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

Jacob T. Bradt †

October 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 decreases estimates of the welfare costs of the 2010 Deepwater Horizon oil spill by 12 percent.

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

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

*I thank Daniel Phaneuf and two anonymous referees for their invaluable support improving this paper. I also thank Joe Aldy, Myrto Kalouptsidi, Cathy Kling, Robin Lee, Kostas Metaxoglou, Ariel Pakes, Frank Pinter, Devesh Raval, Jim Stock, Mo Van de Sompel, 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 excellent 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. Blake Barr provided excellent research assistance constructing measures of state motor fuels taxes. I gratefully acknowledge funding from the Vicki Norberg Bohm Fellowship. I am responsible for all remaining errors.

†The University of Texas at Austin, McCombs School of Business; jacob.bradt@austin.utexas.edu.

1 Introduction

Outdoor recreation is an industry of both intrinsic and instrumental value. Private spending on recreational activities represents 2.2% of United States GDP annually (Bureau of Economic Statistics, 2023) 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 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. This argument motivates a large literature in the estimation of non-market, environmental amenities, which builds off of Hotelling (1947)'s simple insight that consumption of an outdoor recreation site's amenities requires individuals to incur the cost of a trip to that site. Since recreation sites are just bundles of attributes—including, for example, the types of outdoor activities supported, environmental qualities, and physical amenities, among others—researchers can use observed travel costs to infer how tradeoff differences in environmental amenities.

This paper examines how two common challenges in demand estimation—omitted variable bias and measurement error—affect recreation demand models. Individual-level travel costs, often seen as a strength in this context, can worsen these issues due to non-random residential sorting and difficulties in accurately observing these costs. Numerical simulations reveal that conventional recreation demand estimates are significantly biased. To address this, I propose a simple instrumental variables approach that consistently outperforms traditional models. In an empirical example replicating English et al. (2018), this method reduces estimates of the welfare costs of the 2010 Deepwater Horizon oil spill by 12%.

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.

Despite this, researchers often assume that individual-specific travel costs are exogenous in recreation demand models; however, travel costs—implicitly a function of residence choice—may correlate with unobserved characteristics. For example, if anglers choose homes near desirable fishing sites, estimates of price responsiveness in recreation demand will be biased, as they are effectively paying less for a higher-value good. Empirical evidence shows that individuals sort based on natural amenities, with long-term migration trends toward environmentally-rich rural areas (Hjerpe et al., 2020; Rickman and Rickman, 2011). Studies also document sorting based on preferences for spatial characteristics, such as climate and environmental quality, echoing Tiebout (1956)-like sorting behavior (Albouy et al., 2016; Bayer and Timmins, 2007; Bayer et al., 2009; Klaiber and Phaneuf, 2010). The hedonic property framework of Rosen (1974)—a workhorse methodology in environmental economics—also implicitly assumes preferences for environmental qualities affect residence location choices.

Researchers have 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 particularly acute: researchers and analysts rarely have access to the true cost of travel associated with realized trips, let alone the full set of travel costs in an individuals' choice set. Analysts must therefore use information on residence location and the location of sites 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 (Lupi et al., 2020) this approach easily leads to measurement error in the price of recreation activities, which produces biased results (Angrist and Krueger, 2001; Hausman, 2001).

Few applications of recreation demand modeling take either omitted variables-induced endogeneity or measurement error in prices seriously, despite good reason to believe that both concerns are non-trivial in this context. 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 inaccurate 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 and find that this relatively straightforward correction substantially outperforms standard approaches to recreation demand estimation.

To demonstrate the relevance of these challenges, I replicate English et al. (2018)'s nationwide model of Gulf Coast recreation demand, which played a key role in compensatory litigation in the aftermath of the 2010 Deepwater Horizon oil spill. I estimate two versions of the model: one ignoring travel cost endogeneity and measurement error, and one addressing them using a two-stage control function approach. Accounting for these issues alters welfare estimates by up to 12%, a significant difference in this policy-relevant context.

These findings relate to a large literature on non-market valuation of environmental amenities. Empirical recreation demand models build on McFadden (1974)'s random utility maximization framework to evaluate a number of different environmental attributes, including 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). A similarly expansive literature leverages the hedonic property framework of Rosen (1974) to estimate the capitalization of non-market environmental amenities in housing 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).

This paper documents a logical inconsistency between these two vast literatures on non-market valuation in environmental economics. On the one hand, recreation demand estimation assumes that residential sorting is independent of the environmental amenities they study; however, the hedonic framework finds that households make residence location choices at least in part based on these factors. Existing work similarly seeks to bridge the gap between these two methods: Phaneuf et al. (2008) point out that conventional hedonic property studies estimating willingness-to-pay for non-market environmental amenities do not fully incorporate recreational use values and propose a two-stage approach to add these values to hedonic estimates. Kuwayama et al. (2022) apply the approach of Phaneuf et al. (2008) to estimate willingness-to-pay for water quality improvements in Tampa Bay, FL and find meaningful differences from standard methods. 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.

Existing research acknowledges these empirical challenges in recreation demand models. Parsons (1991) highlights price endogeneity due to preference-based sorting but offers no solution for structural discrete choice models. Murdock (2006) and Abidoye et al. (2012) develop methods to control for unobserved site-specific attributes with alternative-specific constants, though this does not address individual-level travel cost endogeneity. Timmins and Murdock (2007) use instrumental variables for site congestion but not for price endogeneity. von Haefen and Phaneuf (2008) combine stated and revealed preference data to account for site quality in the presence of unobservable characteristics. Keiser (2019) documents bias from measurement error in environmental attributes. This paper contributes to earlier work by focusing specifically on endogeneity and measurement error in travel costs, offering a simple approach to mitigate these issues.

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 conventional recreation demand models 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 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 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 a vector of observable individual- and alternative-specific attributes and a residual term 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:

$$u_{ijt} = \underbrace{x'_{jt}\beta - c_{ijt}\alpha + \xi_j}_{\equiv v_{ijt}} + \varepsilon_{ijt} \tag{1}$$

where x'_{jt} is a vector of observable site- and choice occasion-specific attributes; c_{ijt} is an idiosyncratic measure of travel cost; ξ_j is an alternative-specific constant that captures average valuations of time-invariant, site-specific attributes; v_{ijt} collects the observable components of utility; 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] For expositional clarity and ease of notation, the target parameters $\theta' = \begin{bmatrix} \beta' & \alpha \end{bmatrix}$ in (1) do not vary across either observable or unobservable attributes in the population.[2]

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_{ijt} is distributed Type 1 Extreme Value (T1EV) across the population, and is independent and identically distributed (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.

Taking this common assumption of an extreme-value error term, $\varepsilon_{ijt} \sim$ T1EV, it is

[1]It is possible to include the alternative-specific constant, ξ_j, and still obtain estimates of consumers' valuation of invariant, site-specific observable by projecting the observable factors of interest on estimates of these site fixed effects (Murdock, 2006).

[2]It is not uncommon in the literature to allow the coefficients on the additive, observable components of utility, θ, to vary across individuals in the population based on observable attributes, unobservable attributes, or both. Several classes of discrete choice models permit such heterogeneity, including the mixed or random parameters logit. Though I do not allow for this heterogeneity in the exposition and results of this paper, it is important to note that the biases and corresponding control function approach outlined herein extend to these richer classes of models as shown by Petrin and Train (2010) and others.

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:

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

where $\mathcal{C} = \{1, ..., J\}$ is the choice set. Estimation proceeds via maximum likelihood, where the log likelihood is defined as:

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

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 likelihood estimate is the value of θ that maximizes (3).

Having recovered estimates of the model parameters, θ, 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.[3] 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}:

$$WTP^q(\theta) = \frac{\beta^q}{\alpha}$$

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

$$\Delta \mathbb{E}[CS_{it}(\theta)] = \frac{1}{\alpha} \left[\sum_j \exp\left(v^1_{ijt}(\theta) \right) - \sum_j \exp\left(v^0_{ijt}(\theta) \right) \right]$$

where v^1_{ijt} and v^0_{ijt} are the observable components of individual i's utility on occasion t calculated for q^1_{jt} and q^0_{jt}, respectively. It is possible to construct empirical estimates of these statistics from observable data and parameter estimates from (3).

[3] 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.

3 Challenges with Travel Cost in The Standard Model

In this section, I present two challenges with the implementation of the standard model outlined in Section 2: non-random 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

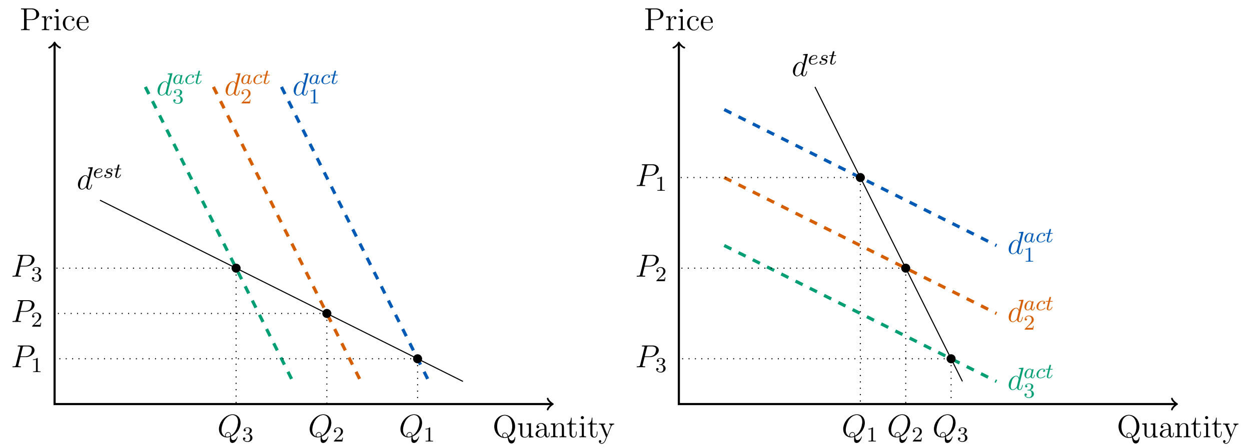
If households factor preferences for certain recreation or environmental amenities into residence location decisions, this can result in non-random, non-zero correlation between observed travel costs and preferences for a specific site. For example, households with a particularly strong idiosyncratic 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.

Households with a particularly strong idiosyncratic preference for a specific site and its attributes may also systematically choose to live further from that site. 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 greater propensity to reside in urban centers far from high value recreation 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 this in a stylized model 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). In both forms of sorting, consumers at (P_1, Q_1) have the highest unobserved preference for recreation whereas consumers at (P_3, Q_3) have the lowest. Ignoring non-random sorting, the analyst 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 have high idiosyncratic preferences, the analyst underestimates households' responsiveness to recre-

Figure 1. Biased Recreation Demand Estimates with Two Sorting Patterns

(a) Sorting towards Recreation Sites **(b)** Sorting away from Recreation Sites



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 analyst 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. Consumers with demand curve d_1^{act} have the highest idiosyncratic preference for recreation, whereas consumers with demand curve d_3^{act} have the lowest: consumers represented by d_1^{act} demand higher levels of recreation at all price levels compared with consumers represented by d_2^{act} and d_3^{act}

ation costs as shown in Figure 1b, mistaking inward shifts in household demand curves for movement along the recreation price gradient.

