Visiting America's Best Idea: Demand for the U.S. National Park System

Andrew Earle*

November 2022

Abstract

The U.S. National Park System protects some of the world's most spectacular resources and attracts 300 million visits each year, generating surplus for visitors and supporting local economies. I create a versatile and unified framework to analyze demand for U.S. national parks. Combining nationally representative surveys with park-level visitor counts, I estimate a discrete-choice model of visitation from 2005 through 2019. The model controls for changing travel costs and inter-park substitution while permitting the use of causal inference techniques. I apply the framework to analyze how long-run average temperatures and short-run temperature shocks impact demand. Visiting a park generates the most surplus when average high temperatures fall between 70°F and 85°F. Relative to this ideal range, visiting when the temperature is 30°F reduces willingness to pay by $503 per-trip, while visiting when the temperature is 95°F reduces willingness to pay by just $107. Positive temperature shocks increase willingness to pay when temperatures are less than 80°F and have no discernible impact at warmer temperatures. The results provide insight to the welfare impacts of climate change, and the framework is broadly applicable to management challenges facing the national parks.

*Department of Economics, Michigan State University (earleand@msu.edu)
I thank Leslie Richardson for providing data. I thank Soren Anderson, Frank Lupi, Roger von Haefen, Nick Parker, Randy Rucker, Margaret Walls, Arthur Wardle, and seminar participants from Michigan State University, Camp Resources, the National Park Service, and PERC for their insightful comments.

1 Introduction

In 1916, the National Park Service was created to conserve the United States' most significant sites, scenery, and wildlife ("Organic Act of 1916"). More than 100 years later, the National Park System includes over 400 parks, which collectively attract 300 million visits each year. The parks generate surplus for visitors and support nearby gateway communities (Cullinane Thomas, Flyr, and Koontz 2022). They are so quintessentially American that Ken Burns featured them in a documentary titled, "America's Best Idea."

Conserving important and popular resources creates management challenges. At the most famous parks, large crowds create air pollution that rivals metropolitan areas (Keiser, Lade, and Rudik 2018). In rare cases, parks have turned away visitors because of overcrowding. Crowds also exacerbate challenges caused by stagnant federal funding, contributing to the National Park Service's $21.8 billion deferred maintenance backlog ("NPS 2022"). Recent legislation to boost funding therefore provides critical financial support.[1] Meanwhile, climate change has already begun to alter the resources the National Park Service was created to protect. Sea level rise, wildfires, drought, and extreme weather events, amplified by climate change, pose threats to the United States' most treasured resources. Understanding how these challenges impact visitation aligns with the National Park Service's core mission of conservation for the enjoyment of present and future generations.

This paper creates a versatile and unified framework to analyze demand for the U.S. National Park System. I introduce a random utility maximization model of visitation for 140 national parks, all those protected for their natural resources, in which individuals repeatedly choose which park to visit and whether to drive or fly.[2] The model includes a full set of park-by-month fixed effects. I call these parameters park effects, and they capture the mean utility of visiting a park while controlling for the travel costs needed to get there. In

[1] The Great American Outdoors Act (2020) and the Bipartisan Infrastructure Law (2022) both provide substantial funding to the National Park System.

[2] The system also includes historic sites that I do not explicitly model (e.g. Theodore Roosevelt Birthplace National Historic Site).

plain terms, they provide an index of national park *awesomeness*.

I introduce a two-step estimation procedure that combines two nationally representative telephone surveys, a fifteen-year monthly panel of park-level visitor counts, and a rich collection of statistics describing park attributes. The first step uses maximum likelihood estimation and the contraction mapping introduced by Berry (1994) to combine the survey and visitor count data. In effect, it filters the visitor counts through the structural model to produce a monthly panel of park effects from 2005 through 2019. In the second step, I regress the park effects on park attributes. I leverage variation both within and between parks to estimate willingness to pay for a variety of attributes, including elevation, land cover, and wildlife presence. Because the first step preserves the panel structure of the visitor counts, causal inference techniques, such as difference-in-differences and event-study designs, can be applied within the structural model.

I demonstrate the versatility of the framework by estimating the impact of temperature on demand. I decompose the temperature impact into long-run differences and short-run shocks. Decomposing the temperature impact captures responses to both gradual warming and the increased frequency of heat waves and cold snaps expected from climate change. I measure long-run differences using monthly ten-year moving averages, which roughly reflect the expected temperature at a park if a visitor were planning their trip more than a few weeks in advance. I measure short-run shocks using the difference between realized temperatures and the monthly averages. By including a flexible set of fixed effects, I identify the impact of long-run differences using within-season variation in a park's average temperature. I identify the impact of short-run shocks using the same within-season variation, while controlling for average temperature.

I find that "bucket list" parks such as Yellowstone, Glacier, and Grand Canyon consistently rank in the top ten of the national parks *awesomeness* index. Observable attributes explain 56% of the variation in this index. Visitors tend to prefer parks with redwood forests, bison, bald eagles, wide-ranging elevation, and shoreline. These attributes vary little across time, posing a challenge for causal inference. Yet, my estimated park effects reveal previously unexplored seasonal variation, making a causal interpretation more plausible for attributes that vary across time. For parks with harsh winters, willingness to pay peaks in the summer months, while parks with more moderate climates provide more stable mean utility throughout the year.

Applying the framework, I find that visiting a park generates the most surplus when average high temperatures fall between 70°F and 85°F. Outside of this range, cold reduces willingness to pay more than heat, and long-run differences impact willingness to pay more than short-run shocks. Compared to the ideal temperature of 70°F, visiting when the temperature is 30°F reduces average willingness to pay by $503, while visiting when the temperature is 95°F reduces average willingness to pay by just $107. Meanwhile, when the temperature is less than 80°F, a positive 5°F shock increases average willingness to pay and has the greatest impact between 40°F and 50°F. I find no evidence that shocks decrease willingness to pay at temperatures above 80°F. Although, I cannot rule out large impacts due to my imprecise estimates. Overall, my results imply that climate change could increase welfare, since visitors dislike extreme cold more than they dislike extreme heat. Of course, this inference holds other park attributes fixed and therefore ignores the potential degradation in park ecology (e.g., species loss) that could result from climate change.

Many previous studies analyze recreational demand for the National Park System, including several that are national in scope. The closest studies to mine use panel data to model visitation. Szabó and Ujhelyi (2021) study how national parks affect a broad suite of local economic outcomes. They find that a national park designation increases visitation by 17% and increases local economic activity along many dimensions.[3] Like me, two studies analyze the impact of temperature on visitation. Henrickson and Johnson (2013) study visitation to 56 national parks using annual data, which fails to capture seasonal differences

[3]Each National Park Service unit carries an official designation (e.g., National Park, National Lakeshore, National Monument). All units are commonly referred to as "national parks," and I refer to all units as parks throughout the paper.

in temperature. Fisichelli et al. (2015) study visitation to 340 parks using monthly data averaged across 35 years, which fails to capture the effect of short-run shocks. My study captures both long-run differences and short-run shocks to monthly temperatures. Additionally, these papers all use reduced-form approaches, which suffer from several limitations in this setting. First, visitors may substitute between parks in response to park closures or changes in attributes, including temperature. In this case, substitute parks are effectively "treated" and are no longer valid controls. I address this concern by modeling the full park demand system, which captures the availability and quality of substitute parks. Second, geographic shifts in population may impact parks differently, violating the common-trends assumption and confounding changes in park attributes. The U.S. population has shifted south and west over the past two decades, closer to many national parks but farther from others, and has become increasingly concentrated in urban areas. I account for these changes by explicitly modeling park visits as a function of where people live in every year relative to every park. Finally, while park visitation is informative, it cannot be used directly to measure welfare losses from park closures, climate change, or other factors. My model provides a structure for calculating these welfare impacts. Like me, Neher, Duffield, and Patterson (2013) estimate willingness to pay for a large number of parks and then explain willingness to pay as a function of attributes in a second-stage regression. But they estimate demand using 58 separate count-data models and a single year of data at each park. In contrast, I estimate demand for 140 parks in a single unified framework, combine microdata and aggregate data, and exploit monthly panel data over fifteen years to generate causal estimates.

