ESTIMATING THE VALUE OF LOST RECREATION DAYS FROM THE DEEPWATER HORIZON OIL SPILL

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

Revised June 2, 2018

Abstract: The 2010 Deepwater Horizon oil spill in the Gulf of Mexico was the largest ever in U.S. waters, eclipsing the 1989 Exxon Valdez spill in terms of the sheer quantity of oil released and the scale and scope of activities impacted. We developed a recreation demand model to monetize economic damages associated with lost shoreline recreational user days attributable to the spill. The unprecedented magnitude of the spill disruption led to a variety of innovations. We estimate a model of shoreline recreation trips to the Gulf Coast region from the general population of the contiguous U.S., combining single and multiple-day trips, calculating travel costs that incorporate detailed information on flying costs and transportation mode choice, and using alternative-specific constants to control for site characteristics. Losses per recreational user day are assessed using utility adjustments that reproduce the decline in recreation observed through on-site counts. Sensitivity analyses demonstrate our lost user day value is robust to changes in income imputation, nesting structure, site aggregation and spill calibration, and show the importance of accounting for flying as a mode choice. Estimated losses from the primary shoreline study are \$520 million (±166) out of the total recreational damages of \$661 million (2015\$).

JEL codes: Q26, Q51

Key words: Recreation demand; Discrete choice model; Natural resource damage assessment

Running title: Recreation Losses from Gulf Oil Spill

* Corresponding author: Frank Lupi, lupi@msu.edu; 517-599-9350; Michigan State University, Department of Agriculture, Food, and Resource Economics, 446 W. Circle Dr, Room 202, East Lansing, MI 48824.

** English: Bear Peak Economics, eric.english@bearpeakeconomics.com. von Haefen: North Carolina State University, rhhaefen@ncsu.edu. Herriges: Michigan State University, jah@msu.edu. Leggett: Bedrock Statistics, chris.leggett@bedrockstatistics.com. Lupi: Michigan State University, lupi@msu.edu. McConnell: University of Maryland, tmcconn@umd.edu. Welsh, Independent Consultant, welsh@mailbag.com. Domanski: ECONorthwest, domanski@econw.com. Meade: National Oceanic and Atmospheric Administration, norman.meade@noaa.gov.
The authors gratefully acknowledge the many important contributions to this study. Roger Tourangeau, Michael Hanemann, Kerry Smith and Jeffrey Wooldridge were instrumental in the design stage. The following individuals were essential to the empirical components: Nathan Braun, Mike Brick, Ismael Flores Cervantes, Eric Horsch, Stanley Presser, Colleen Donovan, Anthony Dvarskas, Fiona Garvin, Megan Lawson, Tim Meernik, Jean Opsomer, Keith Rust, Justin Stein, Sophie Tyack, Carolyn Wagner, and Jacqueline Wilwerth. This work was funded by the National Oceanic and Atmospheric Administration (NOAA). The views presented here are those of the authors and do not represent the views of NOAA or any other state or federal agency.

I. Introduction

On April 20, 2010, British Petroleum's (BP's) Deepwater Horizon (DWH) drilling rig exploded and later sank 50 miles off the Louisiana shore. The accident killed eleven workers and caused a massive oil spill in the Gulf of Mexico. For the next 87 days, oil billowed continuously out of damaged drilling equipment. A U.S. District Court's finding of fact later concluded that 134 million gallons of oil were released (U.S. v. BP et al. 2015), making the DWH spill the largest ever in U.S. waters and an order of magnitude larger than the 1989 Exxon Valdez spill. Oil from the DWH spill washed ashore on beaches and tidal marshes in Texas, Louisiana, Mississippi, Alabama, and Florida. National Oceanic and Atmospheric Administration (NOAA) models predicted during the spill that the oil might reach Florida's Atlantic Coast and Keys. The spill was the most closely followed news item in the U.S. throughout the late spring and summer of 2010 (Pew Research Center 2010; Welsh 2015).

Within days of the spill, NOAA's Office of Response and Restoration initiated efforts to assess recreation-related welfare losses under the authority of the Oil Pollution Act of 1990 (OPA). NOAA assembled a team of economists and survey experts who worked for five years to assess these losses. The effort dwarfed past recreational assessments, launching eight separate surveys and costing tens of millions of dollars. The final estimate of recreation damages was $661 million. This paper describes the assessment of lost shoreline recreation, with emphasis on the shoreline valuation model.

