Hotelling Meets Wright: Spatial Sorting and Measurement Error in Recreation Demand Models*

Jacob T. Bradt †

March 5, 2024

Abstract

Conventional applications of recreation demand models likely suffer from two standard challenges with demand estimation, namely omitted variables bias and measurement error. Idiosyncratic prices in the form of individual-level travel costs can exacerbate these two challenges: the potential for non-random selection into travel costs through residential sorting and the difficulty of observing individual-level travel costs both work to bias traditional model estimates. I demonstrate the magnitude of this potential bias in conventional estimates of recreation demand models. I provide a relatively simple instrumental variables approach to address these two empirical challenges that substantially outperforms traditional estimates in numerical simulations. Replicating English et al. (2018), I find that accounting for potential selection into travel costs and measurement error through the instrumental variables approach changes estimates of the welfare costs of the 2010 Deepwater Horizon oil spill by over 20 percent.

Keywords: Revealed Preference Methods, Recreation Demand, Instrumental Variables Estimator, Discrete Choice Models

JEL Codes: C36, Q26, Q51, Q53, R41

*I thank Joe Aldy, Myrto Kalouptsidi, Cathy Kling, Robin Lee, Kostas Metaxoglou, Ariel Pakes, Frank Pinter, Devesh Raval, Jim Stock, and Richard Zeckhauser as well as seminar and conference participants at the 2021 Association of Environmental and Resource Economists (AERE) Summer Conference and Harvard for invaluable comments and discussion. I obtain data on the Deepwater Horizon oil spill and Gulf Coast beach visitation from the National Oceanographic and Atmospheric Administration's Natural Resource Damage Assessment (NRDA) public repository. I gratefully acknowledge funding from the Vicki Norberg Bohm Fellowship. I am responsible for all remaining errors.

†Harvard University, John F. Kennedy School of Government; jbradt@g.harvard.edu.

1 Introduction

Outdoor recreation is an industry of both intrinsic and instrumental value. Private spending on recreational activities represents 2.1% of United States GDP annually (Bureau of Economic Statistics, 2019) and participation in outdoor recreation activities has increased in recent decades, with visitation to outdoor sites administered by the National Park Service rising by 16% over the period from 2010-2019 (National Park Service, 2020).

In addition to being an economically-meaningful industry, outdoor recreation provides insight into individuals' interactions with the natural environment. Environmental conditions at outdoor recreation sites—air quality on mountaintops, water quality at fishing sites, flora at campsites—generate substantial non-market value to visitors, affecting the quality of visitors' experience. As a result, decisions of which outdoor recreation sites to visit reveal how individuals' value the natural amenities of those sites. Studying consumer demand for outdoor recreation therefore not only sheds light on a consequential industry, but also allows us to understand the value of important environmental qualities which do not trade in market settings and whose value therefore proves difficult to quantify. Understanding these values has important implications for the provision of public goods, the development of regulations, and the litigation of environmental damages.

This argument motivates a large literature in the estimation of non-market, environmental amenities. The logic underlying this approach stems from Hotelling (1947)'s simple insight that consumption of an outdoor recreation site's amenities requires the agent to incur the cost of a trip to that site. The cost of traveling to a given site therefore serves as an implicit price for visits to that site, and site visits serve as quantities demanded. With data on travel costs and visitation patterns for different recreation sites, it is therefore possible to estimate models of demand for site visits. Moreover, it is possible to define recreation sites as bundles of attributes—including, for example, the types of outdoor activities supported, environmental qualities, and physical amenities, among others—so with data on these attributes across sites or over time, researchers can estimate how these amenities affect recreation demand. Estimating recreation demand as a function of travel cost and site attributes, including environmental amenities, allows for the valuation of changes in environmental quality.

In this paper, I examine the potential for two common challenges in the empirical estimation of demand for goods and services to affect conventional applications of recreation demand estimation, namely omitted variables bias and measurement error. Since the seminal work of Wright (1928), economists have recognized the fundamental identification challenge associated with estimating one or more coefficients of a system of simultaneous equations. For example, the empirical relationship between observed prices and quantities for a good reflects a set of equilibrium points on both the supply and demand curves for that good, making it impossible to estimate either the demand or the supply curves from these data alone. More broadly, there are many empirical contexts in which economists have good reason to believe that some unknown, unobserved factor simultaneously affects both the outcome of interest and an observed, explanatory variable. Failing to account for this empirical reality can substantially bias results.

In spite of this, researchers studying recreation demand often implicitly assume that price—defined as individual-specific travel cost—is exogenous in the visitation decision. However, there is ample reason to believe that correlation between travel cost, which is a direct function of individuals' choice of residence location, and unobserved characteristics may exist. For example, if anglers choose their permanent residence based on attributes of neighboring fishing sites or climbers choose to reside near high-value rock climbing locations, then estimates of individuals' responsiveness to price—the travel cost parameter—will be biased as these individuals are effectively choosing a lower price for a higher value good. This has important implications for estimates of the value of non-market, environmental amenities using standard recreation demand models.

There is ample empirical evidence to suggest that individuals do indeed sort—i.e., select their residence location—based on natural amenities and proximity to outdoor recreation opportunities. Work in the regional science and demography literature documents a decades-long trend of in-migration to areas with high environmental qualities in the rural US (Hjerpe et al., 2020; Rickman and Rickman, 2011). Recent empirical work in the spatial and urban economics literature also documents Tiebout (1956)-like sorting on preferences for spatial characteristics (Bayer and Timmins, 2007; Klaiber and Phaneuf, 2010), including climate and environmental qualities (Albouy et al., 2016; Bayer et al., 2009). Moreover, preferences for environmental qualities entering the choice of residence location is an implicit assumption of one of the other main techniques used to value non-market environmental amenities, the hedonic property framework of Rosen (1974). Indeed, much of this literature finds non-trivial capitalization of environmental amenities in housing prices, suggesting that individuals' do consider these factors in choosing where to live (Bishop et al., 2020).

Researchers have also long acknowledged the challenges associated with mismeasured variables in statistical and econometric analysis, which can often lead to attenuation in estimates of relationships of interest (Hausman, 2001). In the recreation demand context, this issue is perhaps particularly acute: while the price of recreation at a particular site is the marginal cost of travel to that site, researchers and analysts rarely—if ever–have access to the true cost of travel associated with realized trips, let alone the full set of travel costs for unselected trips in individuals' choice set. As a result, analysts must use information on residence location and the location of sites in individuals' choice sets in combination with a set of simplifying assumptions to construct estimates of travel costs. Though there is a well-documented set of best practices for doing so—see Lupi et al. (2020)—this approach easily leads to measurement error in the price of recreation activities, which is known to produce biased results (Angrist and Krueger, 2001; Hausman, 2001).

Despite good reason to believe that both omitted variables-induced endogeneity and measurement error in prices are non-trivial in this context, few if any applications of recreation demand modeling take these concerns seriously. Rather, these applications make the strong assumptions that all factors influencing demand for recreation consumption are observed and well-measured. To demonstrate the impact of these assumptions on the inferences we draw from these models, I simulate several site choice datasets that have either non-random selection into travel cost, mismeasured travel costs, or both and find that the standard models that ignore these challenges produce innacurate results.

To address both of these challenges in recreation demand estimation, I adapt a standard econometric approach that accounts for endogeneity in discrete choice models to the recreation demand context. Specifically, I outline how a two-stage control function approach to recreation demand estimation can mitigate concerns of bias introduced by travel cost endogeneity. This approach, first introduced by Heckman (1978), is widely applied in other contexts, including the management literature (Petrin and Train, 2010; Villas-Boas and Winer, 1999). The approach is analogous to the two-stage least squares estimator in linear models, which is known to sufficiently account for bias introduced by omitted variables and measurement error (Angrist and Krueger, 2001). In the first stage, travel cost is regressed in a linear model on a set of instruments which plausibly satisfy instrument relevance and an exclusion restriction and the residuals from this regression are included in estimation of the non-linear discrete choice model of site choice in the second stage.

I demonstrate the effectiveness of this two-stage control function approach using the simulated site choice datasets that have either non-random selection into travel cost, mismeasured travel costs, or both. I find that this relatively straightforward correction substantially outperforms standard approaches to recreation demand estimation.

To demonstrate the empirical relevance of these two challenges and the correction that I propose, I replicate English et al. (2018)'s nationwide model of demand for shoreline Gulf Coast recreation, which the authors use to estimate welfare losses associated with the 2010 Deepwater Horizon oil spill. This particular application is of first-order importance: the recreation demand modeling described in English et al. (2018) played a substantial role in compensatory litigation in the aftermath of the largest oil spill in US history. I estimate two versions of English et al. (2018)'s model, one which ignores potential travel cost endogeneity and measurement error and one which accounts for these issues via the two-stage control function approach. I find that accounting for these common empirical problems alters welfare estimates by as much as 22% in this context, a substantial difference in a policy-relevant setting.

This paper relates to several broad literatures. The first is the expansive literature using recreation demand models to value non-market environmental amenities. Early empirical applications of the travel cost logic—first introduced by Hotelling (1947)—estimate single site demand, mostly using zonal aggregate data on individuals' travel costs (Ward and Loomis, 1986). Later work notes the importance of accounting for substitution across recreation sites, relying on McFadden (1974)'s random utility maximization (RUM) framework to model agents' choice among a discrete set of potential sites (Phaneuf and Smith, 2005). RUM models are the dominant approach for describing consumer preferences for recreation and are used in a number of different contexts, including the valuation of water quality changes (Abidoye and Herriges, 2012; Abidoye et al., 2012; Egan et al., 2009; Smith et al., 1986), fish abundance (Kling and Thomson, 1996; Parsons et al., 2000; Shaw and Ozog, 1999), beach width (Parsons et al., 1999), and a host of physical site amenities (Hicks and Strand, 2000).

Several studies recognize the potential for omitted variables to bias estimates of recreation demand models. Parsons (1991) argues that recreation demand estimation suffers from a price endogeneity issue due to the potential for sorting on preferences for recreation, but does not provide an approach to address this issue in structural discrete choice models of recreation demand. Murdock (2006) and Abidoye et al. (2012) develop estimation procedures that account for unobserved site-specific attributes in the RUM model using alternative-specific constants while also allowing for inference on time-invariant, observed site characteristics. While this approach is feasible in many settings, the control function approach for which I advocate is computationally and conceptually straightforward and easily allows for inference on site-specific attributes. Moreover, the approach of relying on alternative-specific constants does not mitigate the particular source of endogeneity in question since the form of travel cost endogeneity arises over individual decision makers rather than sites.

Noting that congestion—a key site attribute, which is often omitted in recreation demand modelling—is endogenously determined by individuals' site visitation decision, Timmins and Murdock (2007) use an instrumental variables approach to account for this endogeneity in a revealed preference context. von Haefen and Phaneuf (2008) demonstrate that, when available, stated preference data can be combined with revealed preference data to identify site quality effects on behavior in the presence of unobservable site and user characteristics.

This paper also relates to a similarly expansive literature examining the capitalization of non-market environmental amenities in housing prices. Most of the results examining the price effects of environmental qualities use the hedonic property framework of Rosen (1974) to estimate capitalization in home prices. The hedonic framework has been applied to value proximity to hazardous waste sites (Greenstone and Gallagher, 2008), changes in air quality (Bajari et al., 2012; Bento et al., 2014), proximity to shale gas wells (Muehlenbachs et al., 2015), flood risk (Hallstrom and Smith, 2005), and water quality (Keiser and Shapiro, 2019). Generally, these studies find evidence in favor of capitalization of environmental amenities: residential transactions appear to account for a home's exposure to environmental amenities, both positive (i.e., amenities) and negative (i.e., disamenities). These general findings suggest that environmental qualities—the recreation site attributes of interest in most applications of recreation demand modelling—do indeed play a role in determining individuals' permanent residence location.

Several papers do consider both recreation site choices and residence locations when valuing non-market environmental amenities. Phaneuf et al. (2008) point out that conventional hedonic property studies estimating willingness-to-pay for non-market environmental amenities may not fully capture the set values that homeowners derive from non-market environmental amenities. In particular, Phaneuf et al. (2008) argue that recreational use values are not fully incorporated in valuations of environmental amenities derived from conventional hedonic analyses. The authors present a theoretical model which motivates a two-stage revealed preference model in which a recreation demand model is first estimated as a function of the environmental quality of interest, and the resulting estimates of marginal welfare gains from changes in environmental amenities are then incorporated into a standard hedonic property model. Phaneuf et al. (2008) apply this conceptual model to study ecosystem services delivered by a watershed, finding that accounting for recreational use values in a hedonic property model meaningfully increases estimates of welfare derived from the presence of a watershed. Kuwayama et al. (2022) apply the approach of Phaneuf et al. (2008) to estimates willingness-to-pay for water quality improvements in Tampa Bay, FL. I build on this literature by demonstrating the importance of accounting for not only recreation demand in models of residence location choice, but also residence location choices in models of recreation demand.

The remaining sections of the paper are organized as follows. Section 2 presents a standard discrete choice model of recreation demand that follows a class of models commonly found in the literature. Section 3 discusses each of the two challenges with the conventional recreation demand models—non-random and mismeasured travel costs—and provides simulation evidence of the potential bias resulting from these challenges. Section 4 presents the two-stage control function solution and documents the substantial performance gain of this estimation approach in simulated data. Section 5 replicates the recreation demand model application of English et al. (2018) with the control function correction and Section 6 concludes.

2 A Standard Model of Recreation Demand

This section presents a relatively standard discrete choice model of demand for recreation sites. While the literature employs a number of different parametric assumptions for estimating discrete choice models of recreation demand, I focus on a particularly common set of assumptions: the random parameters multinomial logit model. It is important to note that the challenges discussed in Section 3 generalize to many of the other common discrete choice models in the literature.

The basic random utility maximization hypothesis assumes that individuals select the alternative yielding the highest level of utility when facing a well-defined choice set (McFadden, 1974). Let u_{ijt} denote the conditional utility received by individual $i \in \{1, \ldots, N\}$ when selecting alternative $j \in \{1, \ldots, J\}$ on choice occasion $t \in \{1, \ldots, T\}$. The individual selects alternative j if and only if $u_{ijt} > u_{ikt} \forall k \neq j$. Let $y_{ijt} = 1$ if individual chooses alternative j and $y_{ijt} = 0$ otherwise, i.e.

$$y_{ijt} = \begin{cases} 1 & u_{ijt} > u_{ikt} \forall k \neq j \\ 0 & \text{otherwise} \end{cases}$$

Since it is not possible to observe all factors influencing individual site selection decisions, conditional utility is parameterized as a function of observable individual- and alternative-specific attributes, X_{ijt}, and some residual term, ε_{ijt}, which is known to the individual when making their decision, but unobserved by the econometrician. In particular, individual i's conditional utility from visiting recreation site j on choice occasion t is as follows:

$$u_{ijt} = \underbrace{X'_{ijt}\beta_i - c_{ijt}\alpha_i + \xi_j}_{\equiv v_{ijt}(X_{ijt};\theta_i)} + \varepsilon_{ijt} \tag{1}$$

where $X'_{ijt} = \begin{bmatrix} x'_{jt} & c_{ijt} \end{bmatrix}$ is a vector of observable site- and choice occasion-specific attributes, x_{jt}, and an idiosyncratic measure of travel cost, c_{ijt}; ξ_j is an alternative-specific constant that captures average valuations of invariant, site-specific attributes; and ε_{ijt} is an idiosyncratic, unobserved shock to preferences. I include the alternative-specific constant, ξ_j, in line with best practices in the literature to account for unobservable site attributes (Lupi et al., 2020).[1]

[1] It is possible to include the alternative-specific constant, ξ_j, and still obtain estimates of consumers' valuation of invariant, site-specific observables by projecting the observable factors of interest on estimates

The standard approach to fully specify the model given by (1) is to make an assumption on the distribution of the idiosyncratic shocks to preferences, i.e., the residual term ε_{ijt}. While there are several different distributional assumptions made in the recreation demand literature, the most common of these is that ε_{ij} is distributed Type 1 Extreme Value (T1EV) across the population, and is iid across individuals and sites, which corresponds to the logit model. This has a number of desirable properties, including the fact that each individual's resulting choice probability for the different alternatives has a simple, closed-form solution.