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:

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

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

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

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

$$\begin{aligned} u_{ij} &= 1.0x_{ij} - 2.0c_{ij} + 1.0x_j + \xi_{ij} + \tilde{\varepsilon}_{ij} \\ c_{ij} &= 5.0 + 1.0z_{ij} + \rho_{sim}\xi_{ij} + \mu_{ij} \\ \xi_{ij} &= 0.5\tilde{\xi}_{ij} + 0.5\tilde{\xi}_{g(i)} \end{aligned} \tag{6}$$

where

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

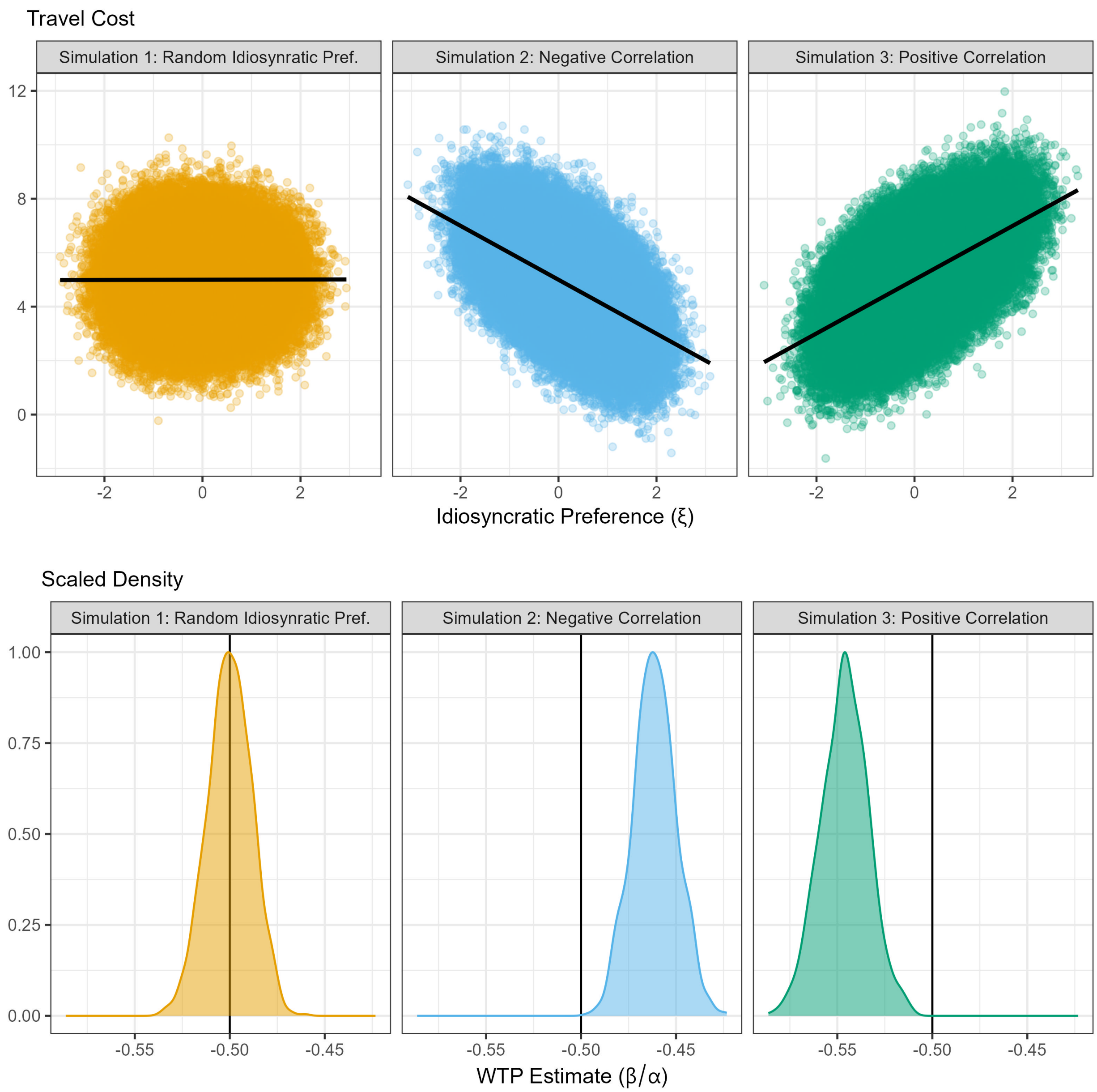
Note that the idiosyncratic unobservable, ξ_{ij}, is the sum of two independent unobservables, an idiosyncratic component, $\tilde{\xi}_{ij}$, and a component that is common across groups, $g(i)$, of individuals, $\tilde{\xi}_{g(i)}$. This allows for non-trivial correlation in unobservable preferences across groups of similar individuals, $g(i)$.[4] 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 10,000 individuals from 10 different origin groups, each of

[4] I explore the robustness of different estimators to different assumptions on the degree of within-group correlation in the unobservable preference, ξ_{ij}, in Section 4.4.

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 and 3.2

whom chooses between 20 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.

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. I estimate the parameters for each simulated dataset via a multinomial logit model with alternative-specific constants, 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.[5]

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. With no correlation between the mean-zero idiosyncratic preference (ξ_{ijt}) and travel cost (c_{ijt}) in Simulation 1, the standard 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 -0.46 in Simulation 2 and -0.55 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 3.40 and -3.71 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, I overestimate households' responsiveness to recreation costs—i.e., the parameter α on travel cost in indirect utility. As I can see from the simulation results in Figure 2, this 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, I underestimate households' responsiveness to recreation costs. As is clear from Figure 2, this results in 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.

[5] 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 cost of travel to a site is the price of consuming that site's amenities. The 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 have access to the true cost of travel associated with realized trips, let alone all trips in their choice set.

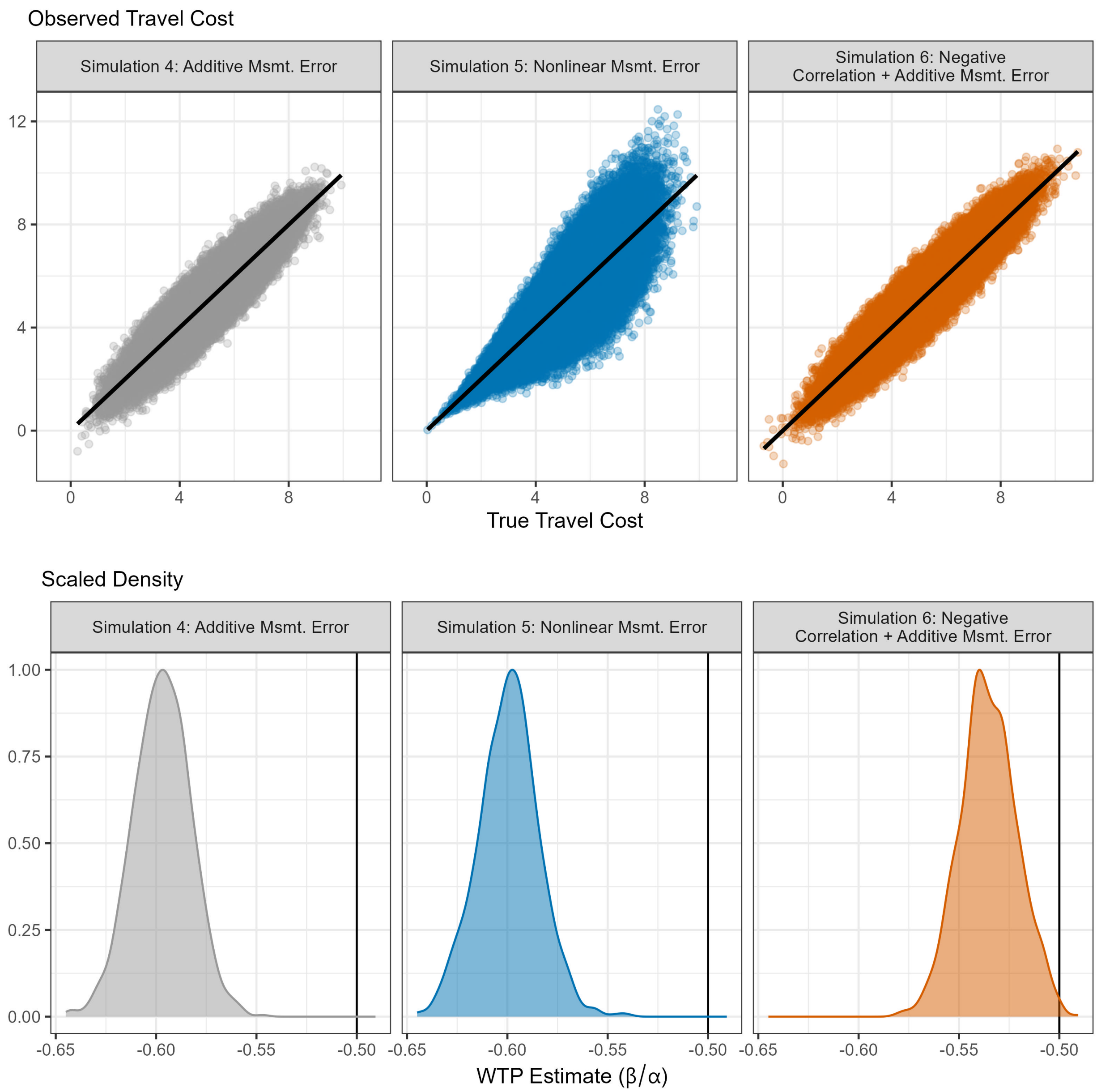
As a result, analysts must use information on residence location and the location of sites in households' choice sets in combination with a set of assumptions to construct estimates of travel costs. While there is a well-defined set of best practices for doing so—see Lupi et al. (2020)—this approach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

Since driving is the main transport mode in recreation demand, analysts must estimate driving distances between households and recreation sites, then convert them into monetary values using marginal fuel costs, tolls, and vehicle depreciation. A common method is to rely on average driving costs from sources like the American Automobile Association (Lupi et al., 2020). However, this likely introduces classical (mean-zero) measurement error due to cost heterogeneity. Similarly, it is often difficult for analysts to observe households' opportunity costs of travel time. As a result, analysts often estimate travel time costs using a fixed proportion (typically one-third to one-half) of average local wages, which, combined with assumptions about travel speed, can also lead to measurement error (Lupi et al., 2020)

Additional features of the travel cost construction process can lead to non-classical measurement error. A practical challenge 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 site and a household increases. By not accounting for potential substitution away from driving as distance increases, conventional applications of this method can generate measurement error that explicitly scales with distance, i.e., non-classical measurement error.

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

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.

