait, J., Adamowicz, W., & van Bueren, M. (2004). Choice and temporal welfare impacts: incorporating history into discrete choice models. *Journal of Environmental Economics and Management*, *47*(1), 94–116. https://doi.org/10.1016/S0095-0696(03)00077-9

Timmins, C., & Murdock, J. (2007). A revealed preference approach to the measurement of congestion in travel cost models. *Journal of Environmental Economics and Management*, *53*(2), 230–249. https://doi.org/10.1016/j.jeem.2006.08.002

United States Environmental Protection Agency. (n.d.). *EPA's Beach Report: 2023 Swimming Season*. Retrieved August 2, 2024, from https://ordspub.epa.gov/ords/beacon2/f?p=BEACON2:DNR

Wolf, D., Chen, W., Gopalakrishnan, S., Haab, T., & Klaiber, H. A. (2019). The Impacts of Harmful Algal Blooms and *E.coli* on Recreational Behavior in Lake Erie. *Land Economics*, *95*(4), 455–472.

A Baseline Model

We compare welfare estimates produced by our model to those produced by a more traditional, "baseline" approach. This baseline approach largely follows English et al. (2018), who value the welfare impacts of the *Deepwater Horizon* oil spill. This appendix provides more details on the baseline model and estimation procedure.

Assume the choice set and choice occasions are the same as in section 2, but specify the utility individual i receives from visiting site j at choice occasion t as

$$U_{ijt} = \delta_j + \beta Closed_{jt} + \alpha^{TC} TC_{ij} + \alpha^X X_{ijt} + \epsilon_{ijt} \tag{10}$$

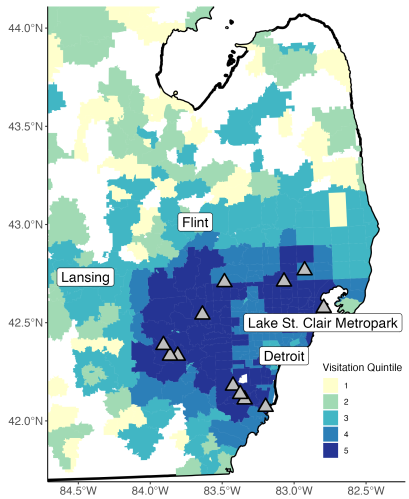
There are two differences between this specification and our preferred specification (equation 1). First, this model includes a cross-section of alternative-specific mean utilities, δ_j, rather than a panel. Second, this model includes the *Closed* indicator variable, which equals to one on days when site j is impacted by a beach closure. All other parameters and variables are defined as in equation 1.

We estimate all parameters in the baseline model using maximum likelihood estimation. Because the beach closure coefficient can be identified separately from site mean utilities, there is no need to apply the two-stage estimation procedure we use for our primary model. Furthermore, given the relatively small choice set in our empirical setting, it is computationally efficient to estimate the cross-section of mean utilities directly, rather than apply a contraction mapping to match predicted and observed visitation shares for the entire sample period.

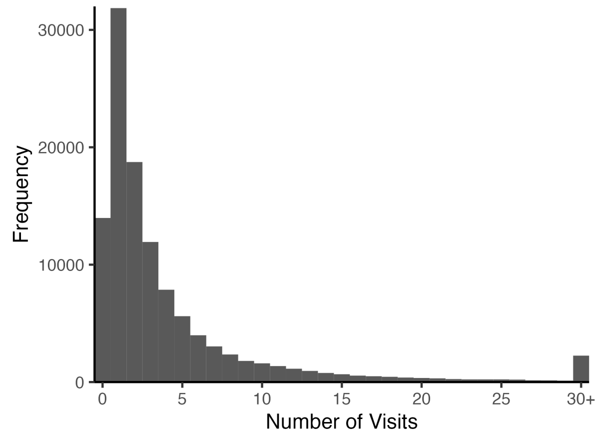
Ex ante, it is unclear how to define $Closed_{jt}$ for Lake St. Clair. If the closures cause a stigma effect, then Lake St. Clair would be impacted even on days when its beach is open. However, it may be difficult to detect this stigma effect before estimating the model. Therefore, we choose two ways of defining the $Closed_{jt}$ variable. First, we set $Closed_{jt}$ equal to one only on days when the beach at Lake St. Clair is closed. We call the results produced by this specification the "Baseline 1" estimates. Second, we set $Closed_{jt}$ equal to one every day after the initial beach closure. We call the results produced by this specification the "Baseline 2" estimates.