Note that (1) allows the coefficients on the additive, observed components of utility, $\theta_i' = \begin{bmatrix} \beta_i' & \alpha_i \end{bmatrix}$, to vary across individuals i in the population. To make this assumption tractable, it is typically assumed that this individual-level heterogeneity follows some parameterized distribution, $f(\theta)$. This specification allows for heterogeneity in preferences for the different observable attributes and, when combined with the assumption that the error term is distributed T1EV, is referred to in the literature as the mixed or random parameters logit. Allowing for individual-level preference heterogeneity has several benefits over a standard logit model without individual-level parameters. Perhaps most importantly, it results in unrestricted substitution patterns. In the standard logit framework, two alternatives with equivalent choice probabilities will have the same substitution patterns. This property is undesirable in many contexts, including recreation demand estimation: simply because two sites have similar probabilities of being visited in the data does not mean that individuals are equally likely to substitute towards them as a result of a change in travel cost or environmental quality.

A common assumption in the literature estimating mixed logit models is that the individual-level parameters are normally distributed, $\theta_i \stackrel{iid}{\sim} \mathcal{N}(\mu, \Sigma)$. The econometrician therefore estimates mean and standard deviation terms for each normally-distributed coefficient with individual-level heterogeneity, thereby providing information on the distribution of preferences for different observed attributes in the population. While the choice of parametric distribution is non-trivial, this approach greatly reduces the dimensionality of the model and makes estimation tractable.

Taking the common assumption of an extreme-value error term, $\varepsilon_{ijmt} \stackrel{iid}{\sim}$ T1EV, it is possible to specify the closed-form choice probabilities in this model. In particular, the probability that individual i chooses site j on choice occasion t is:

$$p_{ijt} = Pr(j \in \arg\max_{k \in \mathcal{C}} u_{ikt}) = \int \frac{\exp(v_{ijt}(X_{ijt}; \theta))}{\sum_{k \in \mathcal{C}} \exp(v_{ikt}(X_{ijt}; \theta))} f(\theta) d\theta \tag{2}$$

where $\mathcal{C} = \{1, ..., J\}$ is the choice set. This random parameters logit probability is a weighted

of these site fixed effects (Murdock, 2006).

average of the logit formula evaluated at different values of the parameters, θ, with the weights given by the density $f(\theta)$. Estimation proceeds via simulated maximum likelihood, where the choice probability (2) is approximated by the average across a large number of draws from, for example, $\theta_i \overset{iid}{\sim} \mathcal{N}(\mu, \Sigma)$ and the simulated log likelihood is defined as:

$$\mathcal{L}(\theta) = \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{j=1}^{J} y_{ijt} \log(\hat{p}_{ijt}(\theta)) \tag{3}$$

where likelihood contributions are summed across individuals $i \in \{1, \ldots, N\}$, choice occasions $t \in \{1, \ldots, T\}$, and alternatives $j \in \{1, \ldots, J\}$. The maximum simulated likelihood estimate is the value of θ that maximizes (3).

Having recovered estimates of the model parameters, θ_i, it is possible to construct estimates of willingness-to-pay for observable attributes, other measures of marginal rates of substitution, or changes in welfare associated with different attribute levels.[2] Often in recreation demand model applications, we are interested in constructing a measure of consumers' marginal willingness-to-pay for a given observable attribute, such as a measure of environmental quality, q_{jt}:

$$WTP_i^q = \frac{\beta_i^q}{\alpha_i}$$

Another statistic that is often of interest in recreation demand applications is the change in consumer surplus resulting from a change in a given environmental quality, say from q_{jt}^0 to q_{jt}^1 $\forall j, t$, which based on the parametric assumptions above is given by:

$$\Delta CS_i = \frac{1}{\alpha_i} \left[\sum_j \exp\left(v_{ijt}(X_{ijt}^1; \theta_i) \right) - \sum_j \exp\left(v_{ijt}(X_{ijt}^0; \theta_i) \right) \right]$$

It is possible to construct empirical estimates of these statistics from observable data and parameter estimates from (3).

3 Challenges with Travel Cost in The Standard Model

In this section, I present two major challenges associated with the implementation of the standard model outlined in Section 2. In particular, I discuss the challenges of non-random

[2]It is important to note that a feature common to all models of discrete choice is that the scale of utility is irrelevant: the alternative with the highest utility is the same regardless of the overall scale of utility (Train, 2009). This means that model parameters are only identified from observed choices and parametric assumptions up to an arbitrary shift in the scale of utility. Since the scale of utility does not affect the ratio of any two model parameter estimates, measures of willingness-to-pay and other ratios of parameters are identified from observed choices and parametric assumptions.

sorting on preferences for outdoor recreation—leading to selection into travel costs—and measurement error in travel costs and the resulting impacts on applications of standard models of recreation demand.

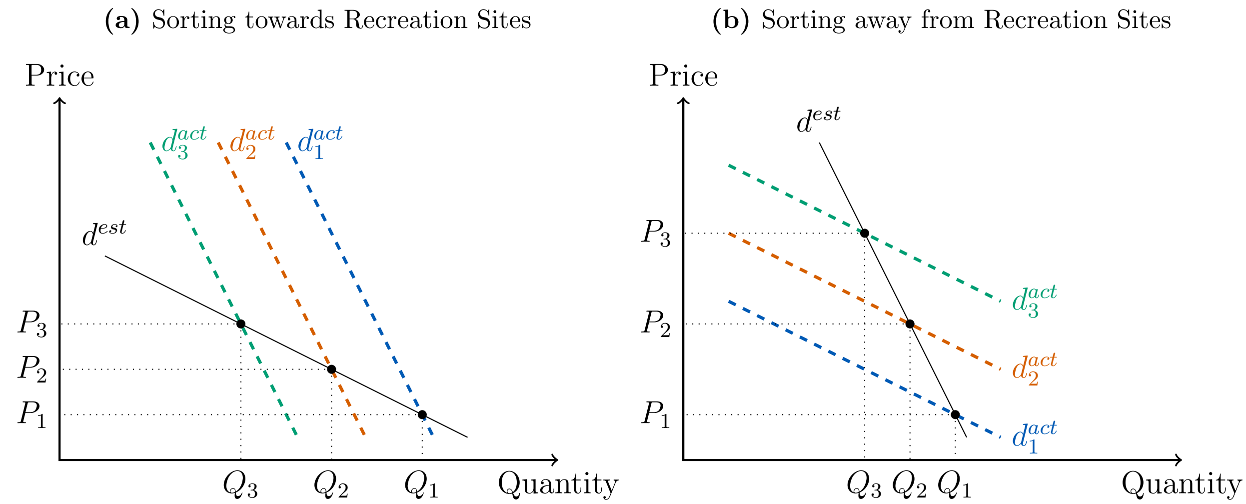
3.1 Selection into Travel Costs

I begin with a discussion of non-random residential sorting leading to selection into travel costs. If households factor preferences for certain recreation or environmental amenities into residence location decisions, this can result in non-random, non-zero correlation between observed travel costs and preferences for a specific site. For example, it is possible that households with a particularly strong idiosyncratic preference for a specific site and its attributes may choose to live closer to that site, all else equal. As Parsons (1991) suggests, this type of sorting towards desirable recreation sites could produce a negative correlation between observed travel cost and the likelihood of visiting a site, the quantity of visits to a site, or some other measure of demand for a site.

It is also possible—though admittedly, less likely—that households with a particularly strong idiosyncratic preference for a specific site and its attributes may choose to live further from that site, all else equal. This form of non-random sorting would result in a positive correlation between observed travel cost and demand for a given site. Such a pattern might emerge in the data if unobserved, idiosyncratic preferences for a site are correlated with another unobserved factor that drives sorting away from desirable sites. For instance, a strong idiosyncratic preference for a remote, pristine recreation site may be correlated with household income, which itself may be associated with a higher propensity to reside in urban centers far from high value recreation sites.

Though the former—sorting towards high-value sites—may be more plausible in practice than the latter—sorting away from high-value sites—it is important to note that any non-random spatial distribution of residences and recreation sites can lead to biased inferences under the standard approach. I demonstrate the bias in the standard approach to demand estimation under each model of residence-recreation location choice in Figure 1. For each form of sorting the analyst observes realized variation in levels of recreation demand at different prices: $(P_1, Q_1), (P_2, Q_2)$, and (P_3, Q_3). Ignoring the potential endogenous relationship between travel cost and recreation demand that arises due to non-random sorting, the analyst then recovers estimates of household recreation demand, d^{est}. As I show in Figure 1a, when households move close to recreation sites for which they have high idiosyncratic preferences, the analyst overestimates households' responsiveness to recreation costs, mistaking outward shifts in household demand curves for movement along the recreation price gradient. Analogously, when households move far from recreation sites for which they

Figure 1. Biased Recreation Demand Estimates with Two Sorting Patterns



Notes: This figure shows two stylized models of the bias introduced by non-random sorting of households on preferences for outdoor recreation. The left panel describes a scenario in which households sort towards recreation sites: households with higher preferences for a particular site's amenities choose to reside closer to that site. The right panel describes a scenario in which households sort away from recreation sites: households with higher preferences for a particular site's amenities choose to reside further from that site. In each scenario, the econometrician observes $(P_1, Q_1), (P_2, Q_2)$, and (P_3, Q_3) and estimates the demand curve d^{est}. The demand curves d_1^{act}, d_2^{act}, and d_3^{act} describe the actual demand curves of households observed at each point in the data consistent with each model of residence choice.

have high idiosyncratic preferences, the analyst underestimates households' responsiveness to recreation costs, mistaking inward shifts in household demand curves for movement along the recreation price gradient. I demonstrate this form of sorting in Figure 1b.

To be more precise about the nature of the endogeneity problem posed by non-random selection into travel costs, I return to the baseline model given by (1). Re-writing the residual term as the sum of two components gives the following specification of household i's conditional utility from visiting recreation site j on choice occasion t:

$$u_{ijt} = v_{ijt}(X_{ijt}; \theta) + \underbrace{\xi_{ijt} + \tilde{\varepsilon}_{ijt}}_{\equiv \varepsilon_{ijt}} \tag{4}$$

where ξ_{ijt} is an unobserved, idiosyncratic preference that is correlated with travel cost and $\tilde{\varepsilon}_{ijt}$ is an unobserved, idiosyncratic, and independent shock to preferences. The endogeneity problem arises due to the fact that travel cost is given by

$$c_{ijt} = w(z_{ijt}; \gamma) + \underbrace{f(\xi_{ijt}) + \mu_{ijt}}_{\equiv \epsilon_{ijt}} \tag{5}$$

where z_{ijt} are some observed instruments that affect travel cost, but not recreation site choice; $w(\cdot)$ is a function with parameters γ that relates z_{ijt} and travel cost c_{ijt}; $f(\cdot)$ is some unknown function that relates the idiosyncratic recreation site preference, ξ_{ijt} to travel cost c_{ijt}; and μ_{ijt} is a mean-zero, idiosyncratic shock to travel costs. In the standard model outlined in Section 2, I assume that $\xi_{ijt}+\tilde{\varepsilon}_{ijt} = \varepsilon_{ijt} \sim$ T1EV, thereby assuming that $\varepsilon_{ijt} \perp c_{ijt}$ which implicitly ignores the fact that $\xi_{ijt} \not\perp c_{ijt}$.

I demonstrate the bias introduced by non-random selection into travel costs in the standard recreation demand model via a set of simulated data generating processes. In particular, I specify different data generating processes based on the model of endogenous travel costs given by (4) and (5), simulate a large number of choice data from each data generating process, and apply the standard recreation demand model of (1) that ignores the endogeneity of travel costs to estimate model parameters based on each simulated choice dataset. Having knowledge of the data generating processes allows me to directly compare the resulting distributions of model parameter estimates with the true values of the target parameters. For expositional clarity, I suppress individual-level heterogeneity in target model parameters in both the assumed data generating processes and the discrete choice model that I apply to the simulated data.

For three distinct simulated data generating processes, I assume that individual i's indirect utility from and travel cost for alternative j follows

$$\begin{aligned} u_{ij} &= 1.0x_{ij} - 2.0c_{ij} + 1.0x_j + \xi_{ij} + \tilde{\varepsilon}_{ij} \\ c_{ij} &= 5.0 + 1.0z_{ij} + \rho_{sim}\xi_{ij} + \mu_{ij} \end{aligned} \tag{6}$$

where

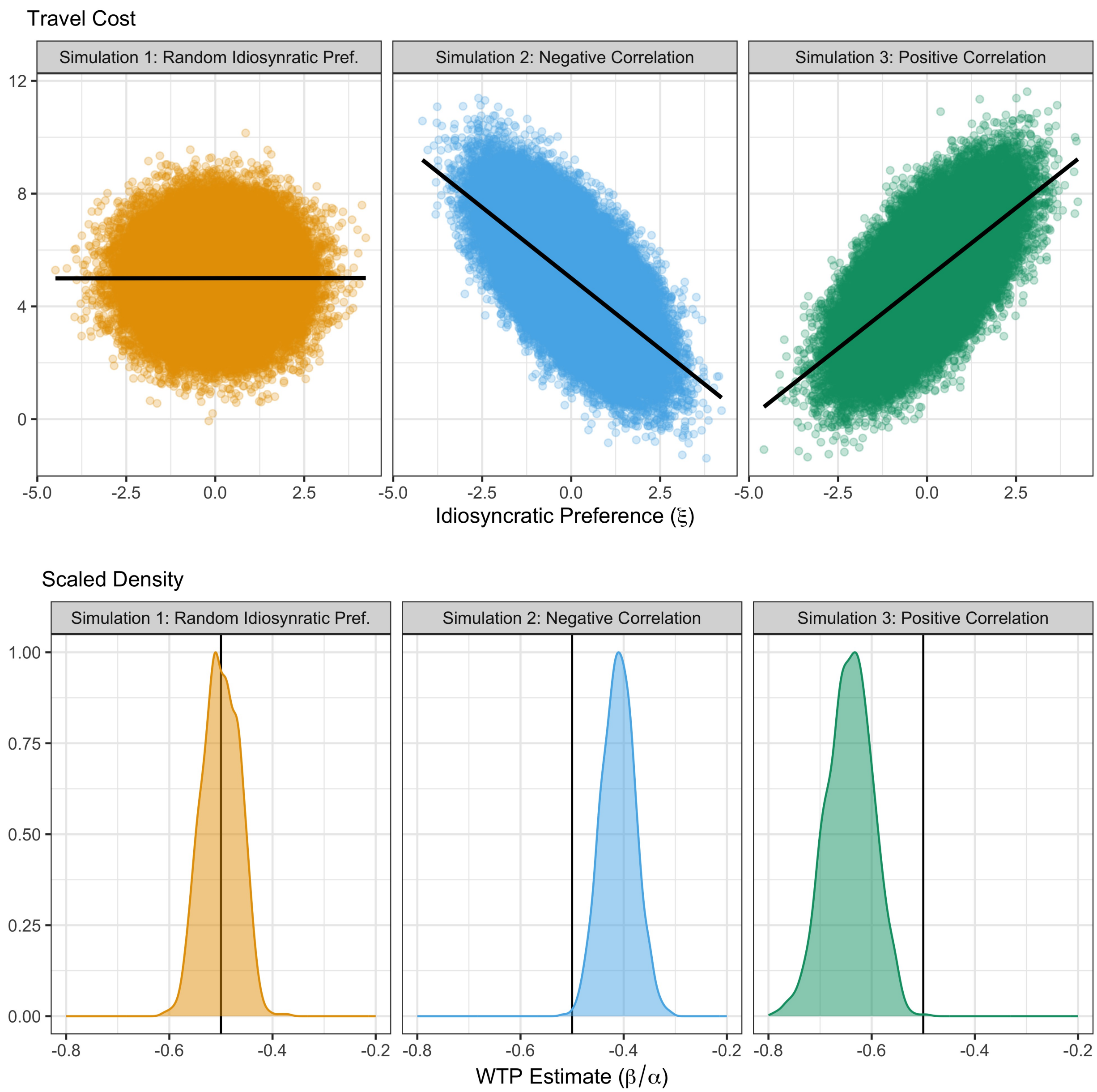
$$\begin{aligned} x_{ij} &\sim U(-1.0, 1.0) & x_j &\sim U(-1.0, 1.0) & z_{ij} &\sim U(-1.0, 1.0) \\ \xi_{ij} &\sim \mathcal{N}(0.0, 1.0) & \tilde{\varepsilon}_{ij} &\sim \text{Gumbel}(0.0, 1.0) & \mu_{ij} &\sim \mathcal{N}(0.0, 1.0) \end{aligned} \tag{7}$$