The recreation demand literature frequently employs random utility maximization models to estimate the value of recreation sites or environmental attributes (Egan et al. 2009; Kolstoe and Cameron 2017; G. Parsons et al. 2020). The estimation procedure I introduce is most closely related to the two stage approach introduced by Murdock (2006), which applies the contraction mapping from Berry (1994) to estimate a full cross-section of site fixed effects. Including a full cross-section of site fixed effects eliminates bias from omitted site characteristics in the first stage, but because Murdock's second stage is a cross-sectional regression, estimates of preferences for site attributes still suffer from omitted variables bias. I better control for omitted site attributes by leveraging panel data econometric techniques in my second stage. In a recent best practices paper, Lupi, Phaneuf, and von Haefen (2020) emphasize the need for more rigorous identification in recreation demand studies, and my procedure provides a versatile method for applying casual inference techniques in random utility maximization models.

By analyzing recreation demand at a national scale, my approach also builds on English et al. (2018), which quantifies the value of recreational losses from the Deepwater Horizon oil spill. The authors use a nationally representative survey with over 41,000 respondents and calculate travel costs of flying and driving. They calibrate their model using visitor counts to estimate preferences under both non-spill and spill conditions. My paper differs in several important ways. First, I generate a monthly panel of park effects, while English et al. estimate a single difference between non-spill and spill conditions. Estimating a monthly panel allows for site-specific and temporal heterogeneity. Second, I explain park effects using park attributes, allowing me to study the welfare impacts of changes in these attributes. Finally, I explicitly model the decision to fly or drive to a recreation site, while English et al. use a weighted average of flying and driving costs when calculating travel costs.

My approach also has substantial practical value. Individual survey efforts are costly, and they rarely occur continuously across multiple years. Yet, consistent visitation data is critical for understanding how visitors value changes in environmental amenities. I show how a structural model and visitor counts can be combined to make up for infrequent survey data. By calibrating my model with visitor counts and annual American Community Survey microdata, I estimate responses to park attribute changes outside the survey period and bridge the gap between costly individual-level surveys.

Finally, my paper contributes to a growing literature valuing the nonmarket impacts of climate change, which are critical and understudied inputs to the social cost of carbon

(Burke et al. 2016). Recent research has valued climate impacts on crime (Hsiang et al. 2017), mortality (Carleton et al. 2022), and recreation (Chan and Wichman 2022; Parthum and Christensen 2022). Within the recreation literature, Dundas and von Haefen (2020) estimate the welfare impacts of climate change on recreational marine fishing with a similar approach to my paper. However, in their model, temperature only affects the participation choice, while I allow temperature to impact both the participation choice and the site choice. This is critical in my setting, where temperatures vary substantially across parks. Like me, Dundas and von Haefen find larger responses to cold than to extreme heat, but I document a narrower range of ideal temperatures and separate responses to long-run differences versus short-run shocks.

Section 2 describes the nationally representative telephone survey, monthly visitor counts, and park attributes data. Section 3 outlines how travel costs are calculated for both driving and flying travel modes. Section 4 presents the model of national park visitation. Section 5 details the two-step estimation procedure and its benefits. Section 6 describes the general results, and section 7 applies the framework to study the impact of temperature on demand. Section 8 concludes.

2 Park Visitation and Attributes Data

The main data sources for this project describe individual-level visitation, park-level visitation, and physical and institutional attributes of the national parks. This section describes each of these data sources.

2.1 Individual-Level Visitation Data

The individual-level visitation data come from the National Park Service's Comprehensive Survey of the American Public. The survey conducts telephone interviews with the primary goal of gauging sentiment towards the National Park System. The roughly fifteen-minute interview includes several questions regarding individuals' visitation history. These questions provide the location of the most recent national park visit and the number of visits over the previous two years. For random subset of the sample, I also observe whether respondents drove or flew on their most recent visit.

Several characteristics of the Comprehensive Survey of the American Public make it a uniquely useful data source for studying national park visitation. It includes both visitors and non-visitors, allowing me to model the extensive-margin the choice to visit or not visit. It is nationally representative. Phone numbers are selected using a regionally-stratified random sampling design, and individuals are randomly selected within each household. The data include weights to account for the regional stratification and match sample demographics statistics to the Census, so weighted mean demographics from the survey closely match mean demographics of the general population (table 1). I use these weights throughout my analysis. Additionally, the survey was conducted twice: once in 2008 and 2009 and again in 2018. The two iterations are similar, with identical visitation history questions. The 2008 and 2009 interviews were split between seasons to account for seasonal variation in visitation. The 2018 survey, citing a lack of seasonality in the 2008 and 2009 data, conducted interviews from June through November.

The Comprehensive Survey of the American Public also has important limitations. First, I observe respondents' home locations imprecisely. In the 2008 iteration, the data include the area code of each respondent's telephone number and their state of residence. In the 2018 survey, the data only include respondents' state of residence. When the area code is within the state of residence, I take the largest city in the area code as the home city for calculating travel costs. When I only observe the state of residence, I randomly sample a home city according to the state's population distribution. Second, the survey does not include any information on visit dates, only that the visits occurred within two years of the interview. This limits my ability to capture seasonal variation in certain parameters, including the travel cost coefficient. I discuss the implications for estimation in section 5. Finally, many

Table 1: Telephone Survey Descriptive Statistics

	Unweighted	Weighted	Census
Age			
18-29	11.8	21.3	23.6
30-39	13.5	16.3	16.9
40-49	16.7	16.7	18.4
50-59	24.1	20.8	17.5
60-69	18.5	14.3	12.0
70+	15.1	10.4	11.3
Income			
Less than $10,000	4.5	6.0	12.6
$10,000 to $25,000	9.5	11.0	15.0
$25,000 to $50,000	20.3	23.2	23.5
$50,000 to $75,000	20.8	22.2	18.9
$75,000 to $100,000	17.3	15.9	13.5
$100,000 to $150,000	15.4	13.1	10.7
Greater than $150,000	12.0	8.3	5.4
Education			
Some high school	3.5	5.5	
High school degree	36.8	46.9	
College degree	35.8	32.0	
Graduate degree	22.8	14.6	
Has child	29.7	35.3	38.4
White, non-Hispanic	75.0	67.9	63.7
Region			
Alaska	14.1	0.2	0.2
DC only	11.6	0.2	0.1
Intermountain	14.9	14.9	15.1
Midwest	14.6	22.9	22.6
Northeast	15.1	22.9	23.2
Pacific	14.8	16.8	17.3
Southeast	14.7	21.8	21.2
Visited in past 2 years	67.9	61.7	
Avg number of visits	9.2	4.7	
Flew (Subsample N = 1537)	13.5	12.6	
Sample size	6762	6762	

Note: The table shows the share of respondents in various demographic groups for the pooled 2008-2009 and 2018 Comprehensive Survey of the American Public survey data compared to statistics from 2010 Census data. Weights are included in the survey and match survey statistics to Census averages. Thus, the weighted variable means align closely with Census means. The unweighted sample tends to be older, richer, and more white, non-Hispanic than the general population.

national parks are never chosen as a respondent's most recent visit. To estimate a full set park effects, I must incorporate visitor count data.

2.2 Park-Level Visitor Counts

The National Park Service publishes monthly, park-level visitor counts in their Visitor Use Statistics database. The counts have a broad temporal and geographic scope, dating back to 1905 for some of the oldest parks and covering 383 national parks. I use counts from January 2005 through December 2019, because this period overlaps closely with the individual-level survey data and American Community Survey microdata.

Counting procedures vary by park and typically involve National Park Service rangers at entry booths and/or strategically placed vehicle counters. Parks use person-per-vehicle multipliers to convert vehicle counts to person counts. Busy peak seasons, available technology, and often remote locations make it difficult to obtain exact counts. Nonetheless, the Visitor Use Statistics are used administratively and in many academic studies of national park visitation (Fisichelli et al. 2015; Keiser, Lade, and Rudik 2018; Henrickson and Johnson 2013).