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:

$$u_{ijt} = x'_{jt}\beta - \hat{c}_{ijt}\alpha + \xi_j + \tilde{\omega}_{ijt} \tag{8}$$

where $\tilde{\omega}_{ijt}$ is an unobserved, idiosyncratic, and independent shock to preferences and $\hat{c}_{ijt}$ is the estimated or observed travel cost for household i to visit site j at time t such that

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

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}. Plugging (9) into (8) shows a correspondence between the measurement error problem and the non-random travel cost issue outlined in Section 3.1:

$$u_{ijt} = v_{ijt}(\theta) + \underbrace{\zeta_{ijt} + \tilde{\omega}_{ijt}}_{\equiv \omega_{ijt}} \tag{10}$$

where I define $\zeta_{ijt} = -\alpha(g(c_{ijt}) + \eta_{ijt})$ as the unobserved component of household utility that is correlated with both household utility and observed travel costs, $\hat{c}_{ijt}$. The symmetry between the non-random sorting and measurement error issues is evident in (4) and (10). Estimating the standard model of Section 2 while treating observed travel costs, $\hat{c}_{ijt}$, as the true travel costs implicitly assumes that $\zeta_{ijt} + \tilde{\omega}_{ijt} = \omega_{ijt} \sim$ T1EV, thereby assuming that $\omega_{ijt} \perp \hat{c}_{ijt}$ which ignores the fact that $\zeta_{ijt} \not\perp \hat{c}_{ijt}$. Kao and Schnell (1987) derive the asymptotic properties of a multinomial logit model with measurement error and show that the standard parameter estimates 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:

<u>Simulation 4:</u> $\hat{c}_{ijt} = c_{ijt} + \eta_{ijt}$ (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 10,000 individuals from 10 different origin groups choosing between 20 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 generated 1000 choice datasets for each of the three simulations, I again make the standard assumption on the distribution of the unobservable residual term that ignores measurement error in travel costs. I estimate the parameters for each simulated dataset via a multinomial logit model with alternative-specific constants, thereby generating distributions of parameter estimates from the standard model of recreation demand for each data generating process. Importantly, however, I now treat the simulated observed travel costs, $\hat{c}_{ijt}$, as the true travel costs when estimating the parameters of the model. I then compare estimates of the marginal rate of substitution between x_{ij} and c_{ij} when evaluating the performance of the standard multinomial logit estimator in the presence of measurement error.

The bottom panel of Figure 3 plots the distributions of willingness-to-pay estimates from the standard logit estimator across the three different measurement error data generating processes. 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 -0.54 and -0.60. 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. I can reject the null hypothesis that the distribution of willingness-to-pay estimates cover the true value of -0.5 for Simulations 4, 5, and 6 with t-statistics of -6.84, -6.62, and -2.50, respectively.

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.

4.1 Control Function Approach and Endogenous Travel Cost

Define $\epsilon_{ijt} = f(\xi_{ijt}) + \mu_{ijt}$ as the sum of the unobservable components in the data generating process for travel costs, c_{ijt}, defined by (5). 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 this unobserved component of travel cost, ϵ_{ijt}. Assuming for now 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:

$$\xi_{ijt} = \mathbb{E}[\xi_{ijt}|\epsilon_{ijt}] + \tilde{\xi}_{ijt} \tag{11}$$

The conditional expectation in (11) is a function of ϵ_{ijt} and can be approximated using a control function:

$$\xi_{ijt} = CF(\epsilon_{ijt}; \lambda) + \tilde{\xi}_{ijt} \tag{12}$$

where λ parameterizes the function $CF(\cdot)$. While (12) allows for many possible parametric assumptions, the simplest assumption is that

$$CF(\epsilon_{ijt}; \lambda) = \lambda\epsilon_{ijt} \tag{13}$$

Substituting (12) and (13) into (4) gives:

$$u_{ijt} = v_{ijt}(\theta) + \lambda\epsilon_{ijt} + \tilde{\xi}_{ijt} + \tilde{\varepsilon}_{ijt} \tag{14}$$

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 (14) 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., $(\tilde{\xi}_{ijt} + \tilde{\varepsilon}_{ijt}) \sim$T1EV. This assumption then allows me to estimate (14) 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 (14) that identifies the true target parameters, θ. However, this requires consistent estimates of the residual term ϵ_{ijt} from the first stage for travel costs (5) which is possible with 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).

With access to a valid travel cost instrument, z_{ijt}, it is possible to implement a two-stage control function estimator that identifies the true parameters 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 variable. 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 multinomial 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 (12) 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 full two-stage procedure (Petrin and Train, 2010).

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. If I ignore potential travel cost endogeneity and assume that true travel costs follow an analogous data generating process as (5), i.e.,

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

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

$$\hat{c}_{ijt} = w(z_{ijt}; \gamma) + \underbrace{h(\zeta_{ijt}) + \mu_{ijt}}_{\equiv e_{ijt}} \tag{15}$$

where once again 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}; $h(\cdot)$ is some unknown function that relates the correlated unobservable from (10), ζ_{ijt}, to observed travel cost $\hat{c}_{ijt}$; and μ_{ijt} is a mean-zero, idiosyncratic shock to travel costs.[6] Assuming that it is possible to observe e_{ijt}, which collects the unobservable terms in (15), I can specify a control function for the unobservable term, ζ_{ijt}. In particular, I can decompose this unobservable error term, ζ_{ijt} into its mean conditional on e_{ijt} and an independent component:

$$\zeta_{ijt} = \mathbb{E}[\zeta_{ijt}|e_{ijt}] + \tilde{\zeta}_{ijt} \tag{16}$$

[6] Note that I defined ζ_{ijt} as $\zeta_{ijt} = -\alpha(g(c_{ijt}) + \eta_{ijt})$ in (10).

The conditional expectation in (16) is itself a function of e_{ijt} and can be approximated using a control function:

$$\zeta_{ijt} = CF(e_{ijt}; \nu) + \tilde{\zeta}_{ijt} \tag{17}$$

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

$$CF(e_{ijt}; \nu) = \nu e_{ijt} \tag{18}$$

Substituting (9), (15), (17), and (18) into (8) gives:

$$u_{ijt} = \underbrace{X'_{ijt}\beta - \hat{c}_{ijt}\alpha + \xi_j}_{v_{ijt}(\hat{c}_{ijt};\theta)} + \nu e_{ijt} + \tilde{\zeta}_{ijt} + \tilde{\omega}_{ijt} \tag{19}$$

where by construction, both $\tilde{\zeta}_{ijt}$ and $\tilde{\omega}_{ijt}$ are idiosyncratic, unobserved, and independent. This final equation has an analogous form to (14) where the analyst now uses the observed travel cost $\hat{c}_{ijt}$ in the observable component of utility, $v_{ijt}(\cdot)$.

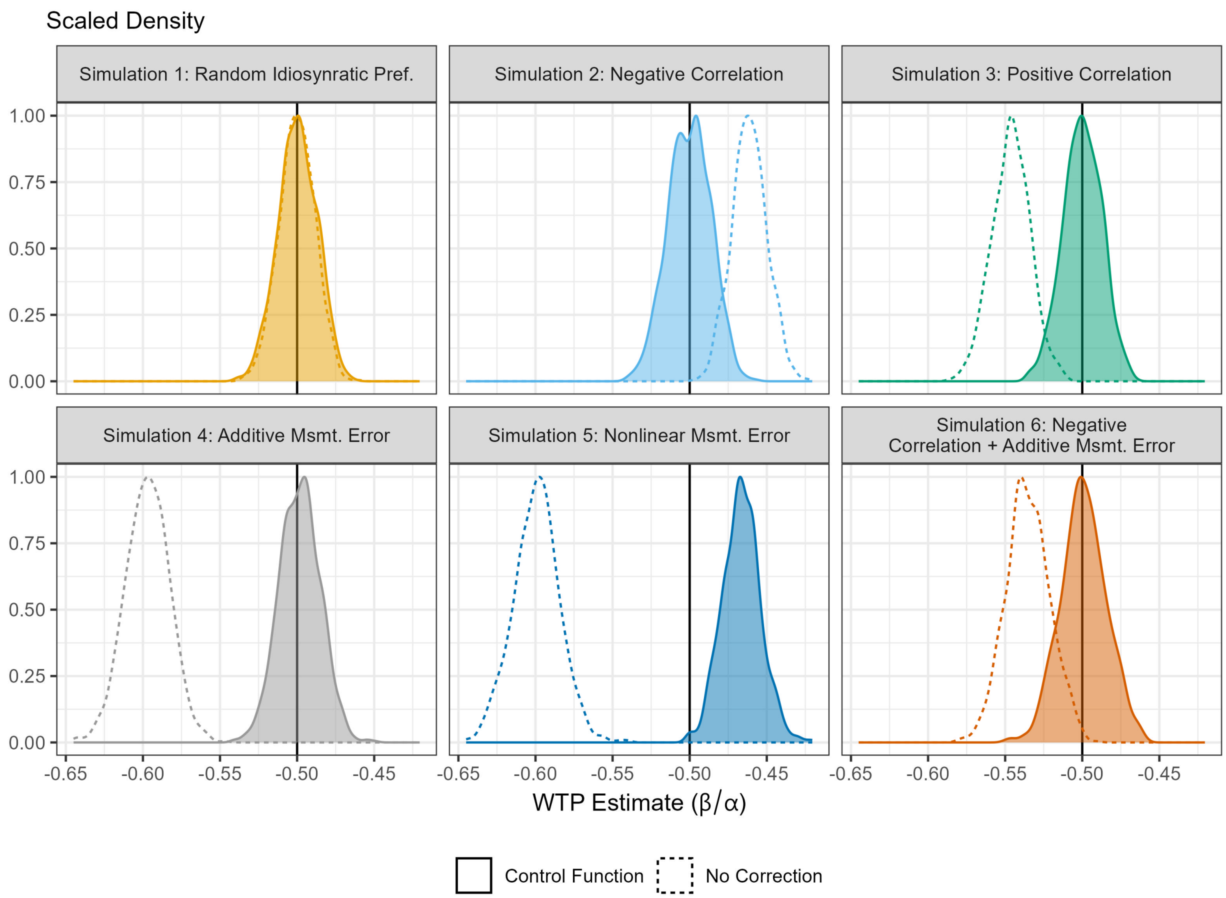
Making an analogous set of parametric and distributional assumptions on the nature of the correlation between observed and true travel costs and assuming the analyst observes a valid instrument for travel cost, z_{ijt}, (19) shows that is possible to account for the measurement error problem and recover 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 implement the two-stage control function estimator using the simulated choice data from 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 estimates. As shown in Table 1, the control function estimator reduces bias in willingness-to-pay estimates by one to two orders of magnitude in the case of Simulations 2 and 3.

The control function also has better coverage of the true willingness-to-pay statistic in the case of the measurement error simulations. In Simulations 4 and 5, which assume additive,

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.

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. The performance of the control function estimator is attenuated with non-classical measurement error in Simulation 5, though the estimator still substantially outperforms the standard logit estimator. I am able to reject the null hypothesis of equivalence between the

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	0.000	-0.004	-0.500	0.000	0.000	0.027
Non-random Sorting Simulations								
Simulation 2	-0.461	0.039	0.002	3.395	-0.500	-0.000	0.000	-0.015
Simulation 3	-0.546	-0.046	0.002	-3.709	-0.500	0.000	0.000	0.014
Measurement Error Simulations								
Simulation 4	-0.597	-0.097	0.010	-6.844	-0.498	0.002	0.000	0.135
Simulation 5	-0.599	-0.099	0.010	-6.619	-0.466	0.034	0.001	2.713
Simulation 6	-0.535	-0.035	0.001	-2.501	-0.499	0.001	0.000	0.063

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 simulations described in Section 3. The table reports the average WTP estimate, average bias, and mean squared error (MSE) across the 1000 simulations as well as the t-statistic for a two-sided test of the null hypothesis that the simulated WTP values equal the true value. Simulations 1 through 3 assume different non-random sorting data generating processes: one where the idiosyncratic preference and travel cost are independent (Simulation 1); one corresponding to a model of household sorting towards desirable recreation sites, where the idiosyncratic preference and travel cost are negatively correlated (Simulation 2); and one corresponding to a model of household sorting away from desirable recreation sites, where the idiosyncratic preference and travel costs are positively correlated (Simulation 3). Simulations 4 through 6 assume different measurement error data generating processes: one where there is additive, mean-zero measurement error in travel costs (Simulation 4); one where measurement error is increasing in travel distance (Simulation 5); 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 (Simulation 6).

the willingness-to-pay estimates from the control function estimator in Simulation 5 and true value based on the t-statistic reported in Table 1 of 2.71. 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 the results for Simulation 6 in Figure 4.