Figure 1: Visitation frequencies by ZIP code and passholder

(a) Visitation by ZIP Code



(b) Number of visits by annual passholder



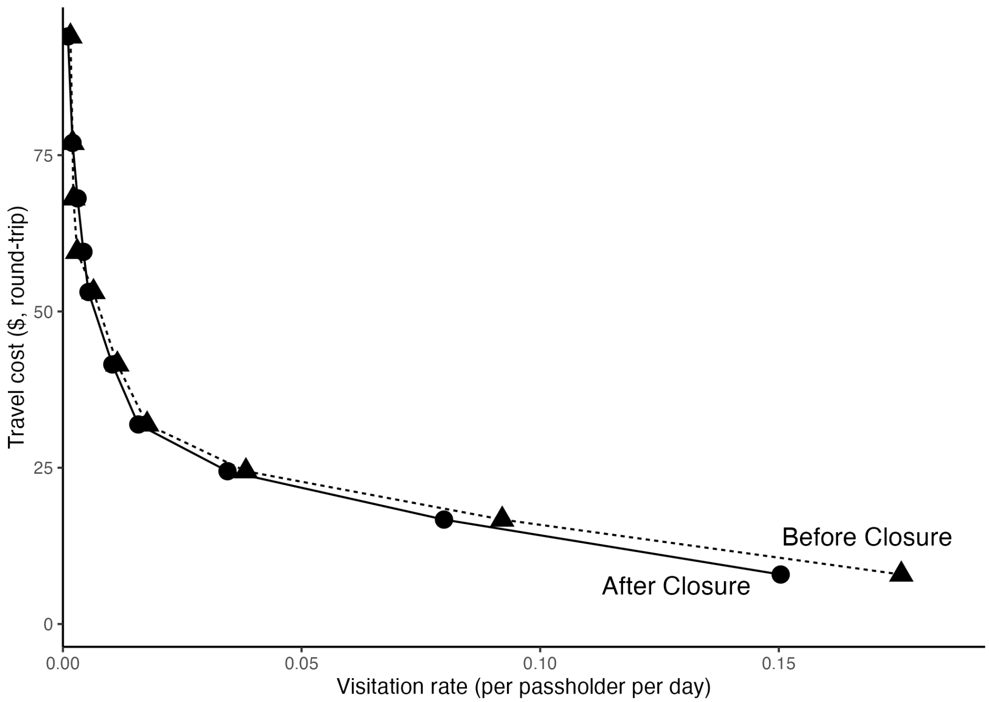
Note: Panel (a) shows Metroparks (grey-triangles) and visitation by ZIP code (darker shading indicates higher visitation). High visitation ZIP codes tend to be nearby Metroparks, often in the Detroit suburbs. Panel (b) shows how many times annual passholders visited one of the six parks in our choice set. The mass of annual passholders with zero visits visited at least one of the parks we omit from the choice set.