I adjust the raw visitor count data to make them suitable for recreation demand modeling. This process accounts for re-entry, group size, international visitation, and the primary purpose of trips using on-site surveys from the Visitor Services Project. I use 105 on-site surveys conducted at 69 different parks between 1995 and 2019. For parks that have not conducted an on-site survey, I impute missing information based on observable park attributes. With the park-specific trip information, I multiply raw visitor counts by the share of domestic visitors and the primary purpose of respondents' visits, then divide by the per-trip re-entry rate and average group size.

Table 2: Park Attribute Data Sources

Source	Variables
USGS National Map	Elevation range, mean elevation, trail miles, number of lakes > 40 acres, area of lakes > 40 acres
NPS Administrative Data	Designation (Park, Lakeshore, Seashore, etc), acreage, coastal, miles of shoreline, species presence
2004 NLCD	Share of land by landcover type, mode landcover type, landcover diversity (Herfindahl-Hirschman Index)
Census	Road miles, population density of overlapping counties
NCEI	Monthly avg max temperature, days with precipitation > 0.1", monthly 10-year avg temperature

Note: The table shows data sources for park attributes and variables generated from them. NPS Administrative Data include the *NPSpecies* database, Annual Acreage Reports, and a 2011 Resource Report on Shoreline length. NCEI data come from weather station-based Global Summary of the Month reports. NLCD - National Land Cover Database, NCEI - National Center for Environmental Information.

2.3 A Statistical Atlas of the National Parks

To understand visitor preferences for park attributes, I consolidate several datasets describing the national parks themselves. Table 2 shows the full list of data sources and the variables I generate from them.

3 How much does it cost to visit the national parks?

This section describes the procedure for computing travel costs. I calculate travel costs at a quarterly frequency for every individual in the nationally representative telephone surveys, as well as every individual in American Community Survey microdata between 2005 and 2019. I use these microdata to calibrate the model outside the survey period. Travel costs include the time and money required for individuals to drive or fly to each of the national parks. These calculations closely follow English et al. (2018), which also computes driving and flying travel costs at a national scale.

To compute driving travel costs, I calculate the driving mileage and time from each respondent's home location to each national park using PC*Miler. I multiply mileage by the marginal cost of driving, which I calculate with per-mile maintenance and tire costs from annual AAA reports and regional gas prices from the Energy Information Agency. For every twelve hours of driving, I add the average U.S. hotel rate. I also make a standard assumption that the cost of travel time is one-third of a respondent's hourly wage rate.

Flying travel costs include travel time, plus the cost of driving to the origin airport, airfare, and the cost of driving from the destination airport to the park, which may include rental car prices. Quarterly average airfare data come from the U.S. Department of Transportation's Consumer Airfare report, which includes the average airfare for flights between city markets, rather than individual airports. I use the 2012 average rental car price from English et al. (2018), adjusting for inflation to approximate rental car prices in other years. For each individual-park pair, I compute travel costs for all routes originating at one of the four city-markets closest to the respondent's home and ending at one of the four city-markets closestto the park. I then select the cheapest of these routes as the individual's travel cost of flying to the park.

Figure 1 shows the flying and driving travel costs for a subset of the telephone survey sample. Driving travel costs increase approximately linearly with one-way driving distance, with different slopes for each income bin. On average, flying is more expensive than driving for trips under 1,600 miles but is cheaper at longer distances, matching calculations from English et al.

4 A model of national park visitation

In this section, I outline a model describing the choices of which national park to visit and how to travel. By jointly modeling the choice of park and the choice of travel mode, I build on both the recreation demand literature, which typically focuses solely on location choice, and the transportation literature, which has a rich history modeling travel mode choices

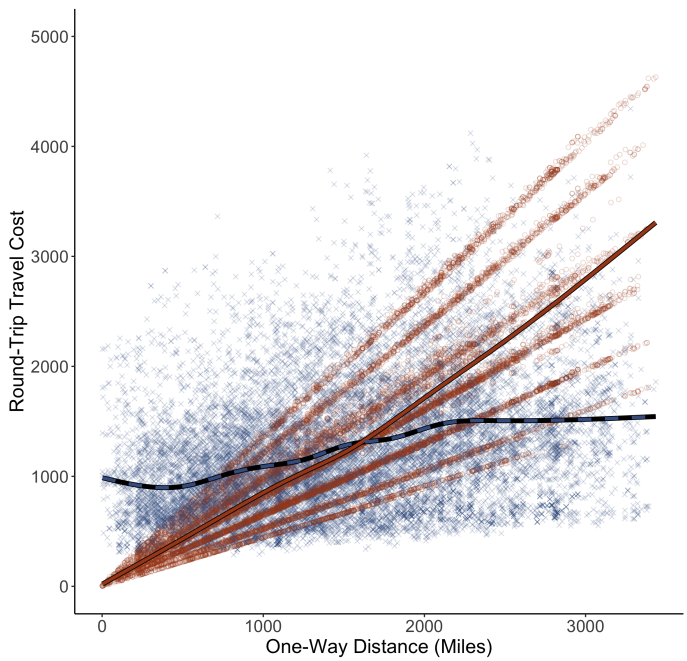

Figure 1: The graph plots calculated round trip travel costs on one-way driving distance for a three percent subset of the 2008 survey sample. Brown circles show driving travel costs, and blue x's show flying travel costs. Lines show average travel costs by distance for both driving (solid) and flying (dashed). On average, flying travel costs increase more gradually with distance.

(McFadden 1974).[4] The model also shares similarities with work by Chintagunta, Dubé, and Goh (2005), which allows for time-varying mean utilities when modeling demand for margarine.

Suppose that each month individuals choose whether to visit a national park, which national park to visit, and whether to drive or fly to the park. Denote the set of national parks $\mathcal{J} = \{1, 2, ...J\}$ and the set of travel modes $\mathcal{M} = \{D, F\}$, where D and F indicate driving and flying, respectively. Let $j = 0$ denote the outside option, each individuals' best way of spending the month that does not involve visiting a national park. Because visits to the National Park System's historic sites are included in the data but differ from visits to nature-centered national parks, I group visits to historic sites into a second outside option, $j = J+1$. Given this choice set, let U_{ijmt} denote the utility individual i receives from visiting national park j using travel mode m during month t, where

$$U_{ijmt} = \begin{cases} \delta_{0t} + \epsilon_{i0t} & j = 0 \\ \delta_{jt} + \beta_{TC} TC_{ijDt} + \epsilon_{ijDt} & j \in \{1, ..., J\}, m = D \\ \delta_{jt} + \beta_F + \beta_{TC} TC_{ijFt} + \epsilon_{ijFt} & j \in \{1, ..., J\}, m = F \\ \delta_{J+1,t} + \epsilon_{i,J+1,t} & j = J+1 \end{cases} \quad (1)$$

$$\equiv \begin{cases} V_{0t} + \epsilon_{i0t} & j = 0 \\ V_{ijmt} + \epsilon_{ijmt} & j \in \{1, ..., J\}, m \in \{D, F\} \\ V_{J+1,t} + \epsilon_{i,J+1,t} & j = J+1 \end{cases} \quad (2)$$

In equation (1), coefficient β_{TC} represents the marginal disutility of travel costs, and β_F represents the fixed cost of flying relative to driving. For $j \in \{1, ..., J\}$, I call the park-by-month fixed effect, δ_{jt}, the park effect. It captures the mean utility provided by park j in month t after controlling for travel costs. Ranking the park effects produces a national park

[4] An exception, Hausman, Leonard, and McFadden (1995) model the travel mode choice in a recreation demand context. They create a model to quantify the recreational use losses of the *Exxon Valdez* oil spill.

awesomeness index. I decompose the park effects further:

$$\delta_{jt} = X_{jt}\alpha + \nu_{jt}, \tag{3}$$

where X_{jt} contains observable park attributes; α is a coefficient vector, and ν_{jt} is unobservable.