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.

4.4 Alternatives to the Control Function Approach

An existing best practice for implementing recreation demand models is to include site- or alternative-specific constants (Lupi et al., 2020). Murdock (2006) motivates the use of alternative-specific constants in recreation demand models as a means of addressing a price endogeneity due to the non-uniform distribution of residences and recreation locations. While the use of alternative-specific constants is indeed important to account for site-specific, unobservable factors influencing recreation decisions, it does not mitigate the particular source of endogeneity on which I focus: the non-random sorting of individuals based on unobservable preferences for recreation arises over both individual decision makers and sites rather than just sites alone. Indeed, the biased willingness-to-pay estimates reported in Figures 2 and 3 come from multinomial logit estimators that include alternative-specific constants.

A less common, but perhaps promising alternative to the control function approach relies on the use of recreator origin fixed effects. For example, Dundas and von Haefen (2020) use a spatially-broad origin fixed effect in their model of fishing site demand with a similar residence choice decision in mind. If individuals make residence location decisions in part based on their idiosyncratic preferences for specific recreation sites, then individuals residing in proximity to one another are likely to have correlated unobservable preferences for recreation. A set of fixed effects based on an aggregate geographic area linking plausibly similar individuals may therefore help directly control for—albeit in a course manner—the unobservable, idiosyncratic preference that gives rise to the motivating endogeneity problem.

I test the performance of this alternative approach across the six simulations discussed above by adding origin fixed effects to the baseline estimator. The data generating process for each simulation described by (6) and (7) assumes a degree of correlation in unobservable preferences by making ξ_{ij} a function of an idiosyncratic component and a group-specific component. This in turn ensures that individuals from the same origin group face a similar set of travel costs across alternatives—i.e., are located in proximity to one another—given the assumed data generating process for costs.

I report the full results from this alternative estimator with origin fixed effects alongside those from the baseline and control function estimators in Table A1. Interestingly, the approach using origin fixed effects offers a minimal performance improvement over the baseline estimator in the case of Simulations 2 and 3, and no performance improvement in the case of Simulations 4, 5, and 6. I test whether this is sensitive to the strength of within-group correlation in the unobservable preference term by running Simulation 2 with different values of ρ_ξ in the following data generating process for ξ_{ij}:

$$\xi_{ij} = (1 - \rho_\xi)\tilde{\xi}_{ij} + \rho_\xi \tilde{\xi}_{g(i)}$$

where higher values of $\rho_\xi \in [0, 1]$ imply a greater degree of within-group correlation in ξ_{ij}. As I show in Appendix Figure A1, while the control function estimator provides consistent estimates of willingness-to-pay for any degree of within-group correlation, the origin fixed effect approach only provides performance gains over the baseline estimator in the case of particularly high levels of within-group correlation of the unobservable preference term. Varying the number of groups in the data generating process for Simulation 2 while holding the strength of the within-group correlation fixed has little impact on the relative performance of the estimators as shown in Appendix Figure A2. Thus, while there are certainly cases in which this alternative offers an improvement over the baseline of no correction, in settings where non-random sorting or measurement error in travel costs are plausible concerns analysts can more effectively reduce bias by implementing the two-stage control function approach.

4.5 Robustness of the Control Function

I examine the robustness of the control function approach along two dimensions which are likely important in practice: functional form misspecification and weak instruments. As I outline in Sections 4.1 and 4.2, implementing the control function approach requires making an assumption on the functional form of the relationship between the first stage residual and the correlated unobservable—i.e., the conditional expectations in (11) and (16). In practice, I assume a linear control function. In constructing the simulated data generating processes, I assume that costs are linear in the observed instrument and both costs and indirect utility are linear in the unobservable preference term; however, it is possible in practice that there are nonlinearities in these relationships, which may affect the performance of my assumption of a linear control function.

I explore the robustness of this assumption by simulating versions of the Simulation 2 data generating process with different forms of nonlinearity in either the unobservable preference term or the observable instrument. Appendix Figure A3 plots the distribution of willingness-to-pay estimates from the baseline and control function estimators when the unobservable preference, ξ_{ij}, enters one or both of travel costs and indirect utility as a nonlinear function. I test both quadratic and exponential functional forms and find that regardless of the form or location of the nonlinearity, the control function estimator has better coverage of the true willingness-to-pay value than the baseline estimator. Note that in implementing the control function across each of these alternative data generating processes, I maintain the assumption of a linear control function.

These results reveal a strength of the control function approach: it only requires an assumption about the distribution of the unobservable preference conditional on the residual from a first stage regression of costs on instruments, where the residual is itself an arbitrary function of the unobservable preference term.[7] Since a given conditional distribution can imply many different joint distributions, the approach is sufficiently general to allow arbitrary and distinct functional forms for the unobservable term in utility and cost and still recover unbiased second stage estimates, assuming it is possible to obtain unbiased estimates of the first stage residual (Train, 2009).

Appendix Figure A4 plots the distribution of willingness-to-pay from each estimator when the observed instrument, z_{ij}, enters costs as a nonlinear function. Interestingly, the functional form of the true relationship between costs and the observable instrument has a material impact on the performance of the control function approach: despite offering better coverage of the true willingness-to-pay statistic than the baseline estimator, the control function estimates are now biased, with a mean absolute bias of 3.1% and 1.8% of the true value for quadratic and exponential relationships between costs and the instrument, respectively. This result is driven by the fact that the first stage regression no longer delivers unbiased estimates of the structural first stage residual. This suggests that practitioners should think carefully about the correct functional relationship between cost and their observable instruments when implementing the control function approach.

A second challenge in empirical applications is the potential for weak instruments. A large literature in econometrics documents that conventional approaches to instrumental variables estimation and inference are unreliable in the presence of weak instruments—instruments that are weakly correlated with endogenous regressors (Andrews et al., 2019; Stock et al., 2002). I explore the robustness of the control function approach to weak instruments by varying the strength of the observable instrument, z_{ij}, in the Simulation 2 data generating process and comparing the performance of the baseline and control function estimators.

I report the results from this robustness check in Appendix Figure A5: perhaps unsurprisingly, the control function estimator has far lower bias except for cases where the instrument has little to no correlation with travel costs. Fortunately, the instrumental variables literature develops a rule-of-thumb cutoff for the F-statistic that tests joint nullity of excluded instruments in the first stage regression that appears to also be helpful in the case of the two-stage control function: at levels of instrument strength in Appendix Figure A5 for which

[7] The exposition in Sections 4.1 and 4.2 implicitly assumes such a conditional distribution by assuming that I can decompose the unobservable into its mean conditional on the first stage residual and deviations around this mean and then placing assumptions on the distribution of these components.

the control function shows minimal bias, the first stage F-statistic exceeds the rule-of-thumb cutoff of 10.0 (Stock et al., 2002), suggesting a similar heuristic can be applied in this setting. Taken together, these results highlight the importance of the first stage regression in determining the performance of the control function approach. Valid instruments and unbiased first stage estimates are necessary to deliver accurate results in the second stage.

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. I examine the performance of the two-stage control function approach in the context of a recreation demand model that monetizes 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 a 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.

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 Oil Pollution Act of 1990. This 5-year, multi-million dollar effort involved both primary data collection and recreation demand modeling. English et al. (2018) summarize the main results from the recreation demand modeling component of this project, which finds that the monetary losses from foregone shoreline recreation following the DWH spill totaled \$661 million. The recreation demand model underlying English et al. (2018) is the primary focus of this section.

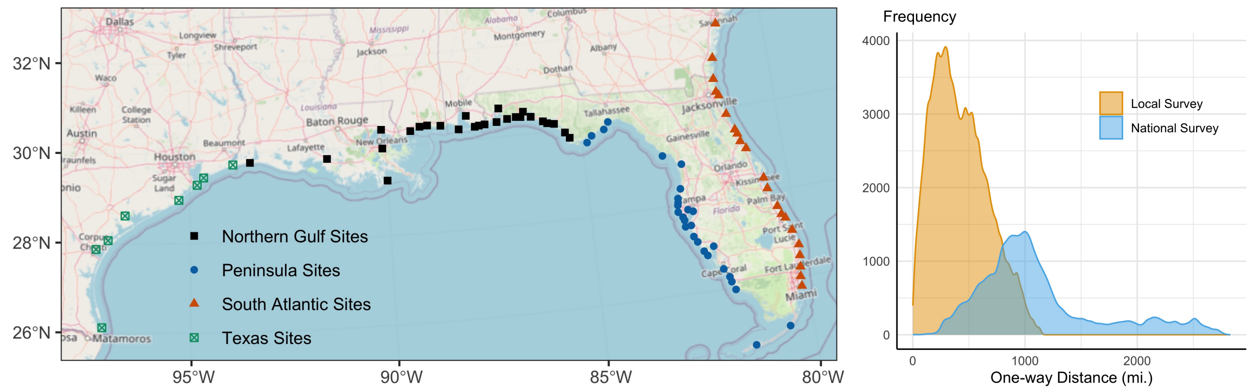
5.1 Empirical Setting and Data

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 conducted over the three years beginning immediately after the DWH spill. 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) and help calibrate monetary losses resulting from the DWH spill as I describe below.

The second source of primary data—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

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 from lost recreational user days, English et al. (2018) identify the Northern Gulf and Peninsula Sites as the affected sites in respondents' choice set.

on samples of adults in the continental US, included 43,000 interviews 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 the precise location of any recent shoreline recreation visits to the study area. English et al. (2018) weight the phone survey data to represent the target population, namely Gulf Coast recreational users that reside in the contiguous US.[8]

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

[8]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 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).

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 travel costs for every respondent-shoreline site combination.

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. I obtain the data that English et al. (2018) use directly from NOAA's Natural Resource damage Assessment (NRDA) public repository.[9] I refer the reader to English et al. (2018) for additional information on the empirical setting and data.

5.2 Model

English et al. (2018) use a nested logit model to characterize demand for shoreline recreation on the Gulf Coast. The authors 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 $i \in \{1, \ldots, N\}$ when selecting Gulf Coast shoreline site $j \in \{0, \ldots, J\}$, 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.[10] In the version of English et al. (2018)'s model that I implement, individual i's conditional utility from visiting shoreline site

[9] 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).