Table 1: Descriptive Statistics

Variable	Mean	Std dev	Min	Max
Travel variables (683,820 passholder-park pairs)				
Driving distance (miles, one-way)	38.1	18.7	0.8	182.3
Driving time (hours, one-way)	0.87	0.38	0.04	3.85
Travel cost ($s, round-trip)	50.8	25.2	1.7	269.9
Demographic variables (113,970 passholders)				
Median household income	85,758	24,390	21,180	164,250
$50k < median household income < $75k	0.329	0.470	0	1
$75k < median household income < $90k	0.210	0.407	0	1
Median household income > $90k	0.418	0.493	0	1
Share Black	0.084	0.124	0	0.957
Share Hispanic	0.036	0.033	0	0.710
Share other non-white race	0.099	0.071	0	0.506
Share with bachelor's degree	0.227	0.077	0.023	0.417
Share with graduate degree	0.158	0.100	0	0.531
Share with child	0.400	0.053	0	0.621

Note: Descriptive statistics for travel variables exclude the outside option. All demographic variables are ZIP code-level statistics from the 2021 American Community Survey 5-year data. We assign each individual the demographic information associated with their home ZIP code. We split race into three categories, white, Black, and other, and we classify ethnicity as either Hispanic or non-Hispanic. Share with bachelor's degree represents the share of the population whose highest degree is a bachelor's degree.

Figure 2: Lake St. Clair visitation rates by travel cost decile



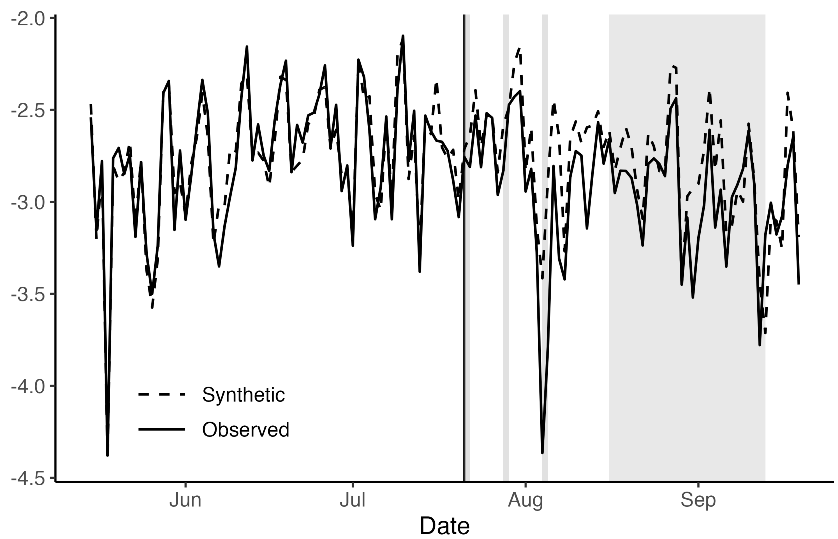
Note: Figure 2 shows visitation rates on the two weekends immediately before and the two weekends immediately after the first beach closure. Each point represents the visitation rate (number of visits to Lake St. Clair per passholder per day) for a different travel cost decile. The vertical axis coordinate is the average travel cost for each decile.

Table 2: First Stage Estimates

	Estimate	Standard Error
Travel cost ($10s)	-0.4406	0.0118
Dissimilarity coefficient	0.4373	0.0110
Interacted with outside option		
$50k < median household income < $75k	-0.3123	0.0029
$75k < median household income < $90k	-0.3460	0.0024
Median household income > $90k	-0.6031	0.0030
Share Black	-0.0699	0.0007
Share other non-white race	-0.2717	0.0004
Share Hispanic	0.2877	0.0002
Share with bachelor's degree	1.1612	0.0004
Share with graduate degree	0.5516	0.0005
Share with child	0.4444	0.0003
Log-likelihood	-2,492,800.4	