Assume the error term, ϵ_{ijmt} follows a Generalized Extreme Value distribution with a two-level nested structure with the no-visit alternative in its own nest. This assumption allows error terms for visit alternatives to be correlated and relaxes the Independence of Irrelevant Alternatives assumption imposed by conditional logit models. The nested logit model still imposes the Independence of Irrelevant Alternatives assumption within the visit nest. Under this nesting structure, the probability of choosing each alternative has a closed form:

$$P_{ijmt} = \begin{cases} \dfrac{\exp(V_{0t})}{\exp(V_{0t}) + (\sum_{k=1}^{J+1}\sum_{n\in\mathcal{M}}\exp(\frac{V_{iknt}}{\lambda}))^{\lambda}}, & \text{if} \quad j = 0 \\ \dfrac{\exp(\frac{V_{ijmt}}{\lambda})}{\sum_{k=1}^{J+1}\sum_{n\in\mathcal{M}}\exp(\frac{V_{iknt}}{\lambda})}\dfrac{(\sum_{k=1}^{J+1}\sum_{n\in\mathcal{M}}\exp(\frac{V_{iknt}}{\lambda}))^{\lambda}}{\exp(V_{0t}) + (\sum_{k=1}^{J+1}\sum_{n\in\mathcal{M}}\exp(\frac{V_{iknt}}{\lambda}))^{\lambda}}, & \text{if} \quad j \in \{1, ..., J\} \\ \dfrac{\exp(\frac{V_{i,J+1,t}}{\lambda})}{\sum_{k=1}^{J+1}\sum_{n\in\mathcal{M}}\exp(\frac{V_{iknt}}{\lambda})}\dfrac{(\sum_{k=1}^{J+1}\sum_{n\in\mathcal{M}}\exp(\frac{V_{iknt}}{\lambda}))^{\lambda}}{\exp(V_{0t}) + (\sum_{k=1}^{J+1}\sum_{n\in\mathcal{M}}\exp(\frac{V_{iknt}}{\lambda}))^{\lambda}}, & \text{if} \quad j = J+1 \end{cases} \tag{4}$$

For visit alternatives, the choice probabilities include two terms. The first indicates the probability of visiting a specific park using a specific travel mode, conditional on choosing a visit alternative. The second term indicates the probability of choosing any visit alternative. If an individual chooses not to visit, then they do not select a specific park and travel mode. Thus, the no-visit choice probability has only one term. The literature often refers to the parameter, λ, as the dissimilarity coefficient. For consistency with random utility maximization, λ is bounded between zero and one. Higher values of λ indicate more dissimilar alternatives in the visit nest, and λ equal to one simplifies the probabilities to match the conditional logit model.

5 A Two-Step Approach to Estimate Demand

This section describes the procedure to estimate demand for national park visitation. My procedure relates most closely to Murdock (2006), which introduced the two-step approach for estimating recreation demand models. It is also similar to the maximum likelihood estimator proposed by Berry, Levinsohn, and Pakes (2004), which combines micro and macro level data to estimate demand for automobiles. However, the particular features of the survey and visitor count data motivate a distinct procedure.

5.1 Step 1a: Maximum Likelihood with a Contraction Mapping

To begin, I estimate the parameters in equation (1). These include the marginal disutility of travel costs, the premium for flying relative to driving, and observable heterogeneity in preferences for park attributes. Because the individual-level survey data do not include the date of respondents' visits, I estimate a constant park effect for each survey period in this step and drop the t subscript. I recover the monthly panel of park effects in Step 1b (described below). Given seasonal travel patterns, the constant parameter restriction likely introduces bias from mispecification.

I specify a three part likelihood function to fully incorporate the visitation information contained in the survey data. Using the choice probabilities from equation (4), the likelihood of observing individual i's visitation history is:

$$L_i(\beta, \delta) = \underbrace{(\Pi_{j=0}^{J} \Pi_{m \in \mathcal{M}} P_{ijm}^{y_{ijm}})}_{(1)} \underbrace{(1 - P_{i0})^{v_i}}_{(2)} \underbrace{(P_{i0})^{24-1-v_i}}_{(3)} \tag{5}$$

The first term represents the likelihood of individual i's most recent visit. For this visit, I observe the park chosen, as well as the travel mode for a subset of respondents. The second term represents the likelihood for all other visits in the two years prior to the interview, where v_i indicates the number of visits in the past two years in addition to the most recent visit. The third term represents the likelihood from all non-visits in the two years prior to the interview. If an individual never visits a national park in the two years prior to their interview, then the model interprets this as choosing the no-visit alternative for each of the 24 months prior to their interview.

When maximizing the log likelihood function, I constrain the visitation shares predicted by the model to match the visitation shares observed in the visitor count data. I impose the constraint by applying the contraction mapping introduced by Berry (1994):

$$\delta^{n+1} = \delta^n + ln(s) - ln(\hat{s}(\delta^n, \beta)) \tag{6}$$

As the optimization routine iterates over values of β, the contraction mapping solves for the unique vector of park fixed effects, δ, that matches observed and predicted visitation shares. Leveraging the contraction mapping has several practical benefits. First, it allows me to simultaneously incorporate visitation information from individual-level surveys and park-level visitor counts. Second, the contraction mapping solves for δ, so the optimization routine must search only over β, reducing the estimation time. Finally, many parks are never chosen in the survey data, so their park effects cannot be identified from the survey data alone.

By estimating a full set of park effects, I control for bias from unobserved park attributes when estimating the travel cost parameter. In industrial organization settings, firms set prices and likely charge a higher price for products with desirable unobservable attributes. The correlation between price and unobserved attributes biases naive estimates and has led to the widespread use of instrumental variables. In the recreation demand literature, travel costs are not set by firms directly, but unobserved attributes, such as remoteness, may be correlated with travel costs. By including a full set of park effects, I control for unobserved attributes when estimating the travel cost coefficient (Murdock 2006). This approach is not possible in industrial organization settings, because the price variable does not vary across individuals.

Geographic sorting remains an identification concern (G. R. Parsons 1991). Individuals who value national parks may choose their residential location to reduce their travel costs. If individuals with low travel costs value national parks more highly than those far away and would visit more often even conditional on travel costs, then the marginal disutility of travel cost will be overstated. Sorting would bias travel cost coefficient estimates downward and away from zero, subsequently biasing willingness to pay estimates towards zero. Few travel cost papers address potential bias from sorting, and because of the limited use of travel cost methods at a national scale, the magnitude of the potential bias is unclear.

5.2 Step 1b: Calibrating Park-Month Effects Over 15 Years

In this step, I recover a monthly panel of park effects by applying the contraction mapping month-by-month, from January 2005 through December 2019. Calibration outside the survey period raises several important concerns. The geographic distribution of population may change significantly over the fifteen-year sample period. To account for this, I calibrate the model using annual American Community Survey (ACS) microdata samples (Ruggles et al. 2021). The microdata contain key demographic variables, such as income, family structure, and age, and they capture annual changes in the geographic population distribution. The ACS microdata also represent the national population, like the telephone survey data.

The calibration also requires assumptions on stability of β across time. In this paper, I assume β is constant across the entire fifteen-year calibration period. While this is not necessary, the assumption is based on empirical results. In a preliminary robustness check, I allow β to vary in the 2008 and 2018 individual-level survey periods and obtain similar estimates both periods. Dundas and von Haefen (2020) also provide evidence supporting this assumption. They allow travel cost coefficients to vary annually and obtain fairly stable estimates from 2004 through 2009.

I continue the estimation procedure by calculating individual choice probabilities for each individual in the ACS microdata, P_{ijmt}. Summing choice probabilities across individuals and travel modes yields the expected number of visits to each park. Dividing the expected number of visits by the sum of individual weights yields predicted visitation shares for each park. Beginning with January 2005, I apply the contraction mapping to obtain the unique vector of park effects that matches the predicted and observed visitation shares, given $\hat{\beta}$. Iteratively applying the contraction mapping month-by-month produces a full panel of park effects through December 2019.