For the three simulations, I make the following assumption about ρ_{sim} to fully specify the data generating process of (6) and (7):

<u>Simulation 1:</u>	$\rho_1 = 0.0$	(No endogeneity)
<u>Simulation 2:</u>	$\rho_2 = -1.0$	(Sorting towards sites)
<u>Simulation 3:</u>	$\rho_3 = 1.0$	(Sorting away from sites)

I simulate 1000 unique choice datasets for each of the three above data generating processes. Each dataset consists of 1000 individuals—i.e., choice occasions—choosing between 100 alternative sites. The top panel of Figure 2 plots the empirical relationship between the idiosyncratic, unobserved preference (ξ_{ij}) and the idiosyncratic, observed cost variable (c_{ij}) for an example simulated choice dataset for each data generating process.

Figure 2. Bias from Non-random Sorting



Notes: This figure plots example relationships between households' unobserved, idiosyncratic preference (ξ_{ijt}) and travel cost from a single simulated dataset (top) as well as the distribution of estimated willingness-to-pay (WTP) across 1000 simulated datasets (bottom). The figure shows example relationships and WTP estimates for three assumed data generating processes: one where the idiosyncratic preference and travel cost are independent (left); one corresponding to a model of household sorting towards desirable recreation sites, where the idiosyncratic preference and travel cost are negatively correlated (center); and one corresponding to a model of household sorting away from desirable recreation sites, where the idiosyncratic preference and travel costs are positively correlated (right). The true value of the the willingness-to-pay statistic is shown as the vertical black line in the bottom panel. The full data generating process for each of simulations 1 to 3 are described in Section 3.1.

Having simulated 1000 choice datasets for each of the three simulations, I then make the standard assumption that $(\xi_{ij} + \tilde{\varepsilon}_{ij})$ ∼T1EV—i.e., ignore the data generating process for travel costs. This allows me to estimate the parameters of the linear indirect utility model in (6) for each simulated dataset via a multinomial logit model, thereby generating distributions of parameter estimates from the standard model of recreation demand for each data generating process. Note that since the scale of indirect utility is in general not identified in discrete choice models, I compare estimates of the marginal rate of substitution between x_{ij} and c_{ij}—a measure of willingness-to-pay for x_{ij}—when evaluating the relative performance of the standard assumption of travel cost exogeneity across the three distinct data generating processes.[3]

The bottom panel of Figure 2 plots the empirical distributions of willingness-to-pay estimates from the standard logit estimator across the three sets of simulations. Unsurprisingly, with no correlation between the mean-zero idiosyncratic preference (ξ_{ijt}) and travel cost (c_{ijt}) in Simulation 1, the logit model estimator performs well, with an average willingness-to-pay across all 1000 simulated samples equal to the true value of -0.5. Introducing non-zero correlations between the unobserved preference and travel cost results in poor coverage of the true target statistic: the average willingness-to-pay across all 1000 simulated samples is -0.41 in Simulation 2 and -0.64 in Simulation 3. Moreover, I can reject the null hypothesis that the willingness-to-pay estimates equal the true value of -0.5, with t-statistics of 2.77 and -3.21 for Simulations 2 and 3, respectively.

The pattern of the bias in the standard logit model parameter estimates for Simulations 2 and 3 matches the predictions of Figure 1. With sorting towards recreation sites, we observe a negative correlation between travel costs and recreation demand as households select desirable sites to which they are proximate. As we can see from the simulation results in Figure 2, this results in overestimation of households' responsiveness to recreation costs—i.e., the parameter α on travel cost in indirect utility—which in turn leads to an underestimation of the (magnitude of the) willingness-to-pay statistic, since this statistic involves dividing by the travel cost parameter. Similarly, with sorting away from recreation sites, we observe a positive correlation between travel costs and recreation demand as households select remote, desirable sites. As is clear from Figure 2, this results in underestimation of households' responsiveness to recreation costs, which in turn leads to an overestimation of the (magnitude of the) willingness-to-pay statistic. Thus, regardless of the direction of the relationship between unobserved, idiosyncratic preferences and travel cost, the phenomenon of non-random sorting in these contexts presents a substantial challenge to standard discrete choice models of recreation demand.

[3]See Train (2009) and the discussion in Section 2.

3.2 Measurement Error in Travel Cost

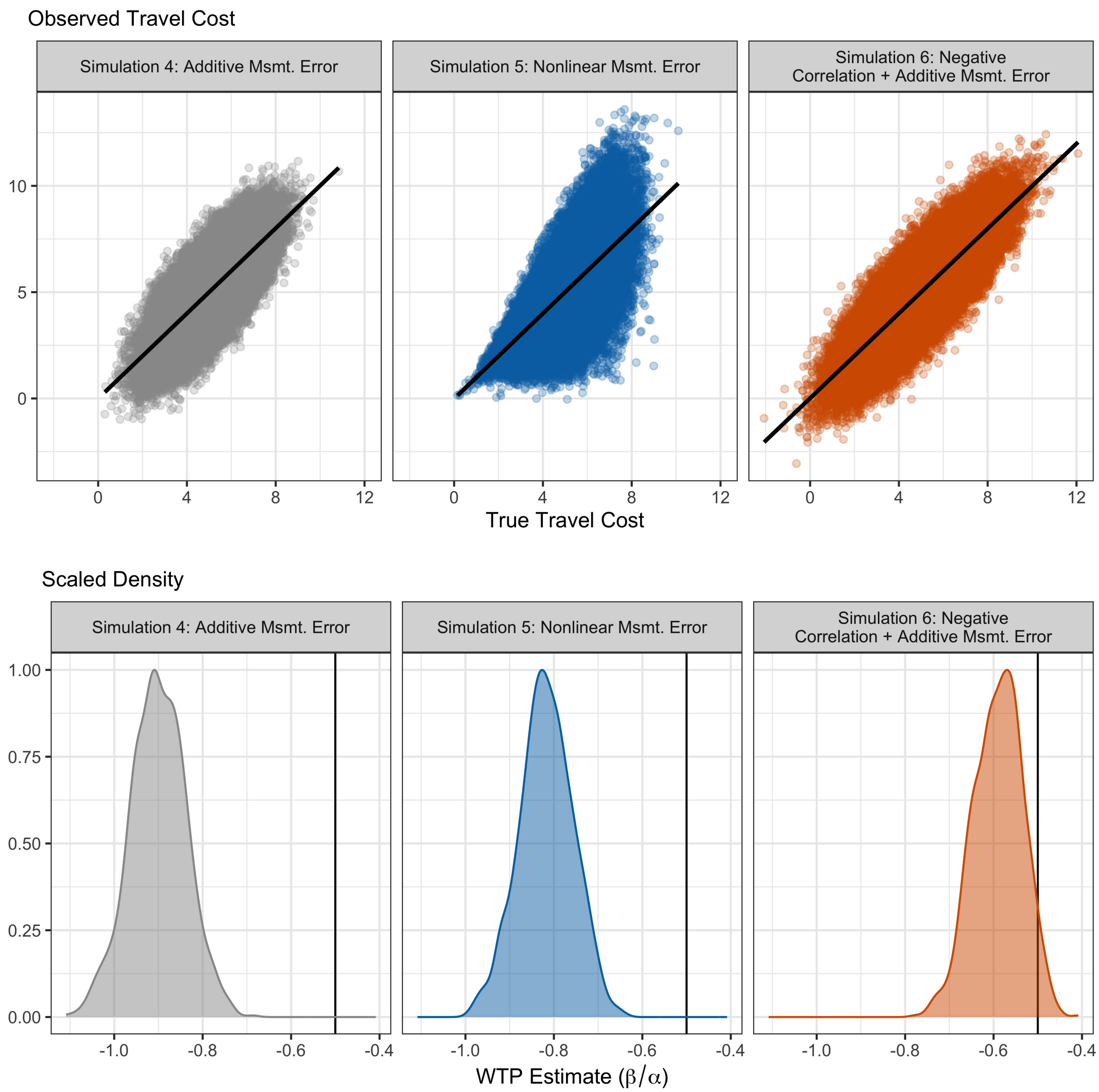
I turn now to a second issue in standard applications of recreation demand estimation, namely measurement error in travel costs. The basic logic of the recreation demand model assumes that the marginal cost of travel to a site is the price of consuming that site's amenities. The marginal cost of travel is inclusive of both the direct monetary costs of site visitation, including fuel, tolls, vehicle depreciation, or airfare, as well as the value of travel time, which is the value of foregone wages directly associated with travel. Unfortunately, analysts rarely—if ever—have access to the true cost of travel associated with realized trips, let alone the full set of travel costs for unselected trips in households' choice set.

As a result, analysts must use information on residence location and the location of sites in households' choice sets in combination with a set of assumptions to construct estimates of travel costs. While there is a well-defined set of best practices for doing so—see Lupi et al. (2020)—this approach to estimating travel costs can clearly result in both classical and non-classical (i.e., non-random) measurement error.

In particular, since driving is the main mode of transport in most recreation demand contexts, once analysts have calculated estimates of driving distances between households' residence location and all recreation sites in their choice sets, they must then use estimates of marginal fuel costs, tolls, and vehicle depreciation to translate these driving distances to monetary values. A common approach is to use data on average marginal driving costs published by automotive clubs such as the American Automobile Association (Lupi et al., 2020). While this may provide reasonable estimates of average per-mile driving costs, there is likely substantial heterogeneity in actual marginal driving costs that results in classical—i.e., mean zero—measurement error. Similarly, it is often difficult for analysts to observe households' opportunity costs of travel time. As a result, analysts often use some fixed proportion—usually between one third and one half (Lupi et al., 2020)—of average wages in the ZIP code or county of residence to calculate households' value of travel time. This combined with necessary assumptions about the speed of households' travel over driving distances can similarly introduce classical measurement error in travel costs.

Additional features of the travel cost construction process can lead to non-classical measurement error. As a best practice, analysts should condition on the road network between households' residence location and all recreation sites in their choice sets when calculating driving distances. While there are a growing number of tools to do so, including Google Maps and Open Source Routing Machine, this can be computationally and financially expensive. As a result, analysts often use measures of the linear distance between households' residence location and recreation sites as a proxy for driving distance. This almost surely *underestimates* the true driving distance. Moreover, the degree to which this underestimates true

Figure 3. Bias from Measurement Error in Travel Cost



Notes: This figure plots example relationships between households' true and observable travel costs from a single simulated dataset (top) as well as the distribution of estimated willingness-to-pay (WTP) across 1000 simulated datasets (bottom). The figure shows example relationships and WTP estimates for three assumed data generating processes: one where there is additive, mean-zero measurement error in travel costs (left); one where measurement error is increasing in travel distance (center); and one where there is both additive, mean-zero measurement error in travel costs as well as non-random sorting towards desirable recreation sites, where the idiosyncratic preference and true, unobserved travel costs are negatively correlated (right). The true value of the the willingness-to-pay statistic is shown as the vertical black line in the bottom panel. The full data generating process for each of simulations 4 to 6 are described in Section 3.1.

driving distance likely increases in the magnitude of the actual driving distance between a residence and recreation site, resulting in non-classical measurement error. Another issue in these settings is the assumed travel mode: as the distance to different sites in a household's choice set increases, so too does the probability that they choose to fly to a given site rather than drive. Though there are several noteworthy exceptions, including English et al. (2018), most recreation demand applications either explicitly or implicitly assume that households only drive to sites in their choice set, which can result in *overestimation* of the true cost of travel, particularly as the linear distance between a sites and a household increases.

Both classical and non-classical measurement error can lead to biased inferences from standard discrete choice models of recreation demand: similar to linear models, measurement error in travel costs leads to the familiar attenuation or regression dilution problem which leads to inconsistent, underestimation of model parameters. To be more precise about the nature of the measurement error problem in this context, I return to the baseline model outlined in Section 2. Setting aside the potential for non-random sorting by households, assume that the analyst evaluates the following model of household i's conditional utility from visiting recreation site j on choice occasion t:

$$u_{ijt} = X'_{ijt}\beta_i - \hat{c}_{ijt}\alpha_i + \xi_j + \tilde{\varepsilon}_{ijt} \tag{8}$$

where $\hat{c}_{ijt}$ is the estimated or observed travel cost for household i to visit site j at time t such that

$$\hat{c}_{ijt} = c_{ijt} + \underbrace{g(c_{ijt}) + \eta_{ijt}}_{\equiv \zeta_{ijt}} \tag{9}$$

Thus, the travel cost that the analyst observes is equal to the sum of the true travel cost, c_{ijt}; some function $g(\cdot)$ of the true travel cost; and some mean-zero shock to the true travel cost, η_{ijt}. On inspection, the structure of the measurement error problem described by (8) and (9) is analagous to the non-random travel cost issue outlined in Section 3.1: plugging (9) into (8) shows that both observed travel cost, $\hat{c}_{ijt}$, and household utility are functions of an unobserved shock to true travel cost and—possibly—some unknown function of true travel costs. This induces a correlation between household indirect utility and travel cost that, if the analyst ignores the data generating process (9) and makes a standard assumption such as $\tilde{\varepsilon}_{ijt}$ ∼T1EV will lead to biased estimates of the parameter α. Indeed, literature shows how ignoring the data generating process of (9) leads to biased parameter estimates: for example Kao and Schnell (1987) derives the asymptotic properties of a multinomial logit model with measurement error and shows that the parameter estimates from the standard logit estimator do not converge to the true values.