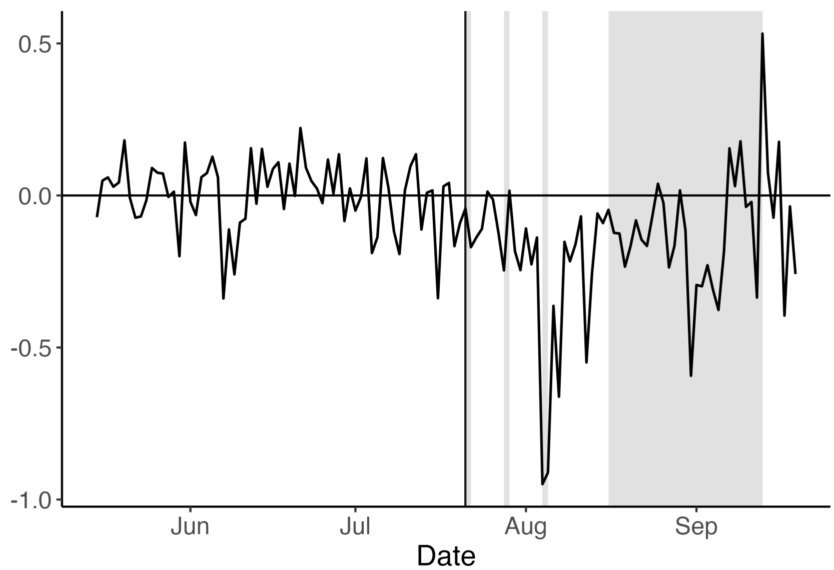
Note: The table shows parameter estimates and standard errors from the first-stage, maximum-likelihood estimation. The omitted race category is white, and the omitted ethnicity category is non-Hispanic. Education variables represent the share of ZIP code residents whose highest educational attainment is a bachelor's degree or a graduate degree. The omitted education category is "less than a bachelor's degree".

Figure 3: Observed and Sythetic Mean Utilities



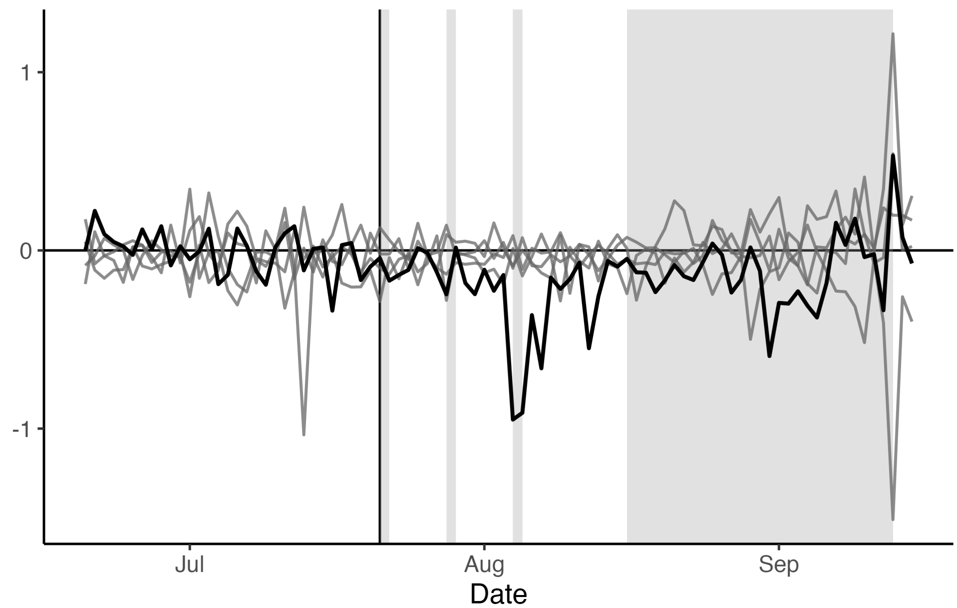
Note: The figure shows observed (solid) and synthetic (dashed) mean utilities for Lake St. Clair Metropark. The synthetic mean utilities represent the mean utilities that Lake St. Clair would have provided in absence of the beach closures. Grey shading indicates dates when Lake St. Clair experienced a beach closure.