Under OPA, responsible parties are required to restore, rehabilitate, replace or acquire the equivalent natural resources and/or services to what was lost. Full recovery can take years, and the public suffers losses while resources are diminished. Under existing law, these "interim losses" are compensable as part of a natural resource damage assessment (NRDA) and pursuable by natural resource trustees on the public's behalf. Our assessment of recreation losses was conducted on behalf of NOAA and other federal and state trustees, and represents a significant component of lost-use values arising from the Gulf oil spill. Given this focus, we did not measure other lost-use values such as lost consumer surplus from seafood consumption. We also did not measure private losses such as property value impacts[1] or losses to commercial fishing enterprises, hotels and other businesses. Nonuse losses to the public were measured in a separate assessment (Bishop et al. 2017).

To monetize the lost interim recreational use value of the spill, we developed a comprehensive strategy involving primary data collection and demand modeling grounded in neoclassical welfare economics. Our damage assessment employed two sources of data that are unprecedented in their scale and scope for a study of recreation value: 1) infield surveys of recreational activity on site, comprised of aerial photographs and ground counts and interviews; and 2) telephone surveys of adults in the lower 48 states. The infield surveys are comprised of 129,000 in-person interviews, 35,000 on-site counts and nearly 500,000 aerial photographs, conducted for three years beginning May 2010. They are the basis for estimating the number of recreation user days in the Gulf coast area by year, month and region, from May 2010 through May 2013, that is, during and after the spill impact. This permits estimation of lost user days due to the spill by year, month and area (Tourangeau et al. 2017). The phone surveys, based on samples of adults in the lower 48 states, include 244,000 mail survey screeners and 43,000 telephone interviews. The surveys gather information on recreational trips to the Gulf area after the spill impact has passed. They form the basis for estimating the shoreline recreational demand model for sites in the greater Gulf of Mexico. Damages are calculated as the product of lost user days – estimated using the infield surveys – and the value of a lost user day – estimated using the phone survey data.

The disruption to recreation generated by the DWH spill led to a variety of challenges that were addressed in our approach for assessing losses. Recreation sites along the entire Gulf of Mexico faced potential impacts that lasted several months or longer, and many of these sites were popular destinations for visitors from throughout the U.S. The standard travel cost model focusing only on day trips from the local region (e.g., Parsons et al. 1999; Lew and Larson 2008) would have been inadequate for assessing the impact of a disruption of this scale. Faced with the challenges of creating a national demand model

[1] A concern with incorporating declines in property values in our damage assessment is that these losses would in part capture recreation losses which should be capitalized into the value of the housing stock. Thus, simply adding recreation and property value losses together would represent a form of double counting.

with a large geographical area of sites, we developed a series of innovations that should have lasting significance for recreational demand modeling. We estimate a model of shoreline recreation trips to the Gulf Coast region from the lower 48 states, combining single and multiple-day trips, calculating travel costs that incorporate detailed information on driving costs, flying costs and transportation mode choice, and using alternative-specific constants (ASCs) to control for site characteristics. The extent of the market and scope of impacted sites make the model unique. In addition, the demand model is estimated using data collected after the recreational impacts of the spill had dissipated, and the spill-induced demand shift is determined by calibrating the model to the estimates of lost user days from the infield counts. This approach improves upon Hausman, Leonard, and McFadden (1995), where declines in recreation were based solely on survey responses, and Stratus Consulting (2010), where declines in recreation were based on survey responses to estimate a percentage decline in recreation that was then applied to total counts from onsite surveys.

Past oil spills have played an important role in the field of environmental valuation. Efforts to assess nonuse values lost from the 1989 Exxon Valdez disaster (Carson et al. 1992, 2003), for example, led to an enduring debate about the measurement of those values (Portney 1994; Diamond and Hausman 1994; Kling, Phaneuf, and Zhao 2012). The Exxon-funded study assessing use value losses from the Valdez spill (Hausman, Leonard, and McFadden 1995) was an early application of a combined site choice and participation model to recreational trip data. The assessment of recreational use losses after the 2007 Cosco Busan oil spill in San Francisco Bay (Stratus Consulting 2010) was one of the first to combine telephone surveys with onsite counts to assess losses from a decline in environmental quality.