I demonstrate the bias from different forms of measurement error in the recreation demand context by adapting the data generating process of (6) and (7) to include imperfectly observed travel cost. I specify three distinct data generating processes to examine the bias from measurement error: in the first (Simulation 4), I ignore potential correlation between idiosyncratic preferences and true travel costs by setting $\rho_4 = 0.0$ and assume that the analyst observes the following travel cost:

$$\underline{\text{Simulation 4:}} \qquad \hat{c}_{ijt} = c_{ijt} + \eta_{ijt} \qquad \text{(Classical measurement error)}$$

where $\eta_{ijt} \sim \mathcal{N}(0.0, 1.0)$ and which corresponds to classical measurement error in travel costs. In the second measurement error data generating process (Simulation 5), I again ignore potential correlation between preferences and true travel costs by setting $\rho_5 = 0.0$ and assume that the analyst observes the following travel cost:

$$\underline{\text{Simulation 5:}} \qquad \hat{c}_{ijt} = c_{ijt}(1 + \eta_{ijt}) \qquad \text{(Non-classical measurement error)}$$

where again $\eta_{ijt} \sim \mathcal{N}(0.0, 1.0)$ and which corresponds to non-classical measurement error in travel costs. In the third and final measurement error data generating process (Simulation 6), I allow for both non-random selection into true travel costs and measurement error in observed travel costs. In particular, I set $\rho_6 = -1.0$ and assume that there is classical measurement error of the same classical form as in Simulation 4:

$$\underline{\text{Simulation 6:}} \qquad \hat{c}_{ijt} = c_{ijt} + \eta_{ijt} \qquad \text{(Classical measurement error)}$$

where again $\eta_{ijt} \sim \mathcal{N}(0.0, 1.0)$.

As with the simulations in Section 3.1, I generate 1000 unique choice datasets for each of the three data generating processes defined by Simulations 4 through 6, including both true and observed travel costs. Once again, each dataset consists of 1000 individuals choosing between 100 alternative sites. The top panel of Figure 3 plots the empirical relationship between true travel costs and travel costs observed by the analyst for an example simulated choice dataset for each simulation.

Having simulated 1000 choice datasets for each of the three simulations, I again make the standard assumption that $(\xi_{ij} + \zeta_{ij} + \tilde{\varepsilon}_{ij}) \sim$ T1EV—i.e., ignore the data generating process for travel costs—where ζ_{ij} is the residual from the observed travel cost data generating process. This allows me to estimate the parameters of the linear indirect utility model in (6) for each simulated dataset via a multinomial logit model, thereby generating distributions of parameter estimates from the standard model of recreation demand for each data generating process. Importantly, however, I now use the simulated observed travel costs, $\hat{c}_{ijt}$ to estimate the parameters of the model and then compare estimates of the marginal rate of substitution between x_{ij} and c_{ij} when evaluating the relative performance of the standard multinomial logit estimator in the presence of measurement error.

The bottom panel of Figure 3 plots the empirical distributions of willingness-to-pay estimates from the standard logit estimator across the three different measurement error data generating processes. Unsurprisingly, the presence of measurement error results in poor coverage of the true target statistic: across all three sets of simulations, average estimates of the willingness-to-pay measure are between -0.58 and -0.90. This is consistent with measurement error producing attenuation in estimates of the travel cost parameter: a smaller-in-magnitude travel cost parameter estimate results in a larger-in-magnitude willingness-to-pay measure, all else equal. Moreover, I can reject the null hypothesis that the distribution of willingness-to-pay estimates cover the true value of -0.5 for Simulations 4 and 5, with t-statistics of -6.36 and -5.23, respectively. In the case of Simulation 6, the negative correlation between the idiosyncratic preference, ξ_{ijt}, and true travel costs works to partially offset this trend, resulting in slightly better coverage of the true value of the statistic.

4 Solution: An Instrumental Variables Estimator

While the challenges of non-random sorting on preferences for outdoor recreation and measurement error in travel costs can bias estimates from standard discrete choice models of recreation demand, a relatively simple class of alternative estimators can circumvent these issues. This section presents an instrumental variables estimator that is analogous to two-stage least squares in the nonlinear context of standard discrete choice models. The estimator, referred to in the literature as a two-stage control function approach, is relatively straightforward to implement and, with some simple additional assumptions outperforms the baseline models in numerical simulations. I begin my exposition of this estimator in the case on non-random sorting assuming the analyst perfectly observes true travel costs and then discuss the performance of this estimator with imperfectly observed travel costs.

4.1 Control Function Approach and Endogenous Travel Cost

Based on the model of indirect utility and travel costs outlined by (4) and (5), there exists a non-zero correlation between the unobserved, idiosyncratic preference, ξ_{ijt}, and the residual travel cost term, ϵ_{ijt}. Assuming that it is possible to observe ϵ_{ijt}, I can decompose the unobservable preference term, ξ_{ijt}, into its mean conditional on ϵ_{ijt} and deviations around this mean:

$$\xi_{ijt} = \mathbb{E}[\xi_{ijt}|\epsilon_{ijt}] + \tilde{\xi}_{ijt} \tag{10}$$

The conditional expectation in (10) is a function of ϵ_{ijt} and can be approximated using a control function:

$$\xi_{ijt} = CF(\epsilon_{ijt}; \lambda) + \tilde{\xi}_{ijt} \tag{11}$$

where λ parameterizes the function $CF(\cdot)$. The simplest assumption is that

$$CF(\epsilon_{ijt}; \lambda) = \lambda\epsilon_{ijt} \tag{12}$$

Substituting (11) and (12) into (4) gives:

$$u_{ijt} = v_{ijt}(X_{ijt}; \theta) + \lambda\epsilon_{ijt} + \tilde{\xi}_{ijt} + \tilde{\varepsilon}_{ijt} \tag{13}$$

where by construction both $\tilde{\xi}_{ijt}$ and $\tilde{\varepsilon}_{ijt}$ are idiosyncratic, unobserved, and independent.

It is possible to take the model implied by (13) to the data to recover unbiased estimates of the target parameters, θ. Doing so requires a set of assumptions about the residual terms $\tilde{\xi}_{ijt}$ and $\tilde{\varepsilon}_{ijt}$ as well as access to a set of valid instruments, z_{ijt}, for travel cost. While a number of different assumptions on the structure of the unobserved terms $\tilde{\xi}_{ijt}$ and $\tilde{\varepsilon}_{ijt}$ are plausible, the simplest is to treat each as an error component and assume that their sum is independently and identically distributed T1EV, i.e., $(\tilde{\xi}_{ijt} + \tilde{\varepsilon}_{ijt})$ ∼T1EV. This assumption then allows me to estimate (13) as a logit model with an additional observable variable, ϵ_{ijt}, and parameter, λ. For a discussion of additional possible assumptions on the distribution of these error terms, see Train (2009) and Petrin and Train (2010).

Armed with an assumption on the distribution of the terms $\tilde{\xi}_{ijt}$ and $\tilde{\varepsilon}_{ijt}$, it is possible to specify an estimator from (13) that identifies the true target parameters, θ. However, I need to have consistent estimates of the residual term ϵ_{ijt} from the first stage formula for travel costs (5) in order to do so. This requires access to a valid travel cost instrument, z_{ijt}. In particular, the travel cost instrument must satisfy the following relatively standard conditions: instrument relevance, i.e., $Cov(c_{ijt}, z_{ijt}) \neq 0$, and instrument exogeneity, i.e., $Cov(z_{ijt}, \xi_{ijt}) = 0$. These assumptions are common to other instrumental variables estimators (Angrist and Krueger, 2001).

Assuming that the analyst has access to a travel cost instrument that satisfies the above two conditions, z_{ijt}, it is possible to implement a two-stage control function estimator that identifies the true parameters. Estimation proceeds as follows:

1. First (5) is estimated: this is a regression with the endogenous travel cost variable as the dependent variable and the exogenous instrument, z_{ijt}, as the explanatory vari-

able. While it is possible to flexibly specify the functional form of $w(z_{ijt}, \gamma)$, a simple assumption is that z_{ijt} enters linearly such that the instrument enters (5) additively and the parameters γ are recovered by ordinary least squares. The residuals from this first stage regression provide estimates of ϵ_{ijt}:

$$\hat{\epsilon}_{ijt} = c_{ijt} - w(z_{ijt}; \hat{\gamma})$$

2. In the second step, a discrete choice model—such as a random parameters logit—is estimated with the first stage residual, $\hat{\epsilon}_{ijt}$ entering as an additional term. Estimation follows the same likelihood routine as that for the baseline model described in Section 2.

Thus, with a set of parametric and distributional assumptions on the nature of the correlation between travel cost and the unobserved, idiosyncratic preference term, it is possible to account for the endogeneity problem and recover unbiased parameter estimates. Moreover, while the linear assumption on the expectation of ξ_{ijt} conditional on ϵ_{ijt} may appear strong, it is possible in practice to allow for more flexible specifications of the control function (11) at minimal additional computational cost.

It is important to note that inference is non-trivial in the context of this relatively simple two-stage estimator. A general feature of estimation in multiple stages is that noise from earlier stages of estimation enters later stages, which means that the covariance matrix for the final estimates must reflect this additional source of error. Karaca-Mandic and Train (2003) and Petrin and Train (2010) derive the asymptotic covariance matrix of the second stage estimates in a two-stage control function estimator of a multinomial logit model and demonstrate the importance of accounting for this additional source of error in an empirical setting. It is also possible to adjust the second stage standard errors by bootstrapping the two-stage procedure (Petrin and Train, 2010).[4]

4.2 Control Function Approach and Measurement Error

It is relatively trivial to adapt the exposition in Section 4.1 to the case of measurement error in travel costs. Ignoring potAssuming that true travel costs follow an analogous data generating process as (5), i.e.,

$$c_{ijt} = w(z_{ijt}; \gamma) + \mu_{ijt}$$

[4] In the context of two-stage control function estimation of a recreation demand model, this would involve re-implementing the two-stage procedure across a number of bootstrapped samples of choice occasions and then taking the empirical covariance of second stage parameter estimates across bootstrap samples as the estimate of the second stage covariance matrix.

then it is possible to express observed travel costs (9) as

$$\hat{c}_{ijt} = w(z_{ijt};\gamma) + \underbrace{\mu_{ijt} + \zeta_{ijt}}_{\equiv e_{jmt}} \tag{14}$$

Assuming that it is possible to observe e_{ijt}, I can specify a control function for the unobservable term representing measurement error in travel costs, ζ_{ijt}, which enters both observed travel costs and, as a result, indirect utility. In particular, I can decompose this unobservable error term, ζ_{ijt} into its mean conditional on e_{ijt} and an independent component:

$$\zeta_{ijt} = \mathbb{E}[\zeta_{ijt}|e_{ijt}] + \tilde{\zeta}_{ijt} \tag{15}$$

The conditional expectation in (15) is itself a function of e_{ijt} and can be approximated using a control function:

$$\zeta_{ijt} = CF(e_{ijt};\nu) + \tilde{\zeta}_{ijt} \tag{16}$$

where ν parameterizes the function $CF(\cdot)$. Once again, the simplest assumption is that

$$CF(e_{ijt};\nu) = \nu e_{ijt} \tag{17}$$

Substituting (9), (16), and (17) into (8) gives:

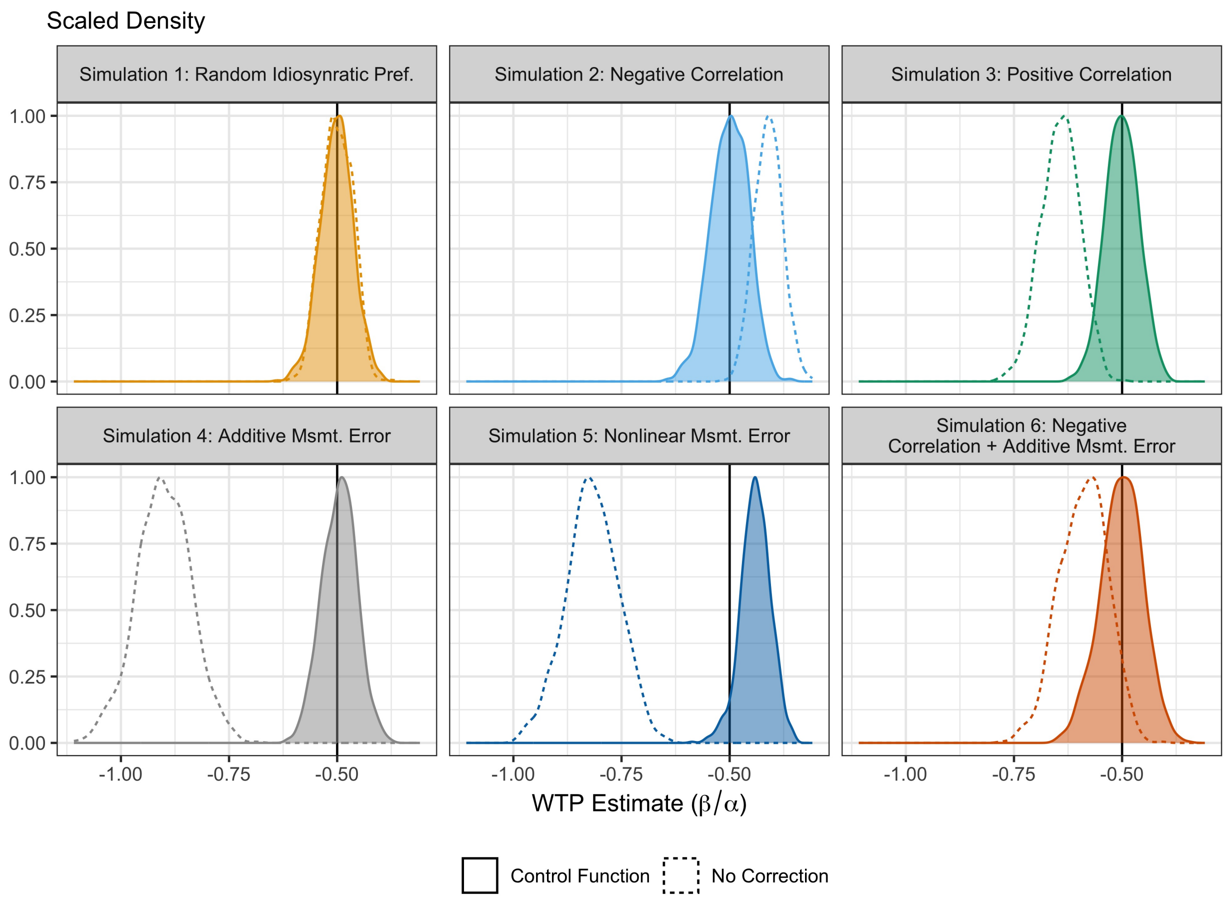
$$u_{ijt} = \underbrace{X'_{ijt}\beta_i - \hat{c}_{ijt}\alpha_i + \xi_j}_{v_{ijt}(X_{ijt};\theta)} + \nu e_{ijt} + \tilde{\zeta}_{ijt} + \tilde{\varepsilon}_{ijt} \tag{18}$$

Making an analogous set of parametric and distributional assumptions on the nature of the correlation between observed and true travel costs and assuming the analyst observes an accurately-measured instrument for travel cost, z_{ijt}, (18) shows that is possible to account for the measurement error problem and recover unbiased parameter estimates. Similar to the solution that I present in Section 4.1, implementing this solution requires estimation in two steps: in the first stage, observed travel costs are regressed on a valid cost instrument which generates estimates of the residual term, $\hat{e}_{ijt}$. In the second stage, these estimates are included in a standard discrete choice model of recreation demand, such as the multinomial logit or random parameters logit.