Figure 4: The gap between observed and synthetic mean utilities



Note: The figure shows the gap between the observed and synthetic mean utilities provided by Lake St. Clair Metropark. This gap represents the causal effect of the closures on Lake St. Clair's mean utilities. Grey shading indicates dates when Lake St. Clair experienced a beach closure.

Figure 5: Placebo test



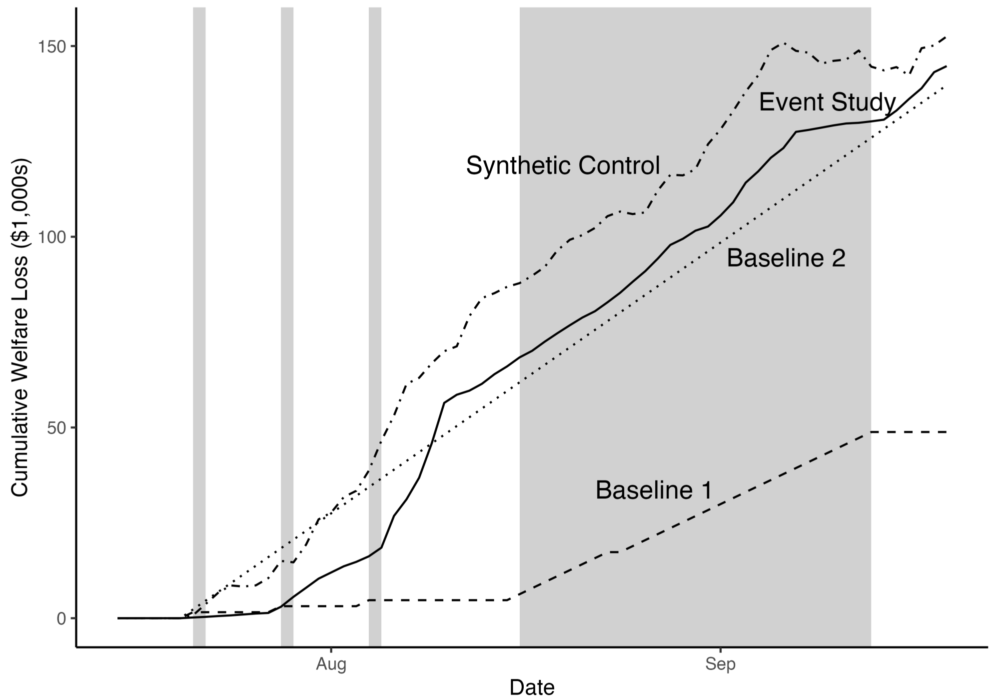
Note: The figure shows the gaps between synthetic and observed mean utilities for Lake St. Clair (black) and each control park (grey). Because the mean utilities provided by the control parks are not impacted by the beach closure at Lake St. Clair, the gaps between synthetic and observed mean utilities for control parks serve as placebos. It is reassuring that no other parks' mean utilities appear impacted by the closure at Lake St. Clair. We remove the gap between synthetic and observed mean utilities for Stony Creek Metropark, because its mean utility is consistently higher than all other parks. Therefore, Stony Creek's mean utilities are poorly predicted by a convex combination of other parks'. Stony Creek has a pre-closure root-mean-squared error four times larger than Lake St. Clair's and more than twice as large as any other park.

Table 3: Welfare Losses by Model

	Baseline 1	Baseline 2	Synthetic Control	Event Study
Aggregate loss ($1,000s)	48.8	139.7	152.5	144.7
Loss per closure day ($)	1,479	4,234	4,622	4,386
Loss per day after initial closure ($)	800	2,291	2,500	2,373
Loss per lost trip ($)	14.82	14.91	14.40	14.69

Note: The table shows estimates of the welfare loss caused by the 2022 Lake St. Clair Metropark beach closures for different models. The "Aggregate Loss" row shows estimates of the total welfare loss. "Loss per closure day" divides the aggregate loss by the number of days where the beach was closed, 33. "Loss per day after initial closure" divides the aggregate loss by the number of days after the first beach closure. "Loss per lost trip" divides the aggregate loss by the number of lost trips to Lake St. Clair Metropark, which varies by model.