The DWH spill, affecting popular beaches and fishing sites throughout the Gulf of Mexico, has likewise attracted considerable attention from resource economists. Several recently published studies have explored spill-related welfare losses for targeted groups of individuals (Alvarez et al. 2014; Whitehead et al. 2016; Glasgow and Train 2017). Alvarez et al. (2014) use National Marine Fisheries Service interviews with anglers at sites throughout the Gulf Coast region to assess recreational fishing losses. They combine several years of pre-spill data with data gathered during the spill year and estimate welfare losses using a site- and season-specific spill dummy defined using data on spill-related federal fisheries closures. The authors estimate total angler losses of $80 million (Alvarez et al. 2016).[2] Whitehead et al. (2018) survey internet panel members residing in 13 states to estimate a single-site travel cost model of shoreline recreation trips to twelve counties in Northwest Florida.[3] Respondents were asked about canceled trips to these counties due to the spill, which the authors use to estimate a demand model under spill conditions. They estimate total losses of $207 million due to canceled trips to Northwest Florida. Finally, Glasgow and Train (2018) use data from a 2011 internet survey to estimate a model of overnight trips to coastal sites throughout the U.S. They develop welfare losses associated with various illustrative Gulf spill scenarios, but none of these scenarios is consistent with the DWH spill.[4] In contrast, our paper presents the first comprehensive estimate of shoreline recreation welfare losses associated with the DWH spill.

The remainder of the paper is organized as follows. The next section briefly discusses the unique legal context that framed our damage assessment, and Section III describes the survey data that support modeling Gulf coast recreation, with particular attention to construction of travel costs in the context of a national model of recreation. Section IV describes the shoreline model that characterizes recreation under non-spill conditions, as well as the procedures used to calibrate the shoreline model to reflect spill conditions so that the value per lost user day can be calculated. This is followed by a discussion of estimation results and a presentation of sensitivity analyses. The paper concludes by placing the model in the context of the shoreline recreation damage assessment and the DWH case, as well as a discussion of some of the lessons learned in conducting NRDA when recreation damages are national in scope.

[2] In a comment, Train (2016) criticizes Alvarez et al. (2014) for not using sampling weights and for using approximations when calculating welfare losses, among other concerns.

[3] The authors estimate that these 13 states comprise nearly 90 percent of all trips to their study site.

[4] One scenario uses estimated trip reductions borrowed from Trustee technical memoranda (Tourangeau and English 2015), but these Trustee estimates incorporate all shoreline recreation trips, while the Glasgow and Train data exclude trips taken by residents of Texas, Louisiana, Arkansas, Mississippi, Alabama, Florida, Georgia, and Tennessee.

II. Legal Context

Our recreation damage assessment was developed in support of a legal claim against BP and other defendants. Working on the public's behalf, it was essential that we capture recreational damages from the spill as accurately as possible, while recognizing that in certain instances it can be difficult or costly to do so. The starting point for our assessment was to collect the best possible recreational data. To measure lost user days, our preferred approach was to count losses onsite, and we therefore conducted infield surveys covering the great majority of Gulf coast recreation sites in Louisiana, Mississippi, Alabama and the Gulf side of Florida. Some areas where impacts were likely to be less severe, such as Texas and western Louisiana, were excluded. Certain activities were determined to be less affected or quite small in scale, such as hunting, scuba diving, and river cruises, and these activities were excluded from the assessment. To measure the value of lost user days, our preferred approach was to use a nationally representative sample of recreation visitors and non-visitors to avoid any endogenous sampling issues that can arise with on-site samples and any selection issues that can arise with visitor-only samples. Thus, we collected detailed information about recreational trips from a phone survey of the general U.S. population after the spill effects had dissipated (to measure baseline behavior un-confounded by the spill). Thus, our two-part strategy leverages the strengths of the two data sources. Combined, these data paint a vivid picture of coastal recreation in the Gulf region that is unprecedented in its scale and thoroughness.

We developed a recreation demand model to quantify the value of a lost user day, and the legal context influenced modeling decisions in ways that may not arise in a typical academic study. Given that courts depend on extensive peer review and acceptance in the field to help evaluate whether methods are reliable and valid, we had a strong preference for employing methods that were not only published in the academic literature, but widely used and well established. We were also cognizant of the fact that we would have to explain our model to a judge who might find advanced economic reasoning and methods challenging. We therefore placed significant value on simplicity and transparency in our analysis.