4.3 Performance of the Control Function Approach

I now turn to evaluating the performance of this alternative estimator. I implement the two-stage control function estimator outlined above on the simulated choice data from

Figure 4. Performance of Two-stage Control Function Estimator



Notes: This figure plots the distribution of estimated willingness-to-pay (WTP) using the baseline multinomial logit estimator and two-stage control function estimator across 1000 simulated datasets for each of the 6 simulations described in Section 3. Simulations 1 through 3 assume different non-random sorting data generating processes: one where the idiosyncratic preference and travel cost are independent (left); one corresponding to a model of household sorting towards desirable recreation sites, where the idiosyncratic preference and travel cost are negatively correlated (center); and one corresponding to a model of household sorting away from desirable recreation sites, where the idiosyncratic preference and travel costs are positively correlated (right). Simulations 4 through 6 assume different measurement error data generating processes: one where there is additive, mean-zero measurement error in travel costs (left); one where measurement error is increasing in travel distance (center); and one where there is both additive, mean-zero measurement error in travel costs as well as non-random sorting towards desirable recreation sites, where the idiosyncratic preference and true, unobserved travel costs are negatively correlated (right). The true value of the the willingness-to-pay statistic is shown as a vertical black line.

Simulations 1 through 6 from Section 3. Figure 4 compares the distribution of estimated willingness-to-pay statistics obtained from the standard logit estimator and the two-stage control function estimator for all six simulations. Apart from Simulation 1, which has no correlation between unobserved preferences and travel costs and no measurement error, the estimates from the two-stage control function estimator outperform the baseline logit es-

Table 1. Willingness-to-pay Estimates for Simulated Choice Data

	Baseline Estimator				Control Function			
	Mean	Bias	MSE	t-stat	Mean	Bias	MSE	t-stat
Baseline Simulation								
Simulation 1	−0.500	−0.000	0.001	−0.005	−0.501	−0.001	0.001	−0.034
Non-random Sorting Simulations								
Simulation 2	−0.410	0.090	0.009	2.770	−0.499	0.001	0.002	0.013
Simulation 3	−0.643	−0.143	0.022	−3.205	−0.500	−0.000	0.002	−0.002
Measurement Error Simulations								
Simulation 4	−0.901	−0.401	0.165	−6.357	−0.495	0.005	0.002	0.127
Simulation 5	−0.817	−0.317	0.104	−5.231	−0.440	0.060	0.005	1.724
Simulation 6	−0.589	−0.089	0.011	−1.618	−0.501	−0.001	0.002	−0.026

Notes: This table reports summary statistics for willingness-to-pay estimates using the baseline multinomial logit estimator and two-stage control function estimator across 1000 simulated datasets for each of the 6 simulatinos described in Section 3. Simulations 1 through 3 assume different non-random sorting data generating processes: one where the idiosyncratic preference and travel cost are independent (left); one corresponding to a model of household sorting towards desirable recreation sites, where the idiosyncratic preference and travel cost are negatively correlated (center); and one corresponding to a model of household sorting away from desirable recreation sites, where the idiosyncratic preference and travel costs are positively correlated (right). Simulations 4 through 6 assume different measurement error data generating processes: one where there is additive, mean-zero measurement error in travel costs (left); one where measurement error is increasing in travel distance (center); and one where there is both additive, mean-zero measurement error in travel costs as well as non-random sorting towards desirable recreation sites, where the idiosyncratic preference and true, unobserved travel costs are negatively correlated (right).

timates, as expected. In the case of the two sets of simulations with non-zero correlation between unobserved site preferences and travel costs, the control function estimator reduces bias in willingness-to-pay estimates by one to two orders of magnitude. Furthermore, as shown in Table 1, it is not possible to reject the null hypothesis of equivalence to the true willingness-to-pay value at any conventional level of statistical significance, with estimated t-statistics of 0.013 and −0.002 for the control function estimates in Simulations 2 and 3, respectively.

In the case of the simulations involving measurement error, the specified data generating processes result in quite considerable bias from the standard logit estimator.[5] In Simulations 4 and 5, which assume additive, independent measurement error and non-linear measurement error, respectively, the control function estimator greatly reduces the magnitude of this bias, again by several orders of magnitude. Interestingly, the performance of the control function estimator is slightly attenuated in absolute terms with non-classical measurement

[5] This is an artifact of the specific magnitudes assumed for the data-generating processes; however, the patterns in relative performance across the two estimators holds independent of these choices.

error in Simulation 5, though the estimator still substantially outperforms the standard logit estimator. However, I am still unable to reject the null hypothesis of equivalence of the willingness-to-pay estimates from the control function estimator to the true value at a 95% confidence level based on the t-statistic reported in Table 1 of 1.72.

Unsurprisingly, the performance of the two-stage control function estimator remains strong with both non-random sorting and measurement error in travel costs. As shown in Figure 4, the control function estimator has better coverage of the true willingness-to-pay value than the standard logit estimator across the 1000 simulated choice datasets in Simulation 6. Since the negative correlation between site preferences and true travel costs and the classical measurement error in travel costs bias the travel cost parameter estimate in opposite directions under the standard logit estimator, the relative performance gain of the control function estimator is less than in Simulations 2 through 5; however, as shown in Table 1, the absolute performance of this approach to bias correction is strong, resulting in minimal bias and failure to reject equivalence to the true willingness-to-pay statistic at any conventional level of statistical significance.

Taken together, the results shown in Figure 4 and Table 1 provide strong support in favor of the two-stage control function approach to address a range of potential biases in conventional recreation demand models. Regardless of the source or direction of the bias in practice, this alternative estimator delivers relative performance improvements over conventional estimators of discrete choice models of recreation demand that ignore potential practical issues with travel costs. These simulation results should encourage analysts and practitioners to implement this relatively simple two-stage correction to ensure valid inferences from recreation demand models in empirical settings.

5 Empirical Application: Deepwater Horizon Oil Spill

In this section, I apply the insights of the numerical simulations in Sections 3 and 4 to a real-world empirical setting. In particular, I examine the performance of the two-stage control function approach in the context of a recreation demand model that is used to monetize lost shoreline recreation associated with the 2010 Deepwater Horizon (DWH) oil spill in the Gulf of Mexico. The DWH spill, which occurred following the explosion and sinking of an offshore drilling rig 50 miles off the Louisiana coastline in April 2010, lasted 87 days and resulted in the release of 134 million gallons of oil into the Gulf of Mexico, making it the largest oil spill in US history.

English et al. (2018) estimates that the monetary losses from foregone shoreline recreation following the DWH spill totaled \$661 million. This estimate played a substantial role in compensatory litigation in the aftermath of the event: in response to the DWH spill, the National Oceanic and Atmospheric Administration (NOAA) initiated a process of assessing the recreation-related welfare losses for the purposes of pursuing compensation on the public's behalf under the authority of the Oil Pollution Act of 1990. This 5-year, multi-million dollar effort employed a comprehensive strategy involving primary data collection and recreation demand modeling. The recreation demand modeling component of this broader NOAA-led effort, which is outlined in detail in English et al. (2018), is the primary focus of the empirical application in this section.

5.1 Empirical Setting and Data

The NOAA-led assessment of recreational damages from the DWH spill employed two primary data collection methods, including (1) infield surveys of on-site recreational activities and (2) telephone surveys of adult heads-of-household in the continental US. The infield surveys included 129,000 in-person interviews, 35,000 onsite counts, and nearly 500,000 aerial photographs and were collected over the three years beginning immediately after the DWH spill in May 2010. These infield surveys form the basis of estimates of lost user days due to the spill by year, month, and area, which are described in detail by Tourangeau et al. (2017). While the primary data collected across these different infield survey instruments do not directly enter the recreation demand modeling of English et al. (2018), the estimated reduction in recreational user days over time and across sites reported in Tourangeau et al. (2017) help calibrate monetary losses resulting from the DWH spill as I describe below.

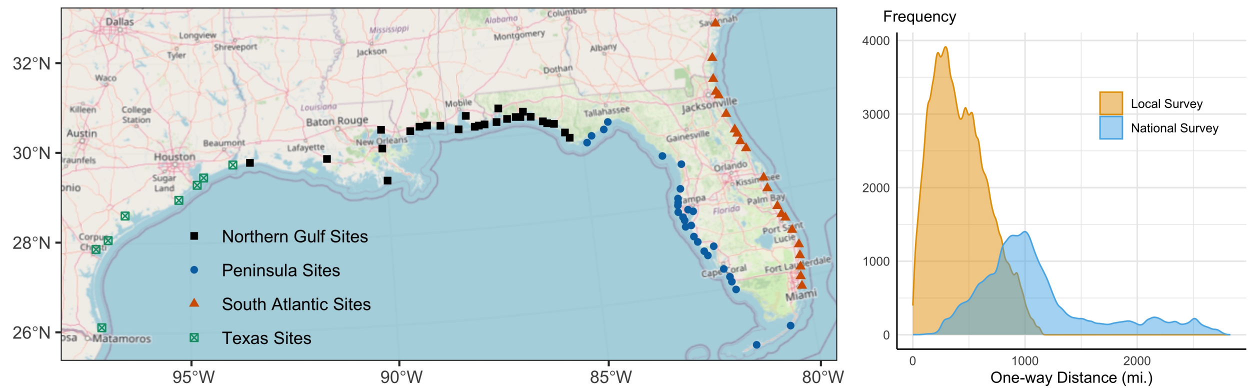
The second source of primary data for this broader damage assessment effort—local and nationwide telephone surveys of adult heads-of-household—form the basis of the choice dataset used to estimate a discrete choice model of demand for recreation at Gulf Coast sites. These phone surveys, which were based on samples of adults in the continental US, included 244,000 mail survey screeners and 43,000 telephone interviews.[6] The telephone surveys were conducted from 2012 to 2013 and collected information on any recent shoreline trips to coastal areas in Texas, Louisiana, Mississippi, Alabama, Florida, and Georgia (henceforth, "study area"). The surveys solicited detailed household demographic data as well as the precise location of any recent shoreline recreation visits to the study area. In the recreation demand modeling described below, the phone survey data are weighted to better reflect the target population, namely Gulf Coast recreational users that reside in the contiguous US.[7]

[6] Mail survey screeners were used to collect phone numbers and identify households that had recreated on the Gulf Coast in the past to allow for oversampling of Gulf Coast recreationists (English et al., 2018).

[7] In particular, English et al. (2018) construct a set of weights that (1) account for sampling probabilities in populated areas; (2) correct for non-response by selected geographies; (3) correct for the intentional oversampling of Gulf Coast recreationists based on the mail screener; (4) post-stratify to match the number

Figure 5. Recreation Sites and The Distribution of Distance in English et al. (2018) Sample

(a) Location of Sites in Sample **(b)** Distance to Sites



Notes: Panel (a) maps the 83 aggregate shoreline sites included English et al. (2018)'s choice set. Panel (b) plots the distribution of one-way driving distances between the full set of shoreline sites and survey respondents' residences separately for the local and nationwide surveys. The shoreline sites are differentiated into four groups: Texas Sites, Northern Gulf Sites, Peninsula Sites, and South Atlantic Sites. For the purposes of assessing monetary damages lost recreational user days, English et al. (2018) identify the Northern Gulf and Peninsula Sites as the affected sites in respondents' choice set.

English et al. (2018) aggregate surveyed households' visit locations into 83 distinct sites that span roughly 2,300 miles of coastline from Texas to Georgia. Figure 5 shows the location of these sites as well as the regional groupings that English et al. (2018) use to analyze the impact of the DWH spill: Texas, Northern Gulf, Peninsula, and South Atlantic. Of the 83 total sites, the authors define the 54 sites in the Northern Gulf and Peninsula regions as those adversely affected by the spill. These 83 aggregate sites represent the full choice set available to respondents when English et al. (2018) construct the data that they use to estimate a discrete choice model of demand for Gulf Coast recreation.

In order to calculate the travel cost associated with visiting the 83 aggregate sites in respondents' choice set, English et al. (2018) follow best practices in the recreation demand literature while also making several noteworthy innovations on existing methods. In particular, recognizing that a non-trivial number of respondents likely chose to fly to their Gulf Coast destination given the geographic scale of the target population, English et al. (2018) solicit mode choice information in their phone survey to generate flying and driving probabilities as a function of one-way driving distance to a given site as well as a subset of household demographics. These flight probabilities are used in combination with driving distances and detailed data on marginal driving costs as well as expected flying costs to construct expected

of households in aggregate geographies included in the target, continental US population; (5) adjust for the number of residents in respondents' household; and (6) re-weight for aggregate observable demographic characteristics. For additional information see English et al. (2018).

travel costs for every respondent-shoreline site combination. Given that the calculation of travel costs has important implications for how we think about possible bias in standard discrete choice models of recreation demand, I will discuss additional aspects of English et al. (2018)'s calculations in detail below; however, I refer the interested reader to English et al. (2018) for a complete description of these calculations.

Given the public funding for this assessment, all data—from processed respondent data to the final data used in recreation demand estimation—are publicly available. In particular, I obtain the data that English et al. (2018) use directly from NOAA's Natural Resource damage Assessment (NRDA) public repository.[8]

5.2 Model

English et al. (2018) use a nested logit model to characterize demand for shoreline recreation on the Gulf Coast. The authors use the survey data collected from contiguous US households from 2012 to 2013 to estimate the parameters of the nested logit model under baseline, non-spill conditions and calibrate the estimated model to match observed declines in recreational user days immediately following the spill in order to compare recreational values during spill and non-spill conditions.

The benefit of the nested logit in this context is that it allows the authors to capture the extensive margin of Gulf Coast recreation demand: the upper nest models households' decision of whether or not to visit a Gulf Coast shoreline site and—conditional on choosing to visit a site—the lower nest models households' choice between sites. Building on the notation of the standard model that I present in Section 2, let u_{ij} denote the conditional utility received by individual $i \in \{1, \dots, N\}$ when selecting Gulf Coast shoreline site $j \in \{0, \dots, J\}$, where site $j = 0$ denotes the outside option of choosing to not visit a Gulf Coast site and $j > 0$ denotes the inside options of the 83 distinct shoreline sites.[9] In the version of English et al. (2018)'s model that I implement, individual i's conditional utility from visiting shoreline site j is $u_{ij} = v_{ij} + \varepsilon_{ij}$ with:

$$v_{ij} = \begin{cases} 0 & \text{for } j = 0 \\ \xi_j - c_{ij}\alpha & \text{for } j \in \{1, \dots, J\} \end{cases} \tag{19}$$

where I normalize the observable component of the flow utility from the no visit option, $j = 0$, to zero; ξ_j is a site-specific constant representing mean valuations of that site; and

[8]The Deepwater Horizon NRDA data are available for download here: `https://www.diver.orr.noaa.gov/deepwater-horizon-nrda-data` (last accessed 2/29/2024).