[10] 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.

j is $u_{ij} = v_{ij} + \varepsilon_{ij}$ with:

$$v_{ij} = \begin{cases} 0 & \text{for } j = 0 \\ \xi_j - c_{ij}\alpha & \text{for } j \in \{1, \ldots, J\} \end{cases} \tag{20}$$

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 c_{ij} is household-site-specific travel cost.[11]

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 (20), this implies choice probabilities of the following form:

$$p_{ij}(\theta) = \begin{cases} \frac{1}{1+\left(\sum_{k=1}^{J}\exp(v_{ik}(\theta)/\rho)\right)^{\rho}} & \text{for } j = 0 \\ \frac{\exp(v_{ij}(\theta)/\rho)}{\sum_{k=1}^{J}\exp(v_{ik}(\theta)/\rho)} \times \frac{\left(\sum_{k=1}^{J}\exp(v_{ik}(\theta)/\rho)\right)^{\rho}}{1+\left(\sum_{k=1}^{J}\exp(v_{ik}(\theta)/\rho)\right)^{\rho}} & \text{for } j \in \{1, \ldots, J\} \end{cases} \tag{21}$$

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 (20) and (21) to the data, estimating the target parameters, ρ and $\theta' = \begin{bmatrix}\alpha & \xi_1 & \ldots & \xi_J\end{bmatrix}$, 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. 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 cost must satisfy the relatively standard conditions of relevance and exogeneity: in other words, the instrument must plausibly affect idiosyncratic travel

[11] 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.

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 certainly 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 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.

A candidate instrument in this first category is a variable that captures the strength of local labor markets. Recent empirical work on models of equilibrium sorting indicates that labor market outcomes play a substantial role in determining households' choice of residence location (Kuminoff et al., 2013). Thus, variables that capture cross-sectional variation in the attractiveness of local labor markets are likely to shift individuals' residence choice, which in turn determines the distribution of travel costs associated with visiting a fixed set of recreation sites. Moreover, conditional on residence choice, such variables are plausibly exogenous to households' choice of recreation site. In the current empirical application, I use the employment rate, median income, and population in the Core Based Statistical Area (CBSA) in which individuals reside to capture cross-sectional variation in the state of local labor markets.[12] These measures, which are intended to proxy the desirability of local labor markets, will be correlated with travel costs based on the pre-existing distribution of labor markets in space relative to the recreation sites of interest.

There are several plausible examples of the second category of instruments, 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 A6. I interact variation in the rate of the gasoline tax in both the state of destination as well as the intermediate states between origin and destination at the time a respondent is surveyed with one-way driving distance as an instrument for travel cost in the present context. Though there is non-trivial cross-sectional variation in gasoline tax rates, I interact these tax rates with one-way driving distances as households are likely more sensitive to variation in gas taxes for trips to sites that are further away and therefore require

[12] I use Core Based Statistical Areas (CBSAs) as these are reasonable proxies for local labor markets given the use of commuting linkages to define their spatial extent. Data on these CBSA-level measures come from the US Census Bureau's 2012 5-year American Community Survey.

more 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.

Another example of an instrument that influences the marginal cost of site visitation directly is the price of crude oil. However, benchmark indices of the price of crude oil are likely correlated with seasonal or macroeconomic trends that also influence recreation demand. I therefore follow the approach of Kilian (2009) to isolate structural supply and demand shocks in the global crude oil market from expected production and consumption as well as seasonal fluctuations or changes in aggregate demand. The approach, which I describe in detail in Appendix B, involves a vector autoregression model using measures of crude oil production, real economic activity, and the price of oil.

The resulting structural oil supply and demand shocks isolate unanticipated movement in crude prices that are due to exogenous shocks in oil production and oil demand, respectively. I interact the resulting crude oil supply and demand shock time series, which I plot in Appendix Figure A7, 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 and households closer to the Gulf Coast shoreline sites likely respond differently to such shocks than households further away from those sites. 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 three separate first stage regressions of travel cost on a set of instruments and alternative-specific constants in Table 2. All results in Table 2 weigh observations using the sample weights constructed by English et al. (2018). Column (1) in Table 2 includes only the oil shock and state gas tax instruments, with estimated coefficients having the expected sign. In column (2), I add the CBSA-level measures and in column (3) I interact these measures with one-way driving distance to allow for a more flexible functional form between the spatial distribution of local labor market quality and travel costs. 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. I use residuals from the regression reported in column (3) of Table 2 when estimating the second stage.

Table 3 reports parameter estimates and bootstrapped standard errors from maximum likelihood estimation of the second stage discrete choice model of visitation to Gulf Coast

Table 2. First Stage Estimates

	Travel Cost		
	(1)	(2)	(3)
Distance ×			
Oil Supply Shock	−0.050 (0.011)	−0.045 (0.010)	−0.031 (0.004)
Oil Demand Shock	0.002 (0.0004)	0.002 (0.0004)	0.001 (0.0002)
Destination Gas Tax	0.017 (0.0002)	0.018 (0.0002)	0.002 (0.002)
Mean Route Gas Tax	0.105 (0.009)	0.093 (0.009)	0.091 (0.007)
log(Employment Rate)		529.4 (14.1)	291.7 (33.1)
×Distance			0.264 (0.022)
log(Median Income)		81.9 (5.74)	227.4 (17.5)
×Distance			−0.114 (0.010)
log(Population)		6.82 (0.577)	−13.6 (1.23)
×Distance			0.018 (0.0009)
Alternative Specific Constants	Yes	Yes	Yes
Observations	3,462,428	3,462,428	3,462,428
R^2	0.356	0.361	0.364
Within R^2	0.353	0.358	0.361
F-statistic	3,798.6	5,799.6	9,917.2

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. All results weigh observations using the sample weights constructed by English et al. (2018).

shoreline recreation sites. 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 by bootstrapping the entire two-stage procedure. The estimated travel cost parameter increases in magnitude relative to the baseline esitmates: households' mean sensitivity to travel cost increases in magnitude from -1.08 in the uncorrected, baseline estimates to -1.21 in the control function estimates, a 12% increase. Thus, it appears as though failing to account for potential non-random selection into 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.[13] 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) requires 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 \$6.17, a 12% decrease 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 translates into an overestimation of the total recreation-based losses from the DWH oil spill of around \$82 million, which suggests that accounting for the challenges 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 based solely 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 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

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

Table 3. Second Stage Estimates

	Parameter	Baseline (1)	Control Function (2)
Travel Cost ($100s)	α	-1.080 (0.075)	-1.206 (0.068)
Nesting Parameter	ρ	0.208 (0.013)	0.160 (0.013)
First Stage Residual	λ		0.843 (0.049)
Alternative Specific Constants		Yes	Yes
Lost User Day Value ($/day)		7.035	6.165
N		41,716	
Sites		83	

Notes: This table reports estimates from the second stage discrete choice model of Gulf Coast shoreline recreation site visitation. 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. Bootstrapped standard errors calculated using 200 replications are reported in parentheses. Standard errors in Column (2) account for noise from the first stage regression by replicating the complete two-stage procedure for each bootstrap sample. The process for calculating lost user day values is described in detail in Appendix C.

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. While it is possible that the true data generating process in the DWH empirical application matches one of Simulations 3 through 6, it is difficult to determine from Table 3 alone which story is most plausible.

The choice of instruments in this empirical application may offer a simple check of what could account for the observed differences between the baseline and control function welfare estimates. Though both sets of instruments—those that shift residence location choices and those that shift travel costs—should effectively handle both issues of non-random sorting and measurement error in travel costs, the motivating logic for each is quite different. For instance, the logic motivating the CBSA-level labor market instruments is that they can account for non-random spatial sorting of households. Though the structural oil market shocks and state gas taxes provide exogenous cost shifters that address the threat to identification of non-random sorting, they are also plausibly orthogonal to the forms of measurement error that are likely present based on the construction of travel costs. So if I assume that the CBSA-level instruments are more effective at accounting for non-random sorting and the cost shifters are more effective at addressing measurement error, I can separately estimate the control function approach with each set of instruments to see how they influence the magnitude and direction of bias.

I report the parameter estimates and calculated lost user day values from this exercise in Appendix Table A2. Though there is little difference between the estimated lost user day value using the full set of instruments and that calculated from a specification with just the oil shock and gas tax instruments—columns (1) and (2) of Appendix Table A2—there is a large difference in this welfare statistic when only using the CBSA-level labor market instruments. This specification, reported in column (3) of Appendix Table A2, suggests a greater degree of overestimation of the magnitude of the lost user day value in the baseline estimator. If we take the differences in these sets of instruments seriously, this suggests that sorting may drive a significant share of the observed bias in the baseline, uncorrected estimates. Moreover, the direction of the bias is consistent with non-random sorting away from desirable recreation sites as indicated by the results for Simulation 3 above, which may be plausible given the scale of the target population.

While this exercise is suggestive, it is in no way conclusive. Importantly, the non-random sorting and measurement error issues are not separable as indicated by their analogous structure in Sections 3.1 and 3.2 and any valid instrument that is correlated with costs but independent of unobserved preferences and/or measurement error should effectively address both issues. Moreover, 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. 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 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.

5.5 External Validity

The degree to which non-random sorting and measurement error affect conventional estimates of recreation demand models is likely context-specific. The magnitude and direction of bias in these models will likely vary based on the target population of decisionmakers; the scope and scale of the choice set under analysis; and the methods used to collect and construct the necessary choice data, among other factors. Though the goal of this empirical application building on English et al. (2018) is simply to demonstrate that meaningful differences between baseline estimates and control function estimates are possible in empirical settings, the remarkable scale of this particular application raises important questions around how well this finding generalizes to more conventional recreation demand applications.

To explore the external validity of the main findings in Table 3, I estimate a version of the baseline nested logit model and the analogous two-stage control function estimator using only choice data from respondents that English et al. (2018) define as local to the Gulf

Coast shoreline sites.[14] This exercise offers a test of the influence of geographic scale, which may plausibly lead to differences in the magnitude and direction of bias in conventional recreation demand estimates. For example, it is possible that, conditional on deciding to live in an area near the Gulf Coast, unobservable preferences for specific shoreline sites are more likely to lead households to sort towards those sites, which would generate bias in the opposite direction to what we observe in Table 3.

I report results replicating the first and second stages of the two-stage control function estimator using just the local respondents in Appendix Tables A3 and A4. I find no difference in the direction of the apparent bias in the conventional estimates using only the local respondents, suggesting that the main findings of this empirical exercise may generalize to smaller scale applications. Interestingly, the level of my calculated lost user day value among the local respondents is lower than that using the nationwide sample; however, it is possible that there are important differences in the reduction in visitation during spill conditions across local and non-local respondents for which this calculation does not account.

Moreover, the scale and scope of these "local" estimates are still quite expansive. Conventional applications of recreation demand estimation are often focused at even smaller scales, for example analyzing lake visitation at the state- or local-level. It is indeed quite possible that the dominant form of non-random sorting in such applications would be for households to locate closer to sites for which they have an unobservable, strong preference, which would result in bias in the opposite direction of what I find in the English et al. (2018) context. Despite the context-specific nature of these empirical challenges, exploring how the magnitude and direction of bias in conventional models varies across applications of different scales remains an interesting area for future research.

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,

[14] This definition includes all respondents from Louisiana, Mississippi, Alabama, and Florida, as well as respondents from Texas and Georgia that are within a half-day drive of a site.

feasible approach to address 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 decreases welfare estimates by 12% 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 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.

Andrews, Isaiah, James H. Stock, and Liyang Sun. 2019. "Weak Instruments in Instrumental Variables Regression: Theory and Practice." *Annual Review of Economics*, 11(Volume 11, 2019): 727–753.

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.

Bureau of Economic Statistics. 2023. "Outdoor Recreation Satellite Account, U.S. and States, 2023." Bureau of Economic Statistics.

Dundas, Steven J., and Roger H. von Haefen. 2020. "The Effects of Weather on Recreational Fishing Demand and Adaptation: Implications for a Changing Climate." *Journal of the Association of Environmental and Resource Economists*, 7(2): 209–242.