The key insight is that applying the contraction mapping to solve for park effects requires predicting visitation shares but not observing individual-level choices. Thus, calculating park effects outside the survey period only requires an estimate of $\hat{\beta}$, a microdata sample, and observed visitation shares.

5.3 Step 2: Estimating Preferences for Park Attributes

In step 2, I estimate equation (3), which explains park effects as a function of park attributes. Because step 1a produces a panel of park effects, causal inference techniques, such as difference-in-differences and event-studies, can be implemented to rigorously identify preferences for park attributes. Previous work has called for more rigorous identification in travel cost models (Lupi, Phaneuf, and von Haefen 2020). My approach offers a method for blending modern econometric techniques with structural demand modeling.

Applying panel data econometric techniques within the structural model has several benefits. In a reduced-form regression, attribute changes at one park may cause visitors to substitute a visit with another park, biasing estimates. The model of park choice controls for the quality of substitute parks when estimating park effects. Therefore, spillovers do not bias estimates. The structural model also provides a framework for calculating welfare impacts. In section 7, I leverage these benefits to estimate preferences for temperatures.

I recover preferences for a broad range of park attributes using a correlated random effects model with a Mundlak device. Some attributes, such as elevation and wildlife presence, do not vary meaningful across a fifteen-year period. Other attributes, such as temperature, vary dramatically across parks and across time. While including a flexible set of fixed effects (e.g. park or park-by-season) has the attractive property of controlling for unobserved attributes constant across time, the fixed effects would subsume preferences for time-invariant attributes. The correlated random effects framework solves this issue. For attributes that vary over time, it recovers identical estimates to a fixed effects approach, and it preserves cross-sectional variation to recover preferences for time-invariant attributes (Mundlak 1978; Wooldridge 2019). Specifically, I estimate:

$$\delta_{jt} = X_j\alpha_0 + X_{jt}\alpha_1 + \bar{X}_{js}\alpha_2 + \nu_{jt}, \tag{7}$$

where X_j includes attributes that do not change over time, X_{jt} includes attributes that vary across time, and $\bar{X}_{js}$ is the mean of of the time-varying attributes at park j in season s for one of four seasons: December to February, March to May, June to August, or September to November.

6 Results: Preferences for U.S. National Parks

Table 3 reports estimates of travel cost, travel mode, and heterogeneity coefficients for several specifications of equation (1). For all specifications and all income groups, the travel cost coefficient is negative and significantly different from zero. Individuals with lower incomes are more sensitive to travel costs, consistent with a diminishing marginal utility of income. Preferences for travel mode differ meaningfully by income group, but all prefer driving to flying, on average. The low income group is willing to pay \$37 extra per-trip to drive rather than fly, while the middle and high income groups are willing to pay \$453 and \$674, respectively. One explanation for the driving premium is flexibility, as driving allows groups

Table 3: Estimates 2008 and 2018 Survey Periods

	(1)	(2)	(3)
Travel cost ($100)	-0.251	-0.250	-0.251
	(0.016)	(0.015)	(0.016)
Fly	-0.945	-1.132	-1.131
	(0.087)	(0.109)	(0.109)
TC x income $<$ $25,000	-0.277	-0.328	-0.326
	(0.022)	(0.027)	(0.027)
TC x income $>$ $100,000	0.127	0.120	0.119
	(0.009)	(0.009)	(0.009)
Flying x income $<$ $25,000		0.914	0.917
		(0.178)	(0.177)
Flying x income $>$ $100,000		0.351	0.241
		(0.131)	(0.138)
Flying x parent			0.568
			(0.117)
Dissimilarity coefficient	0.657	0.662	0.654
	(0.042)	(0.041)	(0.042)

Note: The table shows estimates of the travel mode, travel cost, and heterogeneity coefficients in equation 1 with standard errors in parentheses. Income interaction coefficients are relative to the middle income group.

to adjust their schedule and add side trips. Higher income groups may be willing and able to pay more for this flexibility. Although, the driving premium also reflects factors I do not include in travel costs, such as airport parking fees, the risk of flight cancellations, or the fuel efficiency, reliability, and size of respondents' vehicles. The dissimilarity coefficient is between zero and one, implying the nested logit model is consistent with utility maximizing behavior (McFadden 1979). I use estimates from column 1 when calibrating the panel of park effects, but I plan to use results with more heterogeneity in the future.

Figure 2 shows estimated monthly park effects for two parks: Glacier NP and Great Smoky Mountains NP. The park effects should be interpreted relative to the no-visit option, which is normalized to provide zero mean utility.[5] The consistently negative park effects indicate that potential visitors, on average, prefer the no-visit alternative to visiting a specific

[5]To change the interpretation, regress park-month effects on month-of-sample or park fixed effects and take the residual as the new park-month effect.

park, even when that park has zero travel costs. In the model, individuals will only choose to visit a park if it has a large, positive error term draw. In interpreting this result, it is helpful to note that the no-visit alternative encompasses all ways to spend a month that do not involve visiting a national park, not just staying at home. Additionally, park effect estimates are sensitive to the specified number of choice occasions and the market size. Assuming fewer choice occasions or a smaller market size raises park visitation shares and increases park effects relative to the no-visit alternative.

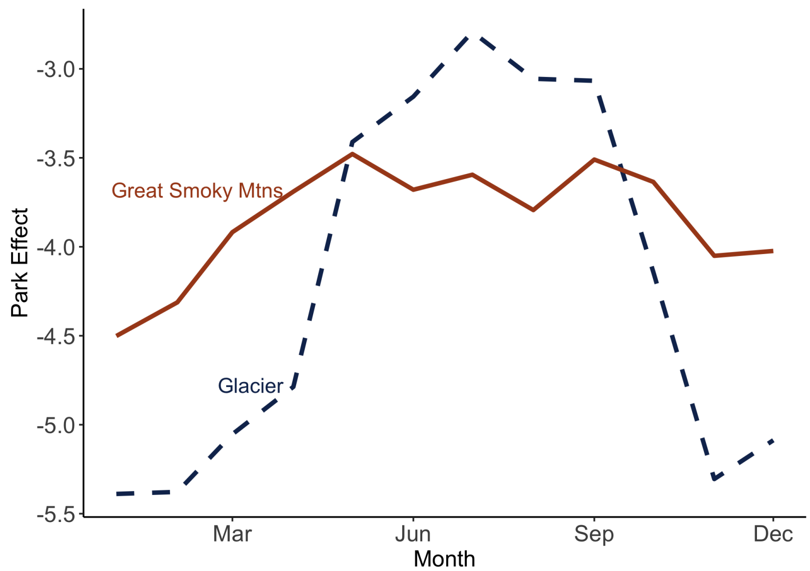

Figure 2: The graph plots estimated park effects for Great Smoky Mountains NP (solid-brown) and Glacier NP (dashed-blue) for each month in 2018. Both exhibit seasonal variation that has been largely overlooked in the recreation demand literature.

Glacier's park effects exhibit dramatic seasonal variation, peaking in the summer and collapsing in the winter. Converting the seasonal difference to dollar terms, potential visitors are willing to pay \$1032 more, on average, to visit in July rather than in January. Great Smoky Mountains displays a flatter peak period and less extreme winter decline. Similar patterns at other parks suggest that temperature and climate drive seasonal variation in park-month effects. I explore this further in section 7.

Results table 4 shows how park attributes impact park effects. Unshaded variables do not vary meaningfully over fifteen years, either due to data availability or geophysical processes.

Conditional on other observable attributes, average willingness to pay tends to be higher at parks with redwood forests, bison, bald eagles, coastline, more roads, more trails, and large elevation ranges, possibly a proxy for quality views. Population density in surrounding counties is positively correlated with the park effects. This reflects the amenities nearby a park, such as restaurants, hotels, and other attractions, but the estimate is likely biased upward, because desirable, unobserved park attributes attract visitors and generate local economic impacts. The land cover coefficient estimates suggest visitors appreciate barren land, which includes rock and sand, more than other land cover types, such as forest, wetland, and grassland. Willingness to pay is lower for parks with grizzly bears and those with more diverse land cover, measured using a standardized Herfindahl-Hirschman Index. These results should be interpreted with caution, because they are identified with cross-sectional variation and susceptible to omitted variables bias. Nonetheless, they provide the most extensive, revealed-preference evidence to date regarding what attracts visitors to U.S. national parks.