[9]Note that in the behavioral choice data, English et al. (2018) do not observe households making visitation decisions on repeat occasions, so I suppress the t subscript that I use in the exposition in Section 2.

c_{ij} is household-site-specific travel cost.[10] Note that in contrast to the more general model presented in Section 2, (19) suppresses individual-specific heterogeneity in preferences.

In their model, English et al. (2018) assume that ε_{ij} follows a generalized extreme value distribution that implies a two-level nesting structure. With the specification of conditional utility (19), this implies choice probabilities of the following form:

$$p_{ij} = \begin{cases} \frac{1}{1+\left(\sum_{k=1}^{J}\exp(v_{ik}/\rho)\right)^{\rho}} & \text{for } j = 0 \\ \frac{\exp(v_{ij}/\rho)}{\sum_{k=1}^{J}\exp(v_{ik}/\rho)} \times \frac{\left(\sum_{k=1}^{J}\exp(v_{ik}/\rho)\right)^{\rho}}{1+\left(\sum_{k=1}^{J}\exp(v_{ik}/\rho)\right)^{\rho}} & \text{for } j \in \{1,\dots,J\} \end{cases} \tag{20}$$

where ρ is a nesting parameter or "dissimilarity coefficient" that proxies for the degree of preference correlation within groups.

I take the nested logit model implied by (19) and (20) to the data, estimating the target parameters, $\begin{bmatrix}\alpha & \rho & \xi_1 & \dots & \xi_J\end{bmatrix}$, via maximum likelihood estimation. This represents the analog to the model estimates of English et al. (2018), which I interpret as representing the baseline, standard approach to discrete choice recreation demand modeling in this context. In light of the potential challenges with this standard approach that I discuss in Section 3, I also implement the two-stage control function estimator in this setting. As outlined in Section 4, this involves regressing travel cost on a set of valid instruments, Z_{ij}, and then plugging the residual from this first stage, $\hat{\mu}_{ij}$, into the second stage nested logit as an observable with an additional target parameter, λ. All that remains to implement this alternative estimator in the context of English et al. (2018)'s model of Gulf Coast shoreline recreation is to identify a set of valid travel cost instruments.

5.3 Valid Travel Cost Instruments

A valid instrument for travel costs must satisfy the relatively standard conditions of relevance and exogeneity: in other words, the instrument must plausibly affect idiosyncratic travel costs, but be independent of households' demand for outdoor recreation. There are likely many possible empirical instruments that analysts can use and—much like other research designs that rely on instrumental variables—the ideal choice of instruments is likely context-specific. However, given the nationwide scale of the current empirical setting, I seek to identify several possible travel costs instruments that may have relatively broad application

[10] English et al. (2018)'s model allows household demographics to enter the flow utility of non-visitation. Given that my primary focus is on the estimation of the travel cost parameter, α, across different estimators, I omit this richer specification. As a result, the estimates of lost user day value that I estimate are not directly comparable to those in English et al. (2018); however, the relative differences that I estimate across estimators are nonetheless of independent interest and should apply to the findings of English et al. (2018) and more broadly.

in the recreation demand literature.

Considering the data-generating process for travel costs and the methods used to construct these measures in practice, there are two broad categories of potential travel cost instruments: those that influence households' choice of residence location, but not recreation site choices; and those that influence the marginal cost of site visitation, but not recreation site choices.

An example of the first category of instruments is a measure of a household's expected time spent commuting to work. This is likely correlated with households' choice of residence location, which in turn determines the distribution of travel costs associated with visitation of a set of recreation sites: as shown in Appendix Figure A1, there is substantial cross-sectional variation in average commuting times across the US, which suggests that households are likely able to factor this into their choice of residence. Conditional on residence choice, it is plausible that a household's commuting time for work has little impact on their choice of recreation site. Moreover, the average values in Appendix Figure A1 likely mask substantial local heterogeneity in commuting times, which suggests that analysts may be able to use granular measures of employment commuting times for smaller-scale recreation demand contexts than the nationwide model on which I focus.

Given the substantial heterogeneity in average commuting time across Zip codes in the US as shown in Appendix Figure A1, I use this as one instrument for travel cost in the present context. In particular, I use average commuting time in respondents' Zip Code Tabulation Area of residence, which I take from the US Census Bureau's 2012 5-year American Community Survey, as an instrument for travel cost in the present model of contiguous US demand for Gulf Coast shoreline recreation.

There are two plausible examples of the second category of instruments, i.e., those that influence the marginal cost of site visitation directly. In the case of the US, there is substantial heterogeneity across states and over time in the level of gasoline tax rates as shown in Appendix Figure A2. I use variation in the rate of the gasoline tax in the state of destination at the time a respondent is surveyed interacted with one-way driving distance as an instrument for travel cost in the present context. Though there is non-trivial variation in gasoline tax rates across the six states in the study area, I interact these rates with one-way driving distance as households are likely more sensitive to variation in destination taxes for further away sites that involve more driving and therefore gasoline consumption. While these tax rates are clearly correlated with households' expected travel costs, this is likely the only channel through which these tax rates influence recreation demand.

An additional example of an instrument that influences the marginal cost of site visitation directly is the price of crude oil. In the present context, I interact the West Texas Intermediate benchmark price with one-way driving distance to each site and use the resulting variable as an additional travel cost instrument. This is clearly correlated with travel cost, but is likely exogenous to unobserved characteristics. While this is not directly testable, it is unlikely that a crude oil price benchmark interacted with driving distance is correlated with factors influencing recreation demand other than directly through travel cost as those factors which determine world oil prices are plausibly different from those which would enter both individuals' recreation decisions. However, it is possible that seasonality in crude oil prices or macroeconomic trends that influence the global crude oil market also affect household recreation demand.

In light of this, I also follow the approach of Kilian (2009) to isolate structural supply and demand shocks in the global crude oil market from seasonal or aggregate demand fluctuations. I describe the method in detail in Appendix B. I interact the resulting crude oil supply and demand shock time series, which I plot in Appendix Figure A3, with one-way driving distance and use these as two additional instruments for travel cost. These instruments follow a "shift-share" logic: shocks to global crude oil demand or supply work to change the relative prices of sites at different distances from a household. Since these shocks to the global crude oil market are plausibly exogenous to household recreation demand, these shifts in relative prices should serve as valid instruments for travel cost.

5.4 Results

I report estimates from four separate first stage regressions of individual- and site-specific travel cost on a set of instruments and alternative specific constants in Table 2. I make a simple functional form assumption for the first stage where each excluded instrument enters travel cost linearly. The four specifications in Table 2 vary the crude oil price instrument(s) and whether sample weights are used. Overall, the estimated coefficients on the excluded instruments in the first stage regressions all have the expected sign and appear to explain a substantial share of variation in travel cost across individuals and sites. The first stage F-statistics of joint nullity all well exceed conventional rule-of-thumb cutoffs for weak instruments employed in the two-stage least squares literature.

Taken together, Table 2 suggests that the selected instruments all satisfy the relevance condition and the resulting residuals likely isolate the component of travel cost that is correlated with unobserved preferences for recreation sites or measurement error. Given the possible concern around the use of the WTI crude oil benchmark discussed in Section 5.3 and to ensure consistency with the observation weights used in the second stage, I use residuals from the regression reported in Column (4) of Table 2 in estimating the second stage.

Table 3 reports parameter estimates from maximum likelihood estimation of the second

Table 2. First Stage Estimates

	Travel Cost			
	(1)	(2)	(3)	(4)
Origin Commute Time	0.328	0.306	1.25	1.25
	(0.091)	(0.090)	(0.096)	(0.098)
Destination Gas Tax × Distance	0.020	0.028	0.015	0.020
	(0.005)	(0.0003)	(0.003)	(0.0002)
WTI × Distance	0.001		0.0010	
	(0.0009)		(0.0006)	
Oil Supply Shock × Distance		−0.042		−0.003
		(0.003)		(0.011)
Oil Demand Shock × Distance		0.0008		0.0006
		(0.0002)		(0.0005)
Alternative Specific Constants	Yes	Yes	Yes	Yes
Sample Weights	No	No	Yes	Yes
Observations	3,462,428	3,462,428	3,462,428	3,462,428
R^2	0.469	0.468	0.346	0.345
Within R^2	0.451	0.451	0.343	0.342
F-statistic	2,568.7	14,752.6	3,120.6	3,058.1

Notes: This table reports estimates from a series of first stage regressions of individual- and site-specific travel cost on a set of instruments and alternative specific constants. See Section 5.3 for a discussion of the different excluded instruments. Columns (1) and (2) do not weight observations whereas Columns (3) and (4) use the sample weights constructed by English et al. (2018) to weight observations.

stage discrete choice model of demand for visits to Gulf Coast shoreline recreation sites. For concision, I omit reporting the 83 site-specific constants for each estimator. Column (1) of Table 3 reports estimates from a standard nested logit model that is analogous to the main estimates in English et al. (2018)—though, as noted in Section 5.2, I do not model observable demographic heterogeneity in preferences for the outside option. Despite the different specification, the travel cost and nesting parameters are quite similar to those reported by English et al. (2018): my estimate of the travel cost parameter is negative in sign, large, and highly statistically significant and my estimate of the nesting parameter implies a similar degree of within-nest correlation in preferences as that found by English et al. (2018).

Column 2 of Table 3 reports estimates from the analogous two-stage control function estimator described in Section 5.2. Reported standard errors for this model adjust for noise in the first stage residuals using the asymptotic covariance formula from Karaca-Mandic and Train (2003). While the estimated nesting parameter with the control function adjustment is similar in magnitude to the baseline estimate, the estimated travel cost parameter substantially increases in magnitude: households' mean sensitivity to travel cost increases in magnitude from -1.08 in the uncorrected, baseline estimates to -1.51 in the control function estimates, a 40% increase. Thus, it appears as though failing to account for potential non-random selection travel costs and/or measurement error in travel costs biases estimates of the travel cost parameter towards zero, suggesting that standard models in this context underestimate households' price sensitivity.

Perhaps unsurprisingly given the direction of the apparent bias in the baseline estimates, the first stage travel cost residuals enter the second stage positively: the estimated parameter on the travel cost residuals is positive, large-in-magnitude, and highly statistically-significant. A positive parameter on the first stage residuals in households' indirect utility indicates that travel costs are, on average, higher than can be explained by observed factors entering the baseline estimates in Column (1).

It is important to note once again that parameter estimates may not be directly comparable across Columns (1) and (2) of Table 3 due to the standard issue with discrete choice models discussed in Section 2, namely the non-identification of the scale of indirect utility.[11] However, the model parameters are not necessarily of independent interest in this context, but rather serve as key inputs into the calculation of the value of a lost user day due to the DWH spill. These welfare statistics are directly comparable across the two sets of estimates given that they do not depend on the scale of indirect utility, much in the same way that the willingness-to-pay statistic is comparable across the different simulations discussed in Sections 3 and 4. Given the lack of data on site visitation decisions during the period of the DWH spill, the process of calculating the lost user day value in English et al. (2018) involves a somewhat involved calibration procedure, which I discuss in Appendix C.

Table 3 reports estimated lost user day values in dollars per user day based on the baseline and control function parameter estimates and the procedure outlined in the Appendix. Using the baseline parameter estimates, I calculate a lost user day value of \$7.04 from the DWH oil spill. Based on the parameter estimates that account for possible endogeneity and measurement error in travel costs, I estimate a lost user day value of \$5.49, a 22% change in the per unit welfare loss resulting from the oil spill. Applying this proportional change to the aggregate welfare loss estimate of English et al. (2018), this translate into an overestimation of the total recreation-based losses from the DWH oil spill of around \$145 million, which suggests that accounting for the threats to travel cost estimation discussed herein is of first-order policy significance.

Unfortunately, it is difficult to determine the precise source of the bias evident in the standard model estimates solely based on the results in Table 3. Indeed, to determine whether any non-random sorting towards or away from desirable recreation sites exists in the data, I

[11]See Section 2 and Train (2009) for additional discussion.

Table 3. Second Stage Estimates

	Parameter	Baseline (1)	Control Function (2)
Travel Cost	α	-1.080 (0.008)	-1.512 (0.032)
Nesting Parameter	ρ	0.208 (0.002)	0.198 (0.003)
First Stage Residual	λ		1.103 (0.046)
Alternative Specific Constants		Yes	Yes
Lost User Day Value ($/day)		7.035	5.487
N		41,716	
Sites		83	

Notes: This table reports estimates from the second stage discrete choice model of demand for visits to Gulf Coast shoreline recreation sites. Parameters are estimated via maximum likelihood estimation. Column (1) employs a standard nested logit model of demand of a similar specification to that in English et al. (2018). Column (2) implements the analogous two-stage control function estimator described in Sections 4 and 5.2. Asymptotic standard errors are reported in parenthesis, with the standard errors in Column (2) adjusting for noise from the first stage regression using the asymptotic formula for the covariance matrix from Karaca-Mandic and Train (2003). The process for calculating lost user day values is described in detail in Appendix C.

would need to specify and estimate a complete model of the residential sorting process—an important exercise, but ultimately one which is outside the scope of this paper. However, the overall direction of the bias is nonetheless informative. In the results for Simulations 3 through 6 reported in Figure 4, we see that the baseline, uncorrected parameter estimates lead to an overestimation of the magnitude of a similar ratio of parameters as the lost user day statistic that I report in Table 3. The form of non-random sorting into travel costs modeled in Simulation 3, where households move away from sites they prefer leading to a positive correlation between idiosyncratic site preferences and travel cost is possible, though unlikely in practice: why, for instance, would households with a lower preference for a specific site sort closer to that site, all else equal?

Thus, it is likely that the bias that I observed in Table 3 is driven by measurement error in travel costs.[12] As I discuss in Section 3.2, there is good reason to believe that there is likely some form of measurement error in travel costs in most if not all applications of recreation demand modeling, even when extreme care is taken in the construction of this variable as is done by English et al. (2018). The necessity even in the most complex of travel cost calculations of simplifying assumptions and the use of aggregate data means that this field is likely observed with error. Indeed, even when the analyst gets travel costs correct

[12] As Simulation 6 demonstrates, it is possible that there is also sorting towards desirable recreation sites in the data that at least partially offsets the direction of the bias from measurement error. Without a more complete model of the residential sorting process, it is impossible to disentangle this from possible measurement error.

on average as was the case in Simulation 4, this can lead to non-trivial attenuation bias in standard model estimates. Thus, while English et al. (2018)'s example of estimating recreation-based welfare losses from the DWH oil spill represents the state-of-the-literature in recreation demand estimation, it also highlights the importance of addressing what are likely important issues with travel cost models in practice.

6 Conclusion

Recreation demand models inform decision-making across a wide range of applications, from regulatory impact analysis and resource management to public health and environmental litigation. Careful estimation of model parameters in these settings is critical to ensure unbiased inferences when making policy, regulatory, and legal decisions.