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. 2019. "The Missing Benefits of Clean Water and the Role of Mismeasured Pollution." *Journal of the Association of Environmental and Resource Economists*, 6(4): 669–707.

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.

Kuminoff, Nicolai V., V. Kerry Smith, and Christopher Timmins. 2013. “The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets.” *Journal of Economic Literature*, 51(4): 1007–1062.

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.

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., 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.

Stock, James H, Jonathan H Wright, and Motohiro Yogo. 2002. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business & Economic Statistics*, 20(4): 518–529.

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.

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[1]

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

—Appendix Section A: Supplemental Figures and Tables

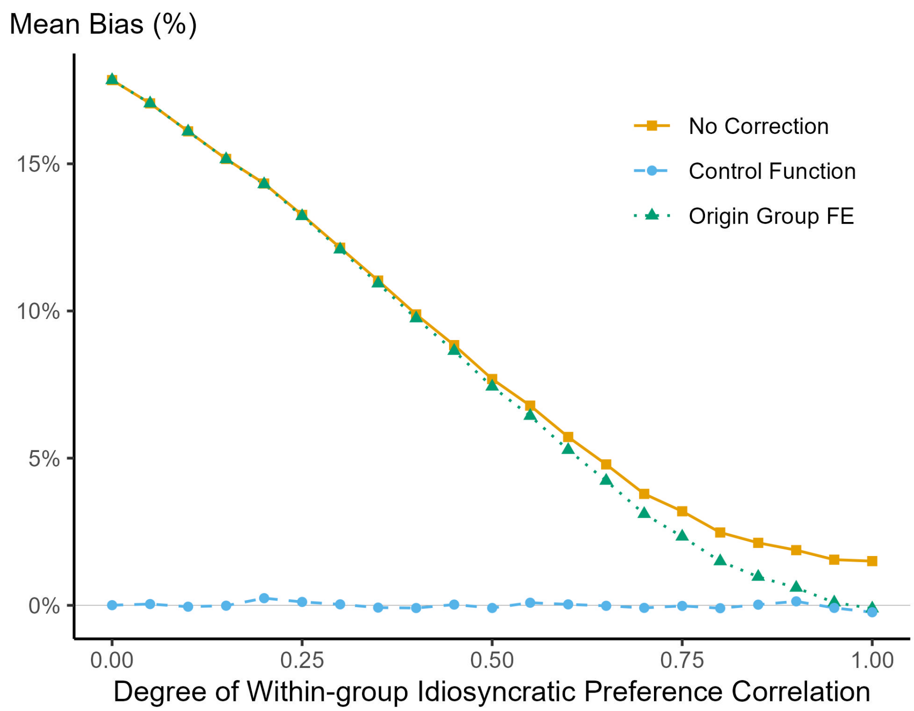
—Appendix Section B: Estimating Structural Oil Market Shocks

—Appendix Section C: Calculating Lost User Day Value

[1]The University of Texas at Austin, McCombs School of Business; jacob.bradt@austin.utexas.edu.

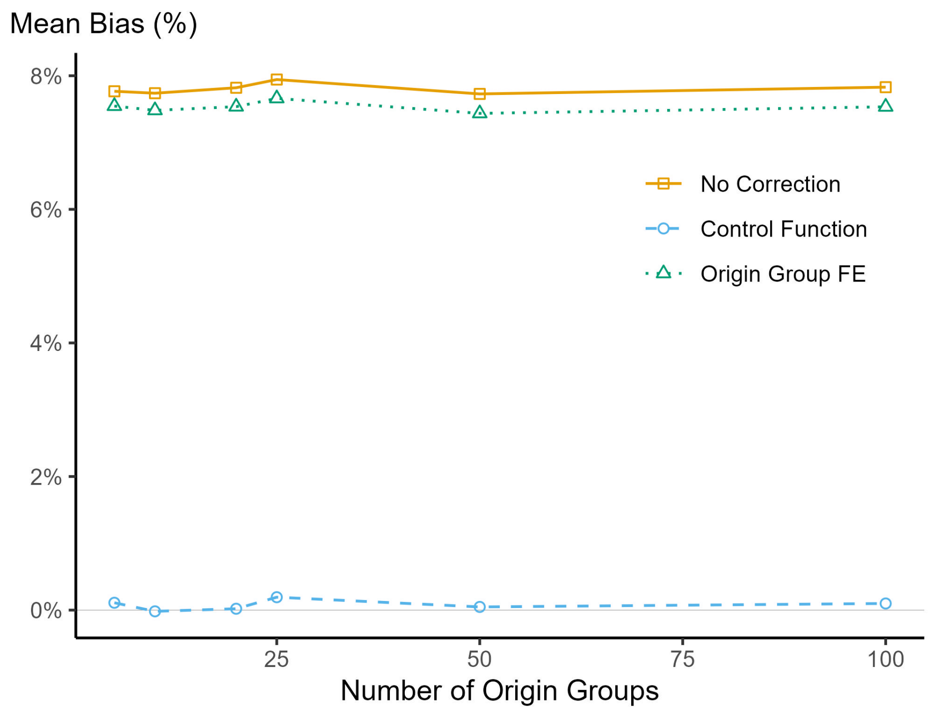
A Supplemental Figures and Tables

Figure A1. Performance of Estimators under Varying Degrees of Within-group Idiosyncratic Preference Correlation



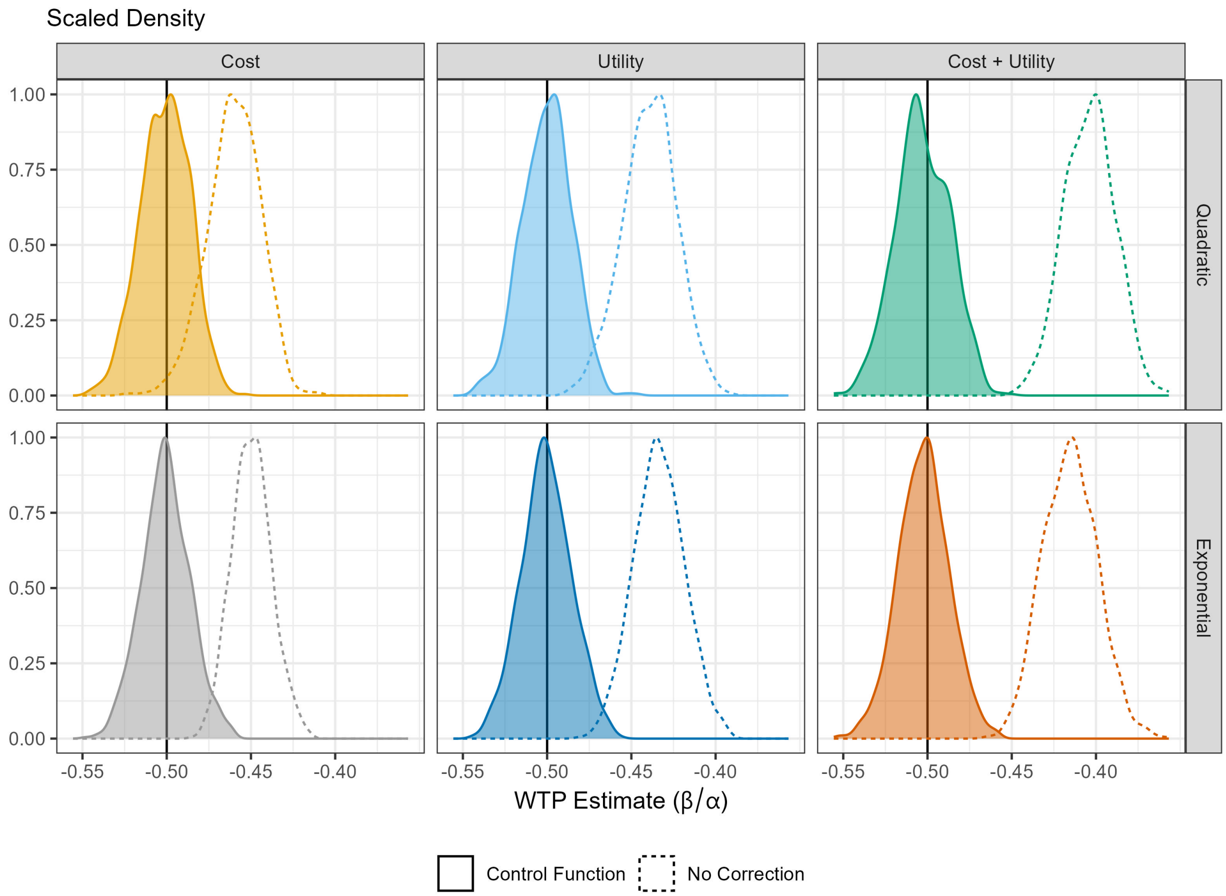
Notes: This figure plots the mean bias of estimated willingness-to-pay (WTP) from three different estimators for different amounts of within-group correlation in the idiosyncratic, unobservable preference, ξ_{ij}, under the Simulation 2 data-generating process: the baseline multinomial logit estimator ("No Correction"), the multinomial logit estimator with origin group fixed effects ("Origin Group FE"), and the two-stage control function estimator ("Control Function"). Values on the horizontal axis correspond to ρ_ξ in the following data generating process for the unobservable preference: $\xi_{ij} = (1-\rho_\xi)\tilde{\xi}_{ij} + \rho_\xi \tilde{\xi}_{g(i)}$, where $\tilde{\xi}_{ij}$ is an idiosyncratic component and $\tilde{\xi}_{g(i)}$ is a component common across groups of individuals, $g(i)$. Going from 0.0 to 1.0 on the horizontal axis therefore results in an increasing degree of within-group correlation of the unobservable preference term. For each estimator and value on the horizontal axis, I show the mean bias of WTP estimates as a percentage of the true WTP value across 1000 simulated datasets, each of which includes 10,000 individuals from 10 different origin groups choosing between 20 different alternatives. The data-generating process is identical to that given by (6) and (7) under the Simulation 2 assumptions, with the exception of the introduction of ρ_ξ.