The shaded coefficient estimates are identified by within-season variation at each park and can also be obtained from a model including park-by-season fixed effects. More rainy days in a month, either historically or contemporaneously, decreases willingness to pay, while acreage changes have minimal impact. The national park designation coefficient reflects the impact of switching a park's designation to "national park" rather than one of the various other designations. Redesignating units as official national parks has been proposed as a possible solution to reduce crowding at other parks by increasing the profile of substitute parks. I provide evidence from a limited set of redesignations; the coefficient is identified from only three parks (Pinnacles, Gateway Arch, and Indiana Dunes) that received their official designation between 2005 and 2019. Based on this evidence, redesignation seems unlikely to cause substitution away from other parks. Szabó and Ujhelyi (2021) examine a broader set of redesignations and find that an official national park designation increases visitation. My results indicate that the impact of redesignation may vary depending on

Table 4: Preferences for Park Attributes

	Coefficient	WTP
Redwoods present	0.872*	347
	(0.508)	
Bison present	0.296	118
	(0.268)	
Bald eagles present	0.152	60
	(0.147)	
Coastal	0.087	35
	(0.262)	
Elevation range (1000 ft)	0.052*	21
	(0.031)	
Land cover share: barren land	0.020**	8
	(0.007)	
Trail miles (10 miles)	0.017**	7
	(0.005)	
Nearby population density (100 per sq mile)	0.008**	3
	(0.002)	
Road miles (10 miles)	0.002	1
	(0.004)	
Land cover share: shrub/scrub	0.003	1
	(0.003)	
Lake acreage (100 acres)	0.000	0
	(0.000)	
Acreage (10k)	0.001	0
	(0.002)	
Trail miles x elevation range	0.000	0
	(0.001)	
Rainy days	-0.005**	-2
	(0.001)	
Land cover share: grassland	-0.009*	-3
	(0.004)	
Average number of rainy days	-0.013**	-5
	(0.006)	
Land cover share: emergent wetland	-0.014	-6
	(0.010)	
Land cover share: mixed forest	-0.016**	-7
	(0.006)	
National Park designation	-0.027*	-11
	(0.014)	
Coastal x elevation range	-0.070	-28
	(0.093)	
Land cover diversity (standardized)	-0.311**	-124
	(0.110)	
Grizzly bears present	-0.806**	-321
	(0.388)	
R-squared:	0.557	

* - Significant at 90% Level, ** - Significant at 95% Level. Estimates for shaded variables are equivalent to estimates from a model including park-by-season fixed effects. Unshaded variables use only between-park variation. Flexible temperature controls are also included. Willingness to pay (WTP) is calculated by dividing each attribute coefficient by the travel cost coefficient for the middle income group and multiplying by 100.

the context and motivate research on recent redesignations to inform the on-going policy discussion.

Even with this broad array of park attributes and temperature controls, roughly 44% of the variation in park effects remains unexplained. Given the unique resources the parks protect, this is not surprising. It is difficult to estimate the value of iconic park attributes, such as Arches' arches or Yellowstone's Old Faithful geyser, which are often idiosyncratic and remain largely unchanged over time.

By capturing mean utilities after controlling for travel costs, monthly park effects provide a national park *awesomeness* index. Table 5 shows the implied ranking for 2018 based on parks' maximum park effect throughout the year. I convert the park effects to a 100-point scale, providing a rating for each park. The maximum park effect between 2005 and 2019 scores 100 and the minimum scores 0. This ranking method offers an attractive alternative to rankings from the popular media that are typically based on travel bloggers' personal experiences or raw visitation counts. Unlike experience-based rankings, it is systematic and incorporates the visitation history of the entire U.S. population. Unlike raw visitation rankings, it controls for the travel costs of reaching a park, isolating the appeal of the park itself.

The top ten ranking includes many of the most famous national parks, such as Glacier, Yellowstone, and Grand Canyon. One surprising result is that Golden Gate National Recreation Area tops the list. Golden Gate provides views of the famous Golden Gate Bridge, beaches, hiking trails, and popular attractions like Alcatraz Island, but for several reasons, its ranking is likely inflated. Although the model controls for the travel costs of accessing each park, it does not control for complementary destinations near a park. Visitors to Golden Gate likely visit other Bay Area attractions on the same trip, while Glacier NP, for example, has fewer convenient complementary attractions. Furthermore, local residents may visit Golden Gate several times per month, or even several times per week. The model's assumption of one choice occasion per month seems appropriate for most people and most

Table 5: Most Awesome National Parks

Rank	Park	Rating
1	Golden Gate RA	97.4
2	Glacier	93.9
3	Yellowstone	92.9
4	Grand Canyon	92.5
5	Grand Teton	91.9
6	Mount Rainier	91.4
7	Acadia	91.0
8	Rocky Mountain	90.9
9	Olympic	90.6
10	Zion	90.5

Note: The National Park awesomeness index combines visitation and travel cost data to rank parks by the mean utility they provide visitors. The ranking reflects each parks maximum park effect throughout 2018.

parks, but it is likely too coarse for local residents. This potentially biases Golden Gate's park effect upward.

7 How does Temperature Impact Demand?

In this section, I apply the model and data infrastructure to study preferences for temperature. This provides insights for understanding how climate change will impact visitation patterns. It also provides a blueprint for applying the the framework to study other challenges facing the National Park System.

To begin, I filter the visitation data through the demand system, as described in sections 5.1 and 5.2. Then, I specify a particular functional form of equation (3). Motivated by Bento et al. (2020), I decompose temperature impacts into long-run differences in average temperatures and short-run shocks. I bin average temperatures to flexibly capture nonlinearity, and I estimate the impacts of shocks separately for each average temperature bin. I write the park effect for park j in month t as

$$\delta_{jt} = \sum_b \alpha^b_{avg} \mathbb{1}(\overline{temp}_{jt} \in b) + \sum_b \alpha^b_{shock} (temp_{jt} - \overline{temp}_{jt}) \mathbb{1}(\overline{temp}_{jt} \in b) + \alpha X_{jt} + \gamma_{js} + \phi_t + \nu_{jt}, \tag{8}$$

where the variable $\overline{temp}_{jt}$ represents the average temperature at park j over the past ten years in the same calendar month as month t (e.g. the previous ten June's if t is June). Thus, $\overline{temp}_{jt}$ is, roughly, the anticipated or expected temperature, if a potential visitor were planning a trip more than a few weeks in advance. The variable $temp_{jt}$ represents the (mean daily high) temperature at park j in month t. It measures the realized temperature in month t. Taking the difference, $(temp_{jt} - \overline{temp}_{jt})$ represents the realized deviation from the moving average temperature in month t, or the short-run temperature shock. Including α^b_{avg} coefficients for each temperature bin allows for a non-linear relationship between average temperatures and monthly park effects. By interacting average temperature bins with the temperature shocks, the α^b_{shock} coefficients allow for the impact of a temperature shock to vary depending on the average temperature.

Equation (8) also includes precipitation controls, X_{jt}, and a flexible set of fixed effects to identify the causal impact of temperature. The γ_{js} is a park-by-season fixed effect whose inclusion allows me to isolate variation within a park-season. Here, I specify four seasons: December to February, March to May, June to August, and September to November. The ϕ_t further controls for system-wide factors affecting demand, such as the number of weekend days in a month. The final term, ν_{jt}, includes unobservable attributes of park j at time t.

Results figure 3 shows that visiting a park provides the most surplus when average high temperatures fall between 70°F and 85°F. Willingness to pay decreases sharply as temperatures become colder. Relative to 70°F, visiting when the temperature is 30°F reduces willingness to pay by $503 per-trip, on average. Moving from the ideal range to extremely hot temperatures also reduces willingness to pay, but not as dramatically. Visiting a park when the average high temperature is 95°F reduces willingness to pay by just $107.