I show that two common empirical challenges previously ignored in the recreation demand literature, namely the potential for non-random residential sorting based on preferences for outdoor recreation and measurement error in travel costs, can substantially bias estimates in entire classes of commonly used discrete choice models. I demonstrate a simple, feasible approach to addressing these two issues simultaneously in empirical applications. In particular, I present an instrumental variables estimator that is analogous to two-stage least squares in the nonlinear context of standard discrete choice models. In a series of numerical simulations, I find that this relatively straightforward correction substantially outperforms standard approaches to recreation demand estimation. Moreover, I demonstrate the relative ease with which analysts can implement this fix by replicating a recent, high-profile application of recreation demand modeling that estimates the welfare losses from the 2010 Deepwater Horizon oil spill in the Gulf of Mexico, finding that accounting for these twin problems alters welfare estimates by as much as 22% in this context.

Based on the findings of this paper, I strongly encourage analysts estimating empirical models of recreation demand to implement instrumental variables estimators of their underlying discrete choice models using the two-stage control function approach. At best, doing so can demonstrate the robustness of estimates from the standard approaches to recreation demand estimation if not document meaningful bias from these estimators. In cases where there is reason to believe non-random sorting on preferences for recreation is important, analysts may want to consider explicitly modeling residence choices and recreation demand; however, the two-stage control function approach discussed herein provides a simple means of calculating unbiased estimates of willingness-to-pay and partial equilibrium welfare.

It is likely that the instrumental variables that I use in the empirical exercise in Section 5 can be applied in other applications of the recreation demand model. More broadly, instruments that influence households' choice of residence location or the marginal cost of site visitation, but not recreation site choices directly, should serve as valid instruments in practice. Those that I use to estimated a nationwide model of Gulf Coast shoreline recreation—households' average commuting time to work; variation in destination state gas taxes; and supply and demand shocks in the global crude oil markets—likely provide an accessible starting point for other applications of recreation demand modeling where more context-specific instruments may be readily available.

Estimating models of recreation demand requires solving important identification challenges. Indeed, this is well acknowledged in countless other applications of demand estimation since the seminal work of Wright (1928), with recreation demand estimation a puzzling outlier. While these findings might be concerning to policymakers and practitioners who rely on the conclusions from recreation demand models, the relatively simple fix for which I advocate in this paper should restore faith in this important methodology moving forward.

References

Abidoye, Babatunde O., and Joseph A. Herriges. 2012. "Model Uncertainty in Characterizing Recreation Demand." *Environmental and Resource Economics*, 53(2): 251–277.

Abidoye, Babatunde O., Joseph A. Herriges, and Justin L. Tobias. 2012. "Controlling for Observed and Unobserved Site Characteristics in RUM Models of Recreation Demand." *American Journal of Agricultural Economics*, 94(5): 1070–1093.

Albouy, David, Walter Graf, Ryan Kellogg, and Hendrik Wolff. 2016. "Climate Amenities, Climate Change, and American Quality of Life." *Journal of the Association of Environmental and Resource Economists*, 3(1): 205–246.

Angrist, Joshua D, and Alan B Krueger. 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *Journal of Economic Perspectives*, 15(4): 69–85.

Bajari, Patrick, Jane Cooley Fruehwirth, Kyoo il Kim, and Christopher Timmins. 2012. "A Rational Expectations Approach to Hedonic Price Regressions with Time-Varying Unobserved Product Attributes: The Price of Pollution." *The American Economic Review*, 102(5): 1898–1926.

Bayer, Patrick, and Christopher Timmins. 2007. "Estimating Equilibrium Models Of Sorting Across Locations*." *The Economic Journal*, 117(518): 353–374.

Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. 2009. "Migration and hedonic valuation: The case of air quality." *Journal of Environmental Economics and Management*, 58(1): 1–14.

Bento, Antonio, Matthew Freedman, and Corey Lang. 2014. "Who Benefits from Environmental Regulation? Evidence from the Clean Air Act Amendments." *The Review of Economics and Statistics*, 97(3): 610–622.

Bishop, Kelly C., Nicolai V. Kuminoff, H. Spencer Banzhaf, Kevin J. Boyle, Kathrine von Gravenitz, Jaren C. Pope, V. Kerry Smith, and Christopher D. Timmins. 2020. "Best Practices for Using Hedonic Property Value Models to Measure Willingness to Pay for Environmental Quality." *Review of Environmental Economics and Policy*, 14(2): 260–281.

Bureau of Economic Statistics. 2019. "Outdoor Recreation Satellite Account, U.S. and States, 2019." Bureau of Economic Statistics.

Egan, Kevin J., Joseph A. Herriges, Catherine L. Kling, and John A. Downing. 2009. "Valuing Water Quality as a Function of Water Quality Measures." *American Journal of Agricultural Economics*, 91(1): 106–123.

English, Eric, Roger H. von Haefen, Joseph Herriges, Christopher Leggett, Frank Lupi, Kenneth McConnell, Michael Welsh, Adam Domanski, and Norman

Meade. 2018. "Estimating the value of lost recreation days from the Deepwater Horizon oil spill." *Journal of Environmental Economics and Management*, 91: 26–45.

Greenstone, Michael, and Justin Gallagher. 2008. "Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program." *The Quarterly Journal of Economics*, 123(3): 951–1003.

Hallstrom, Daniel G., and V. Kerry Smith. 2005. "Market responses to hurricanes." *Journal of Environmental Economics and Management*, 50(3): 541–561.

Hausman, Jerry. 2001. "Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left." *Journal of Economic Perspectives*, 15(4): 57–67.

Heckman, James J. 1978. "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica*, 46(4): 931–959.

Hicks, Robert L., and Ivar E. Strand. 2000. "The Extent of Information: Its Relevance for Random Utility Models." *Land Economics*, 76(3): 374–385.

Hjerpe, Evan, Anwar Hussain, and Thomas Holmes. 2020. "Amenity Migration and Public Lands: Rise of the Protected Areas." *Environmental Management*, 66(1): 56–71.

Hotelling, Harold. 1947. "Letter to the National Parks Service." In *An economic study of the monetary evaluation of recreation in the National Parks.* US Department of the Interior, National Parks Service.

Kao, Chihwa, and John F. Schnell. 1987. "Errors in variables in the multinomial response model." *Economics Letters*, 25(3): 249–254.

Karaca-Mandic, Pinar, and Kenneth Train. 2003. "Standard error correction in two-stage estimation with nested samples." *The Econometrics Journal*, 6(2): 401–407.

Keiser, David A, and Joseph S Shapiro. 2019. "Consequences of the Clean Water Act and the Demand for Water Quality." *The Quarterly Journal of Economics*, 134(1): 349–396.

Kilian, Lutz. 2009. "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market." *American Economic Review*, 99(3): 1053–1069.

Klaiber, Allen H., and Daniel J. Phaneuf. 2010. "Valuing open space in a residential sorting model of the Twin Cities." *Journal of Environmental Economics and Management*, 60(2): 57–77.

Kling, Catherine L., and Cynthia J. Thomson. 1996. "The Implications of Model Specification for Welfare Estimation in Nested Logit Models." *American Journal of Agricultural Economics*, 78(1): 103–114.

Kuwayama, Yusuke, Sheila Olmstead, and Jiameng Zheng. 2022. "A more comprehensive estimate of the value of water quality." *Journal of Public Economics*, 207: 104600.

Lupi, Frank, Daniel J. Phaneuf, and Roger H. von Haefen. 2020. "Best Practices for Implementing Recreation Demand Models." *Review of Environmental Economics and Policy*, 14(2): 302–323.

McFadden, Daniel. 1974. "Conditional logit analysis of qualitative choice behavior." In *Frontiers in econometrics*.

Muehlenbachs, Lucija, Elisheba Spiller, and Christopher Timmins. 2015. "The Housing Market Impacts of Shale Gas Development." *The American Economic Review*, 105(12): 3633–3659. Publisher: American Economic Association.

Murdock, Jennifer. 2006. "Handling unobserved site characteristics in random utility models of recreation demand." *Journal of Environmental Economics and Management*, 51(1): 1–25.

National Park Service. 2020. "Visitation Numbers (U.S. National Park Service)." National Park Service.

Parsons, George R. 1991. "A Note on Choice of Residential Location in Travel Cost Demand Models." *Land Economics*, 67(3): 360–364.

Parsons, George R., Andrew J. Plantinga, and Kevin J. Boyle. 2000. "Narrow Choice Sets in a Random Utility Model of Recreation Demand." *Land Economics*, 76(1): 86–99.

Parsons, G.R., D. Matthew Massey, and Ted Tomasi. 1999. "Familiar and Favorite Sites in a Random Utility Model of Beach Recreation." *Marine Resource Economics*, 14(4): 299–315.

Petrin, Amil, and Kenneth Train. 2010. "A Control Function Approach to Endogeneity in Consumer Choice Models." *Journal of Marketing Research*, 47(1): 3–13.

Phaneuf, Daniel J., and V. Kerry Smith. 2005. "Recreation Demand Models." In *Handbook of Environmental Economics*. Vol. 2, 671–761. Elsevier.

Phaneuf, Daniel J., V. Kerry Smith, Raymond B. Palmquist, and Jaren C. Pope. 2008. "Integrating Property Value and Local Recreation Models to Value Ecosystem Services in Urban Watersheds." *Land Economics*, 84(3): 361–381.

Rickman, Dan S., and Shane D. Rickman. 2011. "Population Growth in High-Amenity Nonmetropolitan Areas: What's the Prognosis?*." *Journal of Regional Science*, 51(5): 863–879.

Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*, 82(1): 34–55.

Shaw, W. Douglass, and Michael T. Ozog. 1999. "Modeling Overnight Recreation Trip Choice: Application of a Repeated Nested Multinomial Logit Model." *Environmental and Resource Economics*, 13(4): 397–414.

Smith, V. Kerry, William H. Desvousges, and Ann Fisher. 1986. "A Comparison of Direct and Indirect Methods for Estimating Environmental Benefits." *American Journal of Agricultural Economics*, 68(2): 280–290.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy*, 64(5): 416–424.

Timmins, Christopher, and Jennifer Murdock. 2007. "A revealed preference approach to the measurement of congestion in travel cost models." *Journal of Environmental Economics and Management*, 53(2): 230–249.

Tourangeau, Roger, Eric English, Kenneth E. McConnell, David Chapman, Ismael Flores Cervantes, Eric Horsch, Norman Meade, Adam Domanski, and Michael Welsh. 2017. "The Gulf Recreation Study: Assessing Lost Recreational Trips from the 2010 Gulf Oil Spill." *Journal of Survey Statistics and Methodology*, 5(3): 281–309.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation.* . Second ed., New York, NY, US:Cambridge University Press.

Villas-Boas, J. Miguel, and Russell S. Winer. 1999. "Endogeneity in Brand Choice Models." *Management Science*, 45(10): 1324–1338.

von Haefen, Roger H., and Daniel J. Phaneuf. 2008. "Identifying demand parameters in the presence of unobservables: A combined revealed and stated preference approach." *Journal of Environmental Economics and Management*, 56(1): 19–32.

Ward, Frank A., and John B. Loomis. 1986. "The Travel Cost Demand Model as an Environmental Policy Assessment Tool: A Review of Literature." *Western Journal of Agricultural Economics*, 11(2): 164–178.

Wright, Philip Green. 1928. *The Tariff on Animal and Vegetable Oils.* Macmillan.

Online Appendix for "Hotelling Meets Wright: Spatial Sorting and Measurement Error in Recreation Demand Models"

Jacob T. Bradt

The following appendices are **for online publication only**:

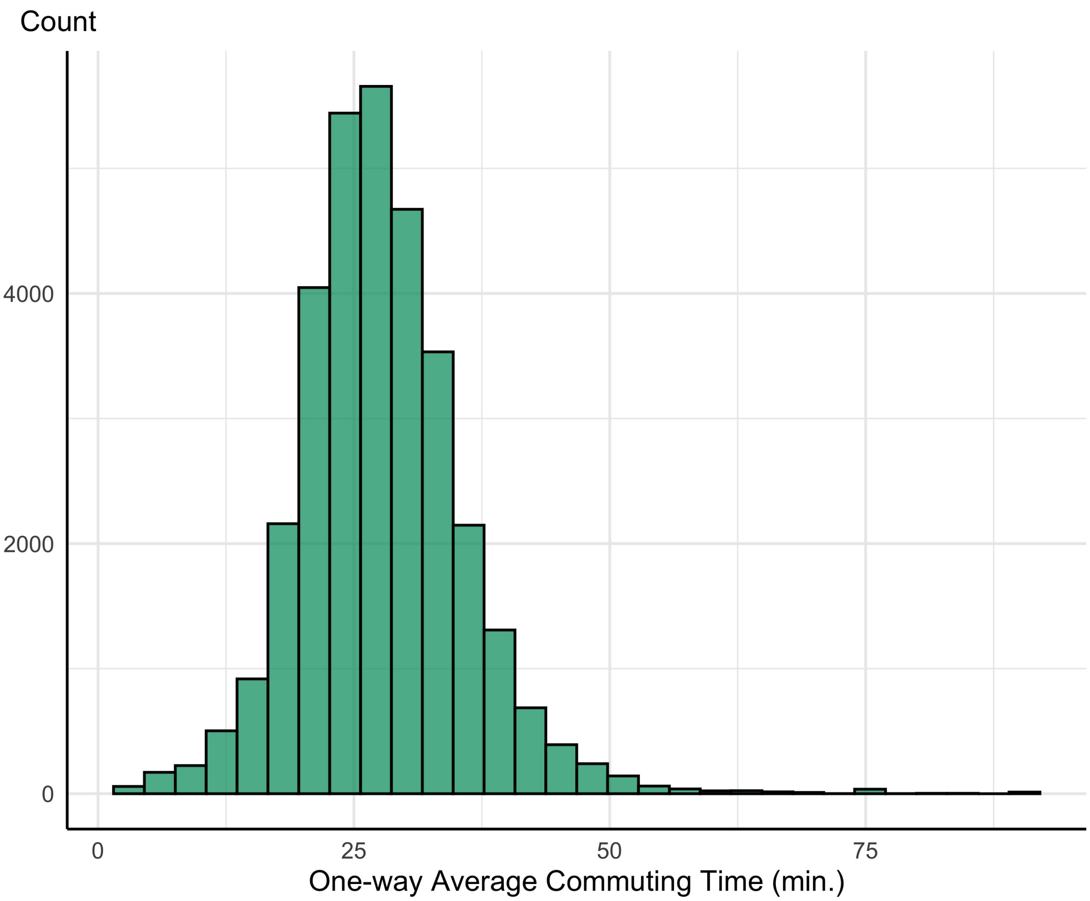
—Appendix Section A: Supplemental Figures and Tables