Figure 6: Cumulative welfare loss across different model specifications



Note: The figure shows the cumulative welfare losses caused by the Lake St. Clair Metropark beach closures estimated by several different models. "Synthetic control" (dot-dash) represents our preferred two-stage synthetic control estimates. "Event Study" (solid) represents the two-stage event study estimates. "Baseline 1" (dashed) represents the traditional approach that assumes Lake St. Clair is impacted only when its beach is actually closed, and "Baseline 2" represents the traditional approach that assumes Lake St. Clair is impacted every day after the first closure. Grey shading indicates dates when Lake St. Clair experienced a beach closure.

Table 4: Regressing welfare losses on weather and a weekend indicator variable

	Synthetic Control	Event Study
Constant	1,698.0***	2,440.4***
	(429.2)	(337.0)
Daily max temperature (F)	115.6	99.2
	(80.7)	(63.3)
Rainy day	527.7	-1,447.1*
	(926.4)	(727.4)
Weekend or holiday	2,297.9***	544.8
	(708.9)	(556.6)
Observations	61	61
Adjusted R^2	0.139	0.051

Note: The table shows estimates from regressions of daily welfare losses on a weekend/holiday indicator variable and weather variables. The first column uses welfare loss estimates produced by applying the synthetic control method in the second stage, and the second column uses welfare loss estimates produced by applying an event study with weekly coefficients in the second stage. "Daily max temperature (F)" is the daily maximum temperature in degrees Fahrenheit. "Rainy Day" equals one if there was more than 0.1 inches of rain and zero otherwise. "Weekend or holiday" equals one on weekends and holidays (Labor Day) and zero otherwise.

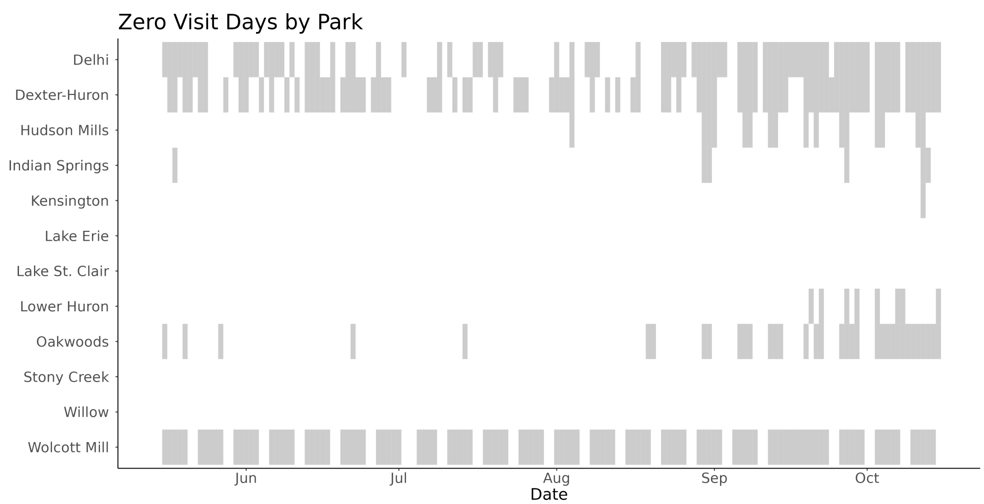
B Online Appendix: Supplemental Figures and Tables