Figure A2. Performance of Estimators under Different Recreator Origin Group Sizes



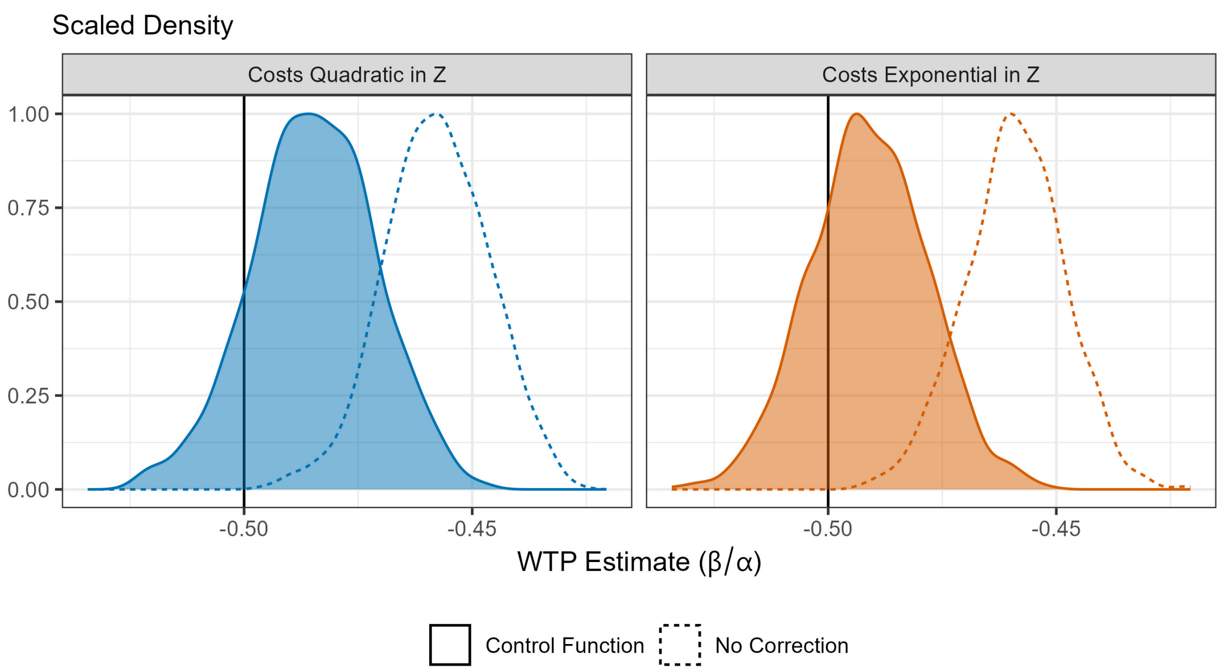
Notes: This figure plots the mean bias of estimated willingness-to-pay (WTP) from three different estimators for different numbers of recreator origin groups under the Simulation 2 data-generating process, holding fixed the number of simulated individuals: the baseline multinomial logit estimator ("No Correction"), the multinomial logit estimator with origin group fixed effects ("Origin Group FE"), and the two-stage control function estimator ("Control Function"). Values on the horizontal axis correspond to the number of different recreator origin groups, $\left|\{g(i)\}_{\forall i}\right|$, used in the data-generating process for the unobservable preference, $\xi_{ij} = 0.5\tilde{\xi}_{ij} + 0.5\tilde{\xi}_{g(i)}$. For each estimator and value on the horizontal axis, I show the mean bias of WTP estimates as a percentage of the true WTP value across 1000 simulated datasets, each of which includes 10,000 individuals choosing between 20 different alternatives. The data-generating process is identical to that given by (6) and (7) under the Simulation 2 assumptions, with the exception of number of groups.

Figure A3. Performance of Two-stage Control Function Estimator with Nonlinear Unobservable Preference in Data-generating Process



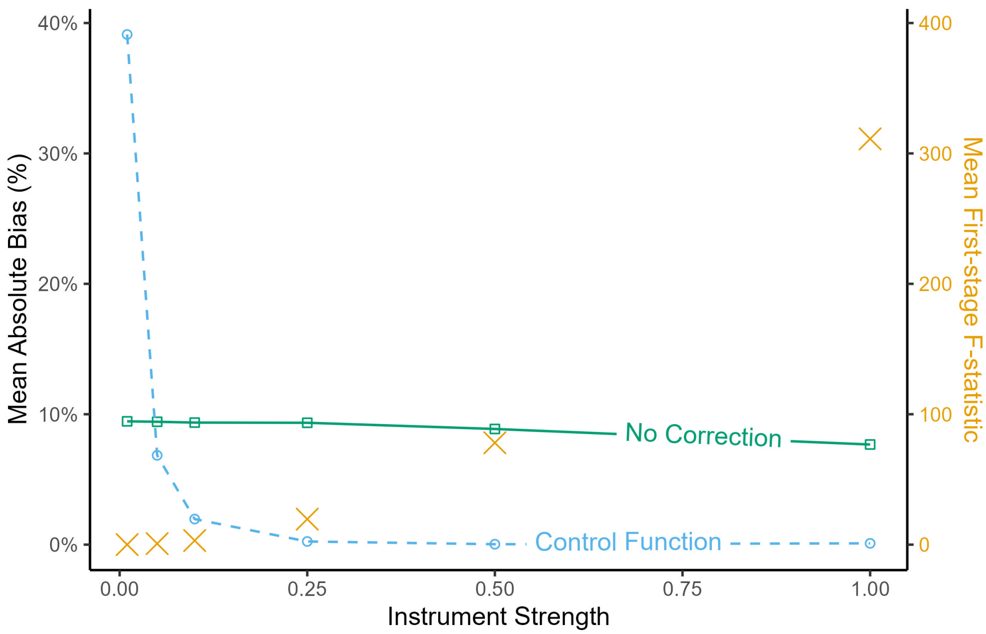
Notes: This figure plots the distribution of estimated willingness-to-pay (WTP) using the baseline multinomial logit estimator and two-stage control function estimator across 1000 simulated datasets for different data-generating processes. Each panel corresponds to a different nonlinearity in the unobservable preference term, ξ_{ij}, under the Simulation 2 data-generating process. Specifically, for each set of simulations (i.e., panel), I replace ξ_{ij} with some nonlinear function, $f(\xi_{ij})$, in the Simulation 2 data-generating process outlined in Section 3. In the left panels, ξ_{ij} enters costs, c_{ij}, in (6) as a nonlinear function; in the center panels, ξ_{ij} enters utility, u_{ij}, in (6) as a nonlinear function; and in the right panels, ξ_{ij} enters both c_{ij} and u_{ij} as nonlinear functions. In the top panels, ξ_{ij} enters as a simple quadratic, i.e., $f(\xi_{ij}) = \xi_{ij}^2 + \xi_{ij}$. In the bottom panels, ξ_{ij} enters as an exponential, i.e., $f(\xi_{ij}) = \exp(\xi_{ij})$. All other aspects of the simulated data-generating process are held fixed as I outline in Section 3, including the fact that $\xi_{ij} = 0.5\tilde{\xi}_{ij} + 0.5\tilde{\xi}_{g(i)}$ and the assumption that $\tilde{\xi}_{ij}, \tilde{\xi}_{g(i)} \sim \mathcal{N}(0.0, 1.0)$. The true value of the the willingness-to-pay statistic is shown as a vertical black line.

Figure A4. Performance of Two-stage Control Function Estimator with Nonlinear Cost Instrument in Data-generating Process



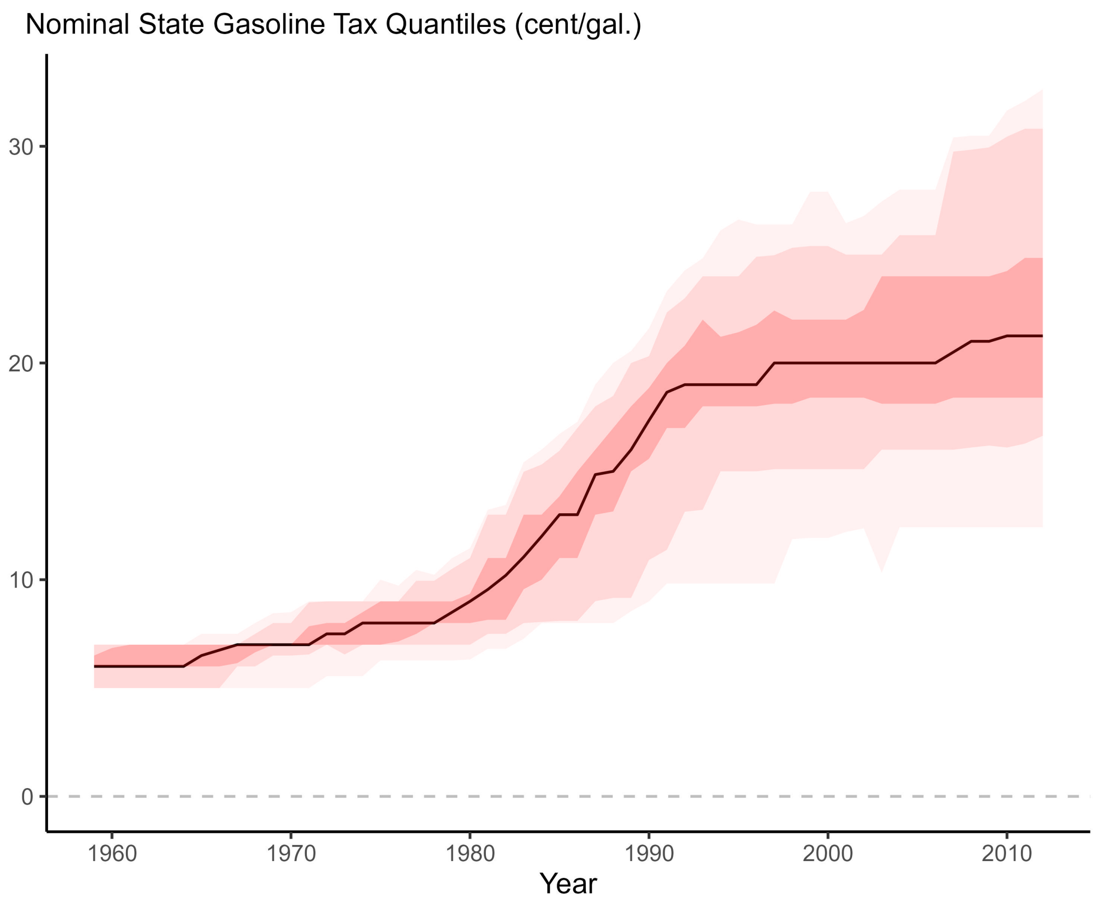
Notes: This figure plots the distribution of estimated willingness-to-pay (WTP) using the baseline multinomial logit estimator and two-stage control function estimator across 1000 simulated datasets for two different data-generating processes. Each panel corresponds to a different nonlinearity in the exogenous cost instrument, z_{ij}, under the Simulation 2 data-generating process. Specifically, for each set of simulations (i.e., panel), I replace z_{ij} with some nonlinear function, $g(z_{ij})$, in the Simulation 2 data-generating process outlined in Section 3. in the left panel, z_{ij} enters costs, c_{ij}, in (6) as a quadratic function, i.e., $g(z_{ij}) = z_{ij}^2 + z_{ij}$. In the right panel, z_{ij} enters c_{ij} as an exponential, i.e., $g(z_{ij}) = \exp(z_{ij})$. All other aspects of the simulated data-generating process are held fixed as I outline in Section 3. The true value of the the willingness-to-pay statistic is shown as a vertical black line.