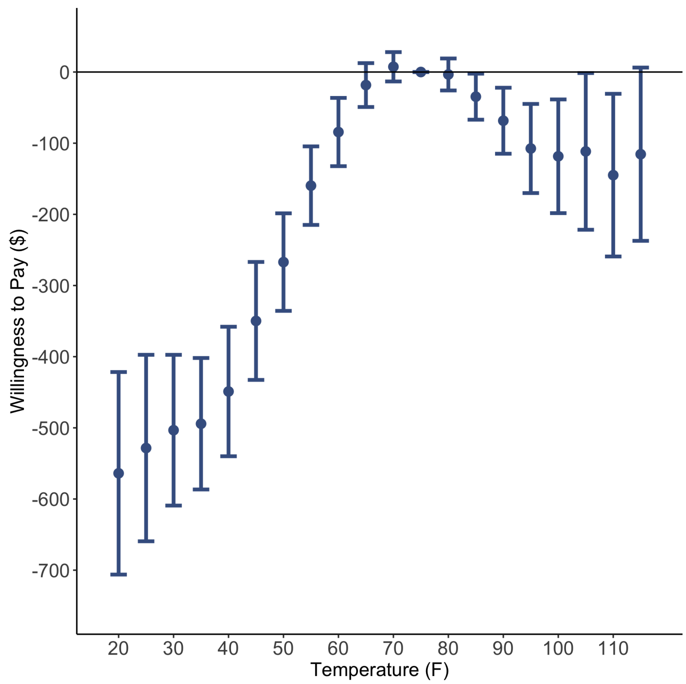

Figure 3: The figure shows potential visitors' willingness to pay for a park visit across long-run average temperatures. All estimates are relative to the 75°F bin.

Preferences for short-run temperature shocks vary intuitively across average temperatures (figure 4). When temperatures are cold, positive shocks increase willingness to pay, peaking at \$6 for a one degree temperature increase between 40°F and 55°F. At ideal temperatures, a small shock has little impact, and at hot temperatures, the impact of a shock is indistinguishable from zero.

Results figure 5 compares the impact of equal-sized average temperature differences and shocks. When temperature differences impact willingness to pay most, at 50°F, a 5°F increase in average temperatures raises willingness to pay by \$107 compared to a \$32 increase for a 5°F shock. At higher temperatures, increasing average temperatures from 90°F to 95°F reduces average willingness to pay by \$39 per-trip, while a short-run shock has little impact. The magnitude of willingness to pay for changes in average temperatures is greater at nearly every temperature. This result may be driven by visitors that plan trips far in advance and cannot observe short-run shocks when making their choice. After committing to their trip, it may be costly to cancel or substitute to another location, minimizing the observable

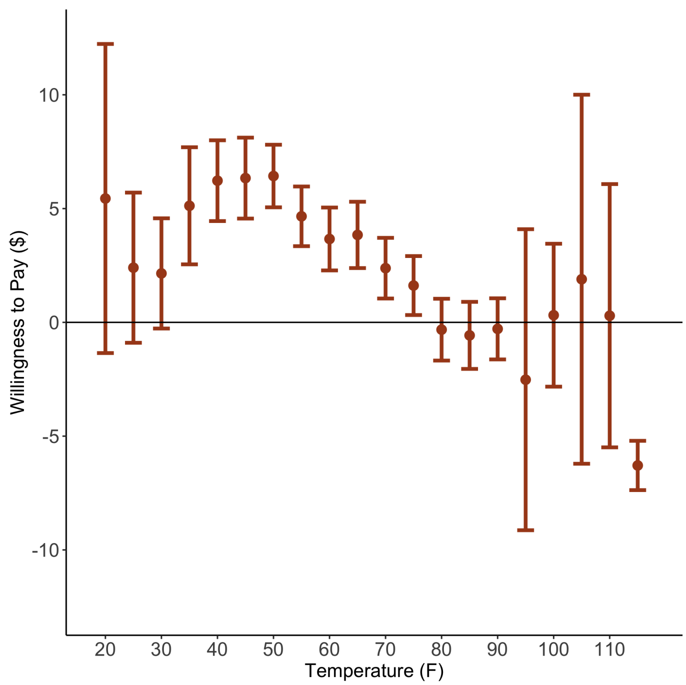

Figure 4: The figure shows potential visitors' willingness to pay for a one degree Farehnheit positive temperature shock.

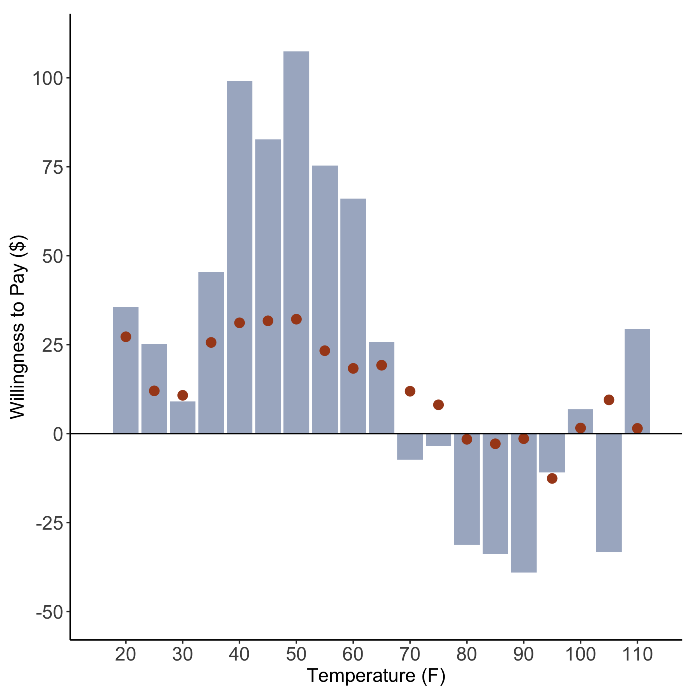

Figure 5: The graph shows the willingness to pay for a five degree increase in long-run average temperature (blue bars) and short-run temperature shocks (brown dots) by long-run average temperature. It provides a uniform change in temperature for comparing estimates displayed in figures 3 and 4.

response to short-run shocks. Visitors may also respond to shocks by changing behavior on their trip, such as spending less time in the park, but I cannot observe within-trip responses in the survey or visitor count data. Thus, my results may understate the impact of short-run temperature shocks.

For both long-run temperatures differences and short-run shocks, the benefits of making cooler temperatures warmer far outweigh the costs of making hot temperatures hotter. This suggests that climate change may increase the surplus generated by the national parks. These findings come with the important caveat that my estimates do not account for how climate change will impact park resources. An increased frequency of wildfires and flooding, sea level rise, and possible species extinctions are just a few ways that climate change will alter the national parks. Understanding these broader climate impacts is important, and my framework provides a strong foundation for future research on these topics.

8 Conclusion

This paper creates a versatile framework to study demand for U.S. national parks. The results describe general preferences for national parks and their attributes. On average, potential visitors are willing to pay $376 more to drive rather than fly to a park, conditional on travel costs. Willingness to pay tends to be higher at parks with iconic wildlife, wide-ranging elevation, and coastline, and it varies dramatically across seasons, especially for parks with harsh winters. The national park *awesomeness* index provides a systematic alternative to existing rankings that controls for travel costs. It ranks many of the most famous parks in the top ten. Observable park attributes explain 56% of variation in the index, meaning idiosyncratic, unobservable, or difficult to quantify attributes play an important role in driving visitation.

I demonstrate the versatility of the framework by estimating willingness to pay for long-run average temperature differences and for short-run temperature shocks. Potential visitors prefer average high temperatures between 70°F to 85°F and respond to long-run differences more than short-run shocks. Long-run average temperatures of 95°F are preferred to 30°F by roughly $400. Without considering changes to park attributes, this suggests warming temperatures may increase the welfare generated by national parks. As the next step for the project, I plan to use my temperature preference estimates to simulate welfare changes under future climate projections, quantifying the welfare impact of climate change on national park visitation.