—Appendix Section B: Estimating Structural Oil Market Shocks

—Appendix Section C: Calculating Lost User Day Value

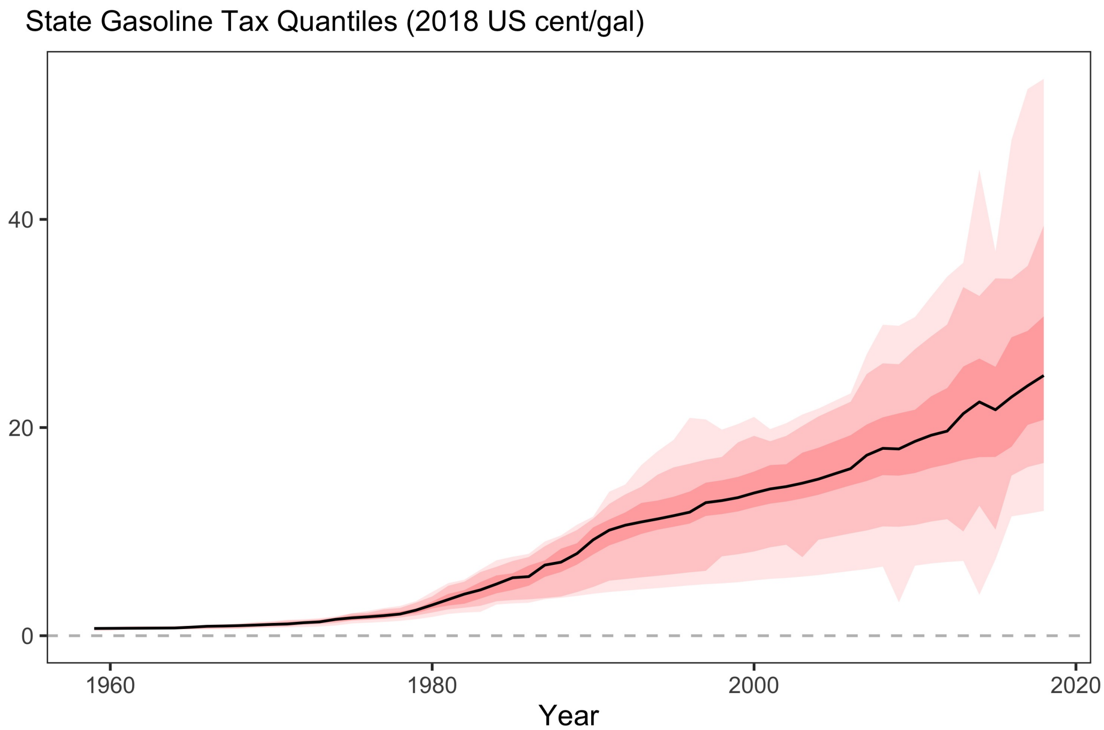
A Supplemental Figures and Tables

Figure A1. Average Commuting Time by Zip Code Tabulation Area, 2012



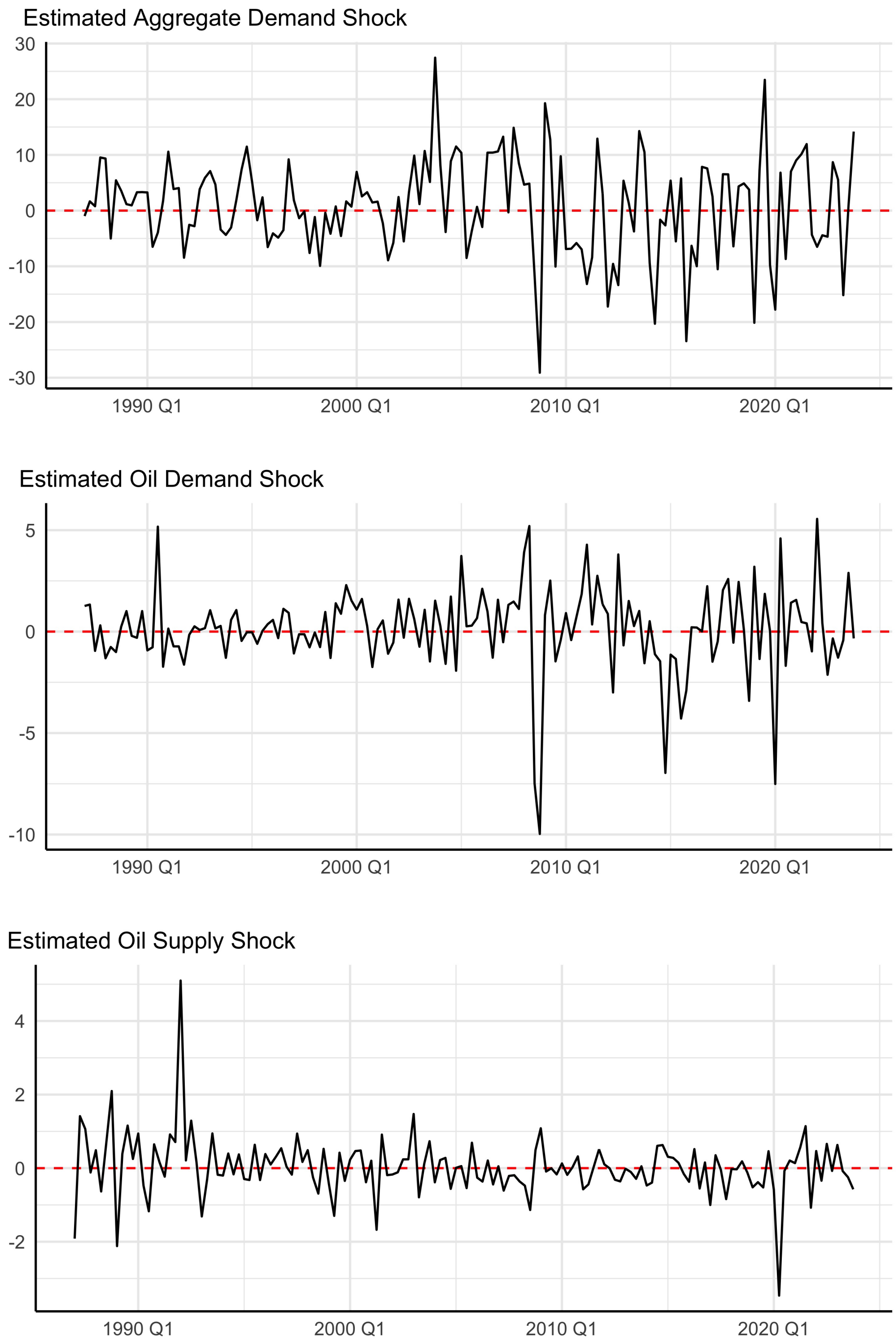
Notes: This figure plots the distribution of average one-way commuting time by Zip Code Tabulation Area for 2012. Data are taken from the US Census Bureau's 2012 5-year American Community Survey.

Figure A2. Variation in Real State Gasoline Taxes over Time



Notes: This figure plots the distribution of gasoline tax rates across US states over time. In particular, the shaded areas show the following annual percentiles of gasoline tax rates: bottom, 20th, 40th, 60th, 80th, and top. The black line shows the median state gasoline tax rate. All values are in 2018 cents per gallon.

Figure A3. Time Series Variation in Structural Shocks from Model of Global Crude Oil Market



Notes: This figure shows estimated structural errors from a vector autoregressive model of the global crude oil market based on Kilian (2009). The model uses global time series data on global crude oil production, real economic output, and the real price of oil to calculate structural shocks to aggregate demand (top), oil demand (middle), and oil supply (bottom).

B Estimating Structural Oil Market Shocks

It is possible that seasonality in crude oil prices or macroeconomic trends that influence the global crude oil market also affect household recreation. As a result, I follow the approach of Kilian (2009) to isolate structural supply and demand shocks in the global crude oil market from factors which may otherwise be correlated with recreation demand. This approach uses a novel measure of global real economic activity as well as global crude oil production to decompose the real price of crude oil into three components: (1) crude oil supply shocks; (2) demand shocks for industrial commodities, a proxy for aggregate demand; and (3) demand shocks for crude oil.

To isolate structural shocks in the global crude oil market following Kilian (2009), I acquire data on monthly global crude oil production and US refiner acquisition costs, which proxies crude oil prices, from the US Energy Information Administration (EIA). I obtain a monthly index of global real economic activity from Kilian (2009), which is available through the Federal Reserve Bank of Dallas.[1] This monthly index, which proxies for global business cycle trends, is derived from a panel of dollar-denominated global bulk dry cargo shipping rates. This index can be viewed as a proxy for the volume of shipping in global industrial commodity markets. Given the importance of freight in international trade, this index provides a strong indicator of global demand pressures and is more closely linked with global real output than other measures such as GDP (Kilian, 2009). I combine data on these monthly time series for the period from January 1985 to October 2023.

Following Kilian (2009), I isolate structural shocks in the global crude oil market from these data using a vector autoregressive (VAR) model. Let $x_t' = \begin{bmatrix} \Delta prod_t & rea_t & rpo_t \end{bmatrix}$ where $\Delta prod_t$ is the percent change in global crude oil production, rea_t is the index of real economic activity from Kilian (2009), and rpo_t is the real price of oil. The structural VAR is as follows:

$$A_0 x_t = \alpha + \sum_{i=1}^{24} A_i z_{t-i} + \varepsilon_t \tag{B1}$$

where ε_t is a vector of serially and mutually uncorrelated structural shocks. Following Kilian (2009), I assume that A_0^{-1} has a recursive structure such that the reduced form errors in

[1] The global real economic activity index is available for download here: https://www.dallasfed.org/research/igrea (last accessed 3/4/2024).

(B1), e_t, can be decomposed as follows:

$$e_t = \begin{bmatrix} e_t^{\Delta prod} \\ e_t^{rea} \\ e_t^{rpo} \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \varepsilon^{\text{oil supply shock}} \\ \varepsilon^{\text{agg demand shock}} \\ \varepsilon^{\text{oil demand shock}} \end{bmatrix} \tag{B2}$$

where $\varepsilon^{\text{oil supply shock}}$, $\varepsilon^{\text{agg demand shock}}$, and $\varepsilon^{\text{oil demand shock}}$ are the structural errors of interest. Thus, with estimates of the autoregressive parameters, A_0, and reduced form errors, e_t, from the empirical implementation of (B1), it is possible to construct estimates of the structural error terms.

This model implicitly assumes several exclusion restrictions. In particular, the model assumes that oil supply does not respond to innovations in oil demand within the same month, i.e., a vertical short run supply curve. Moreover, the model structure assumes that changes in the real price of oil driven by oil-specific shocks will not lower global real economic activity immediately. Any changes in the real price of oil that cannot be explained by unpredictable innovations to global oil production or real economic activity will, by construction, reflect changes in the demand for oil rather than changes to the demand for all industrial commodities.

Figure A3 plots quarterly averages of the estimated monthly innovations to aggregate demand, oil demand, and oil supply. As is clear from Figure A3, the real price of oil is a function of a number of concurrent shocks, each of which is driven by different global phenomena. Several events clearly emerge in Figure A3, including the Great Recession and the Covid-19 pandemic.

C Calculating Lost User Day Value

Given that the shoreline recreation demand model in Section 5.2 is estimated using data from the post-spill period, it recovers preferences for shoreline recreation in the Gulf of Mexico under the baseline or no-spill conditions. In order to calculate estimates of the lost user day value due to the Deepwater Horizon (DWH) oil spill, it is therefore necessary to estimates how demand shifted in response to the spill. Given the lack of site choice data during the spill period—which lasted approximately 5-6 months after the DWH explosion in April 2010—and the relatively extreme nature of the spill conditions, English et al. (2018) uses external information from the onsite counts to infer changes to overall preferences for individual sites induced by the spill.

This external information is based on the analysis of Tourangeau et al. (2017) and is used to infer changes in affected sites' alternative specific constants. In particular, English et al. (2018) use estimates on the proportional reduction in trips to two broad categories of sites—the Northern Gulf and the Florida Peninsula (see Figure 5 in the main text)—to calibrate affected sites alternative specific constants to reflect spill conditions. Letting the estimates of the proportional reduction in visitation be denoted by r_g for $g \in \{$Northern Gulf, Florida Peninsula$\}$, the calibration exercise entails selecting group-level adjustments, δ_g, to the alternative specific constants such that

$$\xi_j^1 = \begin{cases} \xi_j^0 + \delta_{NG} & \text{for } j \in \mathcal{J}_{\text{Northern Gulf}} \\ \xi_j^0 + \delta_{FP} & \text{for } j \in \mathcal{J}_{\text{Florida Peninsula}} \\ \xi_j^0 & \text{otherwise} \end{cases} \tag{C1}$$

where $\mathcal{J}_{\text{Northern Gulf}}$ and $\in \mathcal{J}_{\text{Florida Peninsula}}$ are the sets of sites that fall within the Northern Gulf and Florida Peninsula, respectively; ξ_j^0 are the alternative specific constants estimated under baseline conditions; and ξ_j^1 are the calibrated alternative specific constants under spill conditions.[2]

[2] English et al. (2018) model two distinct spill condition periods, one immediately after the spill in which both affected regions experience a fixed reduction in visits and a later spill condition period where only the Northern Gulf experiences adverse impacts from the spill, with a lower reduction in observed visitation for this region during this later period. For simplicity and for the sake of comparing estimates across the standard and control function estimators, I focus on estimating lost user day values for the first period only since the calculation is analogous during the second period.

To calibrate δ_g, I first calculate spill condition choice probabilities as follows:

$$s_j^1 = \begin{cases} (1 - r_{NG})s_j^0 & \text{for } j \in \mathcal{J}_{\text{Northern Gulf}} \\ (1 - r_{FP})s_j^0 & \text{for } j \in \mathcal{J}_{\text{Florida Peninsula}} \\ s_j^0 & \text{otherwise} \end{cases} \tag{C2}$$

I then iterate over the following contraction mapping until convergence, where for each iteration t, the next iterate is given by:

$$\delta_g^{(t+1)} = \delta_g^{(t)} + \log(s_j^1) - \log(\hat{s}_j(\xi_j^1)) \tag{C3}$$

where—with some abuse of notation—$\hat{s}_j(\xi_j^1)$ is the model-implied market shares based on the choice probability defined in Section 5.2, total households facing each choice occasion, the parameters estimated under baseline conditions; and the alternative specific constants under the spill conditions.

With estimates of the shoreline recreation demand model under baseline conditions and the calibrated alternative specific constants under spill conditions, it is possible to calculate welfare losses following English et al. (2018). In particular, the value per lost trip from the spill conditions is given by:

$$\Delta CV = \frac{\sum_{i=1}^{N} T_i \frac{1}{\hat{\alpha}} \left[\log\left(1 + \left(\sum_{j=1}^{J} \exp\left(\frac{\hat{v}_{ik}^1}{\hat{\rho}}\right)\right)^{\hat{\rho}}\right) - \log\left(1 + \left(\sum_{j=1}^{J} \exp\left(\frac{\hat{v}_{ik}^0}{\hat{\rho}}\right)\right)^{\hat{\rho}}\right)\right]}{\sum_{t=1}^{N} T_i \Big((\hat{p}_{i,NG}^0 + \hat{p}_{i,FP}^0) - (\hat{p}_{i,NG}^1 + \hat{p}_{i,FP}^1) \Big)} \tag{C4}$$

where T_i is the number of choice occasions that observation i represents (i.e., an observation-level weight); $\hat{v}_{ik}^1$ is the calibrated estimate of conditional utility under spill conditions; $\hat{v}_{ik}^0$ is the estimate of conditional utility under baseline conditions; $\hat{\alpha}$ and $\hat{\rho}$ are parameter estimates; $\hat{p}_{i,NG}^0$ and $\hat{p}_{i,FP}^0$ are estimates of the probability of visiting a North Gulf or a Florida Peninsula site under baseline conditions, respectively; and $\hat{p}_{i,NG}^1$ and $\hat{p}_{i,FP}^1$ are calibrated estimates of the probability of visiting a North Gulf or a Florida Peninsula site under spill conditions, respectively. The numerator is based on the standard log-sum formula for the welfare loss due to changes in conditional utility and the denominator gives the change in trips to the North Gulf and Florida Peninsula. To translate (C4) into a value per user day, I divide by the mean number of recreational days per trip.