Table B.1: Visitation by Metropark 2022

Metropark	Visitation	Beach	Kayak Rental	Pool or Splash Pad	Hiking
Stony Creek	201,882	Y	Y		Y
Kensington	197,122	Y	Y	Y	
Lake St. Clair	117,376	Y	Y	Y	
Lower Huron	56,678		Y	Y	Y
Willow	52,218			Y	Y
Lake Erie	37,722			Y	Y
Hudson Mills	35,629		Y	Y	Y
Indian Springs	27,280			Y	Y
Oakwoods	5,175		Y		Y
Dexter-Huron	4,411		Y		Y
Delhi	1,880		Y		
Wolcott Mill	785				Y

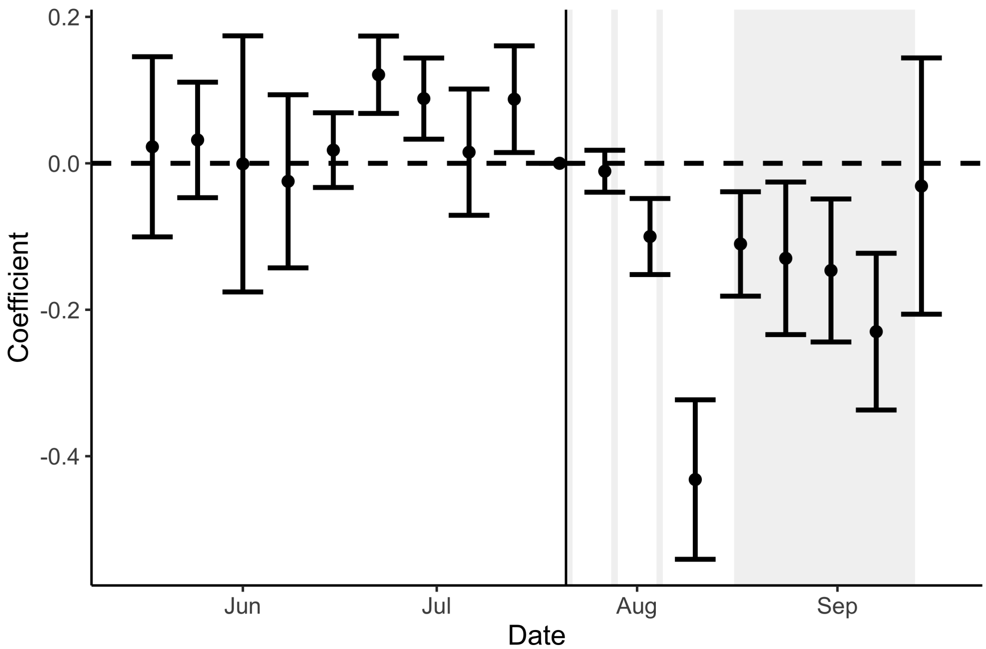
Note: The table shows total visitation by annual passholders between May 15 and October 15, 2022 and whether each park has a few of the common amenities. We do not observe visitation records for Huron Meadows Metropark, and we exclude it from the table.

Figure B.1: Zero visit days



Note: The figure shows all zero visit days shaded in grey – i.e., days where a park recorded no visitors. This indicates days when entry booths were not staffed, rather than days when parks truly received zero visits. We group parks with zero visit days during our sample period (May 15 through September 19) with the outside option. We include Kensington, Lake Erie, Lake St. Clair, Lower Huron, Stony Creek, and Willow Metroparks, which all report no zero visit days in our sample period, as distinct alternatives in the choice set.

Figure B.2: Second stage event study coefficient plot



Note: The figure shows coefficient estimates from a second stage event study regression, where mean utilities are the outcome variable. A separate coefficient is estimated for each week relative to the initial beach closure. Grey shading indicates dates when Lake St. Clair experienced a beach closure.

Table B.2: Weights for Synthetic Lake St. Clair Metropark

Park	Weight
Kensington Metropark	0.003
Lake Erie Metropark	0.040
Lower Huron Metropark	0.462
Stony Creek Metropark	0.475
Willow Metropark	0.020