The model, data infrastructure, and estimation procedure have the potential to generate future research on national parks and recreation demand more broadly. The estimation procedure provides a method to control for changing travel costs and demand system spillovers when conducting causal inference. It filters visitor count data through a structural model, allowing for welfare analysis and counterfactual simulations, and it bridges gaps between individual-survey efforts. The framework is relevant for policy and management decisions surrounding many challenges facing the National Park System, such as crowding, the impacts of wildfires, and infrastructure investments. This is particularly important given recent legislative actions, which provide new resources for the continued conservation of the country's most treasured resources.

9 References

Bento, Antonio, Noah S. Miller, Mehreen Mookerjee, and Edson R. Severnini. 2020. "A Unifying Approach to Measuring Climate Change Impacts and Adaptation." *NBER Working Paper 27247*. https://www.nber.org/system/files/working_papers/w27247/w27247.pdf.

Berry, Steven. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *The RAND Journal of Economics* 25: 242–62. https://www.jstor.org/stable/2555829.

Berry, Steven, James Levinsohn, and Ariel Pakes. 2004. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." *Journal of Political Economy* 112. https://doi.org/https://doi-org.proxy2.cl.msu.edu/10.1086/379939.

Burke, M., M. Craxton, C. D. Kolstad, C. Onda, H. Allcott, E. Baker, L. Barrage, et al. 2016. "Opportunities for Advances in Climate Change Economics" 352: 292–93. https://doi.org/10.1126/science.aad9634.

Burns, Ken. 2009. "The National Parks: America's Best Idea." Public Broadcasting Service.

Carleton, Tamma, Amir Jina, Michael Delgado, Michael Greenstone, Trevor Houser, Solomon Hsiang, Andrew Hultgren, et al. 2022. "Valuing the Global Mortality Consequences of Climate Change Accounting for Adaptation Costs and Benefits." *The Quarterly Journal of Economics* 137: 2037–2105. https://doi.org/10.1093/qje/qjac020.

Chan, Nathan W., and Casey J. Wichman. 2022. "Valuing Nonmarket Impacts of Climate Change on Recreation: From Reduced Form to Welfare." *Environmental & Resource Economics* 81: 179–213. https://doi.org/10.1007/s10640-021-00624-3.

Chintagunta, Pradeep, Jean-Pierre Dubé, and Khim Yong Goh. 2005. "Beyond the Endogeneity Bias: The Effect of Unmeasured Brand Characteristics on Household-Level Brand Choice Models." *Management Science* 51: 832–49.

Cullinane Thomas, Cathy, Matthew Flyr, and Lynne Koontz. 2022. "2021 National Park Visitor Spending Effects: Economics Contributions to Local Communities, States, and the Nation." *National Park Service.* https://www.nps.gov/subjects/socialscience/vse.html.

Dundas, Steven, and Roger von Haefen. 2020. "The Effects of Weather on Recreational Fishing Demand and Adaptation: Implications for a Changing Climate." *Journal of the Association of Environmental and Resource Economists* 7: 209–42. https://doi.org/10.1086/706343.

Egan, Kevin, Joseph Herriges, Catherine Kling, and John Downing. 2009. "Valuing Water Quality as a Function of Water Quality Measures." *American Journal of Agricultural Economics* 91: 106–23. https://doi.org/10.1111/j.1467-8276.2008.01182.x.

English, Eric, Roger H. von Haefen, Joe Herriges, Christopher Leggett, Frank Lupi, Kenneth McConnell, Michael Welsh, Adam Domanski, and Norman Meade. 2018. "Estimating the value of lost recreation days from the Deepwater Horizon oil spill." *Journal of Environmental Economics and Management* 91: 26–45. https://doi.org/https://doi.org/10.1016/j.jeem.2018.06.010.

Fisichelli, Nicholas A., Gregor W. Schuurman, William B. Monahan, and Pamela S. Ziesler. 2015. *PLoS ONE* 10. doi:10.1371/journal.pone.0128226.

Hausman, Jerry A., Gregory K. Leonard, and Daniel McFadden. 1995. "A Utility-Consistent, Combined Discrete Choice and Count Data Model Assessing Recreational Use Losses Due to Natural Resource Damage." *Journal of Public Economics* 56: 1–30.

Henrickson, K. E., and E. H. Johnson. 2013. "The Demand for Spatially Complementary National Parks." *Land Economics* 89 (May): 330–45. http://www.jstor.org/stable/24243681.

Hsiang, Solomon, Robert Kopp, Amir Jina, James Rising, Michael Delgado, Shashank Mohan, D. J. Rasmussen, et al. 2017. "Estimating Economic Damage from Climate Change in the United States." *Science* 356: 1362–69. https://doi.org/10.1126/science.aal4369.

Keiser, David, Gabriel Lade, and Ivan Rudik. 2018. "Air Pollution and Visitation at U.S. National Parks." *Science Advances* 4. https://doi.org/10.1126/sciadv.aat1613.

Kolstoe, Sonja, and Trudy Ann Cameron. 2017. "The Non-Market Value of Birding Sites and the Marginal Value of Additional Species: Biodiversity in a Random Utility Model of Site Choice by eBird Members." *Ecological Economics* 137: 1–12. https://doi.org/10.1016/j.ecolecon.2017.02.013.

Lupi, Frank, Daniel Phaneuf, and Roger von Haefen. 2020. "Best Practices for Implementing Recreation Demand Models." *Review of Environmental Economics and Policy* 14: 302–23. https://doi.org/https://doi.org/10.1093/reep/reaa007.

McFadden, Daniel. 1974. "The Measurement of Urban Travel Demand." *Journal of Public Economics* 3: 303–28.

———. 1979. "Quantitative Methods for Analysing Travel Behaviour of Individuals: Some Recent Developments." In *Behavioural Travel Modelling*, edited by Croom Helm, 279–318. London.

Mundlak, Yair. 1978. "On the Pooling of Time Series and Cross Section Data." *Econometrica* 46: 69–85. https://doi.org/10.2307/1913646.

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

Neher, C., J. Duffield, and D. Patterson. 2013. "Valuation of National Park System Visitation: The Efficient Use of Count Data Models, Meta-Analysis, and Secondary Visitor Survey Data." *Environmental Management* 52 (May): 683–98. https://doi.org/https://doi.org/10.1007/s00267-013-0080-2.

Office of Aviation Analysis. 2015. "Consumer Airfare Report." *US Deparment of Transportation.* https://www.transportation.gov/policy/aviation-policy/domestic-airline-consumer-airfare-report.

Organic Act of 1916. n.d. 16 U.S.C §1 et seq. 1916.

Parsons, George R. 1991. "A Note on Choice of Residential Location in Travel Cost Demand Models." *Land Economics* 67: 360–64. https://doi.org/10.2307/3146431.

Parsons, George, Christopher Leggett, Joe Herriges, Kevin Boyle, Nancy Bockstael, and Zhe Chen. 2020. "A Site-Portfolio Model for Multiple-Destination Recreation Trips: Valuing Trips to National Parks in the Southwestern United States." *Journal of the Association of Environmental and Resource Economists* 8: 1–25. https://doi.org/https://doi.org/10.1086/710714.

Parthum, Bryan, and Peter Christensen. 2022. "A Market for Snow: Modeling Winter Recreation Patterns Under Current and Future Climate." *Journal of Environmental Economics and Management* 113 (May). https://doi.org/10.1016/j.jeem.2022.102637.

Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek. 2021. "IPUMS USA: Version 11.0 [Dataset]." Minnneapolis, MN: IPUMS. https://doi.org/10.18128/D010.V11.0.

Szabó, Andrea, and Gergely Ujhelyi. 2021. "Conservation and Development: Economic Impacts of the US National Park System." *Working Paper*, February.

US Department of the Interior. n.d. "National Park Service Visitor Use Statistics." https://irma.nps.gov/STATS/.

Wooldridge, Jeffrey M. 2019. "Correlated Random Effects Models with Unbalanced Panels." *Journal of Econometrics* 211: 137–50. https://doi.org/10.1016/j.jeconom.2018.12.010.