Figure A5. Performance of Two-stage Control Function Estimator with Weak Instrument



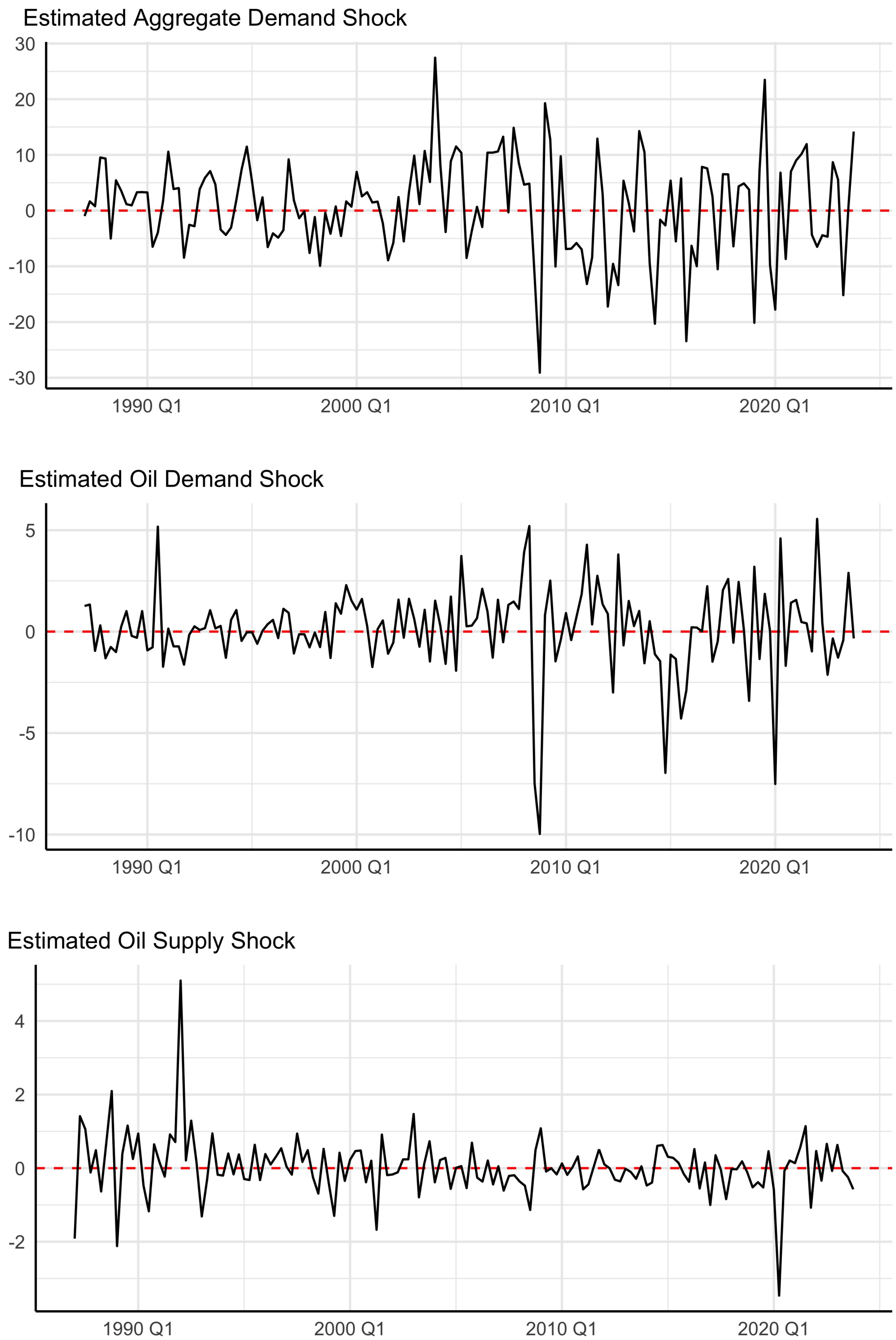
Notes: This figure shows the performance of the two-stage control function estimator in the presence of a weak cost instrument. In particular, this figure plots the mean absolute bias (left axis) of estimated willingness-to-pay (WTP) using the baseline multinomial logit estimator ("No Correction") and two-stage control function estimator ("Control Function") for different levels of instrument strength for the cost instrument, z_{ij}. Each set of simulated datasets follows the structure and assumptions of Simulation 2 as outlined in Section 3, with the exception of the assumed strength of the instrument in the data-generating process for costs, c_{ij}. The horizontal axis corresponds to different levels of instrument relevance in the data-generating process for costs, i.e., the ρ_z in $c_{ij} = 5.0 + \rho_z z_{ij} - 1.0\xi_{ij} + \mu_{ij}$ where the baseline assumption from Simulation 2 is that $\rho_z = 1.0$. Instrument strength is therefore increasing from left to right on the horizontal axis. For each estimator and value of ρ_z shown on the horizontal axis, I show the mean absolute bias of WTP estimates as a percentage of the true WTP value across 1,000 simulated datasets. I also plot the average F-statistic across the first-stage regressions used to implement the control function approach for each simulated dataset (right axis).

Figure A6. Variation in Nominal State Gasoline Taxes over Time



Notes: This figure plots the distribution of gasoline tax rates across US states from 1960 to 2012. In particular, the shaded areas show the following annual percentiles of gasoline tax rates: 5th, 10th, 30th, 70th, 90th, and 95th. The black line shows the median state gasoline tax rate. All values are nominal cents per gallon. I use the geographic variation in gasoline taxes within calendar year 2012 as discussed in Section 5.3 of the text.

Figure A7. 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).

Table A1. Willingness-to-pay Estimates for Simulated Choice Data across Three Estimators

	Baseline Estimator				Origin Group FE				Control Function			
	Mean	Bias	MSE	t-stat	Mean	Bias	MSE	t-stat	Mean	Bias	MSE	t-stat
Baseline Simulation												
Simulation 1	-0.500	-0.000	0.000	-0.004	-0.500	-0.000	0.000	-0.004	-0.500	0.000	0.000	0.027
Non-random Sorting Simulations												
Simulation 2	-0.461	0.039	0.002	3.395	-0.463	0.037	0.002	3.273	-0.500	-0.000	0.000	-0.015
Simulation 3	-0.546	-0.046	0.002	-3.709	-0.544	-0.044	0.002	-3.580	-0.500	0.000	0.000	0.014
Measurement Error Simulations												
Simulation 4	-0.597	-0.097	0.010	-6.844	-0.597	-0.097	0.010	-6.842	-0.498	0.002	0.000	0.135
Simulation 5	-0.599	-0.099	0.010	-6.619	-0.599	-0.099	0.010	-6.619	-0.466	0.034	0.001	2.713
Simulation 6	-0.535	-0.035	0.001	-2.501	-0.537	-0.037	0.002	-2.647	-0.499	0.001	0.000	0.063

Notes: This table reports summary statistics for willingness-to-pay estimates using the baseline multinomial logit estimator, the baseline multionial logit estimator with origin group fixed effects, and two-stage control function estimator across 1000 simulated datasets for each of the 6 simulations described in Section 3. The table reports the average WTP estimate, average bias, and mean squared error (MSE) across the 1000 simulations as well as the t-statistic for the two-sided test of the null hypothesis that the simulated WTP values equal the true value. Simulations 1 through 3 assume different non-random sorting data generating processes: one where the idiosyncratic preference and travel cost are independent (Simulation 1); one corresponding to a model of household sorting towards desirable recreation sites, where the idiosyncratic preference and travel cost are negatively correlated (Simulation 2); and one corresponding to a model of household sorting away from desirable recreation sites, where the idiosyncratic preference and travel costs are positively correlated (Simulation 3). Simulations 4 through 6 assume different measurement error data generating processes: one where there is additive, mean-zero measurement error in travel costs (Simulation 4); one where measurement error is increasing in travel distance (Simulation 5); 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 (Simulation 6).

Table A2. Second Stage Estimates from English et al. (2018) Empirical Application Using Different Combinations of First Stage Instruments

	Instrumental Variables:	Baseline	Cost Shifters	Sorting Shifters
	Parameter	(1)	(2)	(3)
Travel Cost ($100s)	α	-1.206 (0.068)	-1.266 (0.076)	-1.318 (0.071)
Nesting Parameter	ρ	0.160 (0.013)	0.180 (0.015)	0.157 (0.014)
First Stage Residual	λ	0.843 (0.049)	0.855 (0.054)	0.960 (0.060)
Alternative Specific Constants		Yes	Yes	Yes
Lost User Day Value ($/day)		6.165	6.254	5.566
N		41,716	41,716	41,716
Sites		83	83	83

Notes: This table reports estimates from the second stage discrete choice model of demand for visits to Gulf Coast shoreline recreation sites using different combinations of first stage instruments. Parameters are estimated via maximum likelihood estimation. Column (1) implements the two-stage control function estimator described in Section 5 using the baseline set of instruments discussed in Section 5.3, with results matching those reported in Column (2) of Table 3. Column (2) implements the two-stage control function estimator using only those instruments that act as plausible marginal travel cost shifters in the first stage: destination and intermediate state gas taxes interacted with one-way driving distance and structural oil supply and demand shocks interacted with one-way driving distance. Column (3) implements the two-stage control function estimator using only those instruments that act as plausible shifters of residential sorting in the first stage: population, median income, and employment rates at the Core Based Statistical Area of origin, both without interactions and interacted with one-way driving distance. Bootstrapped standard errors are reported in parentheses and account for noise from the first stage regression by replicating the complete two-stage procedure for each bootstrap sample. The process for calculating lost user day values is described in detail in Appendix C.

Table A3. First Stage Estimates from English et al. (2018) Empirical Application Using Local Respondents

	Travel Cost		
Sample:	National (1)	Non-local Respondents (2)	Local Respondents (3)
Distance ×			
Oil Supply Shock	-0.031 (0.004)	-0.057 (1.89)	0.006 (0.0008)
Oil Demand Shock	0.001 (0.0002)	0.002 (0.072)	0.0010 (0.00006)
Destination Gas Tax	0.002 (0.002)	0.0004 (0.002)	0.001 (0.003)
Mean Route Gas Tax	0.091 (0.007)	0.062 (0.006)	-0.010 (0.024)
log(Employment Rate)	291.7 (33.1)	239.5 (22.5)	-58.6 (24.1)
× Distance	0.264 (0.022)	0.193 (0.021)	0.490 (0.046)
log(Median Income)	227.4 (17.5)	190.2 (20.2)	151.3 (7.05)
× Distance	-0.114 (0.010)	-0.079 (0.012)	-0.187 (0.023)
log(Population)	-13.6 (1.23)	-12.6 (1.46)	-5.02 (0.778)
× Distance	0.018 (0.0009)	0.016 (0.001)	0.025 (0.003)
Alternative Specific Constants	Yes	Yes	Yes
Observations	3,462,428	1,112,781	2,349,647
R^2	0.364	0.258	0.606
Within R^2	0.361	0.255	0.588
F-statistic	9,917.2	5,594.8	3,831.2

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. Column (1) includes the full choice data from English et al. (2018) and corresponds to the results reported in column (3) of Table 2, whereas columns (2) and (3) subset to respondents that are not local or are local to the Gulf Coast recreation sites, respectively. All results weigh observations using the sample weights constructed by English et al. (2018) to weight observations, where the weights in columns (2) and (3) are adjusted to reflect the different sample definition.

Table A4. Second Stage Estimates from English et al. (2018) Empirical Application Using Local Respondents

	Parameter	Baseline (1)	Control Function (2)
Travel Cost ($100s)	α	-3.662 (0.916)	-6.736 (1.661)
Nesting Parameter	ρ	0.391 (0.077)	0.722 (0.200)
First Stage Residual	λ		6.826 (1.862)
Alternative Specific Constants		Yes	Yes
Lost User Day Value ($/day)		3.001	1.862
N		28,309	
Sites		83	

Notes: This table reports estimates from the second stage discrete choice model of demand for visits to Gulf Coast shoreline recreation sites using respondents that are local to the sites as defined by English et al. (2018). 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. Bootstrapped standard errors are reported in parentheses. Standard errors in Column (2) account for noise from the first stage regression by replicating the complete two-stage procedure for each bootstrap sample. The process for calculating lost user day values is described in detail in Appendix C.

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.[2] 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

[2] 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 A7 plots quarterly averages of the estimated monthly innovations to aggregate demand, oil demand, and oil supply. As is clear from Figure A7, 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 A7, 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.[3]

[3] 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.

References

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.

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.

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.