Much of the recent recreation demand literature has focused on microeconometric methods that account for preference heterogeneity, dynamics, and endogenous choice set formation (Moeltner and von

Haefen 2011; Provencher and Bishop 1997; von Haefen 2008). Although this literature has demonstrated that incorporating these methods can capture intuitive data features and can improve overall statistical fit (Moeltner and Englin 2004; Thiene, Swait and Scarpa 2017), there is surprisingly little evidence that they significantly and systematically change estimated economic values.[5] Moreover, these methods are often difficult to interpret behaviorally, implement in practice, and can result in models with poor in-sample predictions of demand at individual sites (Klaiber and von Haefen 2018). All of these properties undermine our goals of simplicity and transparency and could erode a judge's confidence in our estimate of recreation damages. As we describe in later sections, we therefore employ modeling techniques that are well-established in the literature when possible. For example, we employ the repeated discrete choice framework to model participation and site choice (Morey et al. 1993, Morey 1999) and assume a two-level nested logit specification in estimation. We also employ the standard assumption that the value of travel time is one-third the hourly income (Cesario 1976). However, we encountered a number of situations where consensus practices have not arisen in the literature. The most important of these involved the treatment of multi-day trips and methods for specifying a travel cost model that is national in scale, including accounting for airline travel. These were essential to valuing recreation over a large geographic area that attracts visitors from throughout the country. Here we used our collective expert judgement to develop innovative methods which were suitable for the recreation damage assessment and should be applicable to valuing recreation in other circumstances as well. Subsequent sections describe these methods in detail.

III. Sampling Strategy and Data Sources

The development of the shoreline model with such a broad scope necessitated a large-scale data collection effort because publically available recreation data were inadequate for reasons such as spatial and temporal coverage gaps and inconsistencies. As mentioned above, lost user days were measured by on-

[5] For example, von Haefen (2008) finds quantitatively similar welfare measures across models with and without endogenous choice sets, as do Thiene, Swait and Scarpa (2017).

site counts of visitors during and after spill conditions, and the value of lost user days was estimated using a model of shoreline recreation estimated using data from the general U.S. population. In this section, we outline the data sources used to estimate our shoreline model.

A. Recreational Trip Data

Data were collected on trips to coastal areas in Texas, Louisiana, Mississippi, Alabama, Florida, and Georgia, known hereafter as the "six-state" area. The data were collected in general population surveys conducted in 2012-13, which distinguished between "local" and "national" subpopulations. While the survey instruments differed slightly for the two subpopulations, as described below, the basic structure of the surveys and the data collected were comparable. For both subpopulations, the surveys collected information from people who did and did not make trips to the six-state area. The local component sampled residents of Louisiana, Mississippi, Alabama, and Florida, as well as portions of Texas and Georgia that were roughly within a half-day drive of the Gulf.[6] The national component sampled residents throughout the lower 48 states that were not covered by the local component. Figure 1 illustrates the location of the survey respondents for the local and national subpopulations, respectively.

For both subpopulations, a short mail screener was used to collect phone numbers and identify past Gulf Coast recreators.[7] All respondents with a valid phone number were eligible to be recruited into the subsequent telephone survey, although respondents who reported taking recreational trips to the Gulf were sampled more intensively to improve sampling efficiency. This oversampling was addressed in the sampling weights, described below. The telephone survey collected detailed trip and demographic data.

---

[6] The "local" component used a dual frame for sampling. One frame consisted of the general population sampled via the United States Postal Service (USPS) delivery sequence file, and the second frame consisted of registered boaters sampled via state boater registration lists. The dual frame allowed over-sampling of likely boating households to improve efficiency of sampling for a boating model, which is described in English et al. (2018b). The "national" component's sample was drawn from residential addresses in the USPS delivery sequence file.

[7] Both the local and national versions of the mail and telephone survey instruments went through extensive development and testing by the NOAA team members, including focus groups and cognitive interviews, as well testing of the computer assisted telephone interview (CATI) tools used to administer the telephone interviews. Phone interviewers were also required to undergo extensive training and certification, including training on the pronunciation and location of relevant geographic locations within the region. The key technical memos describing survey and model estimation procedures are given in Appendix A.

As indicated in Figure 2, telephone interviews in the national sample were conducted across two waves with one starting in July 2012 for trips since January 2012 and another starting in January 2013 for trips since July 2012. The local telephone interviews were conducted in twelve monthly waves from June 2012 to May 2013. Within a wave, a complete record of trips taken during the corresponding reporting period was obtained, and a new sample of respondents was drawn each wave. The local sample had more and shorter waves to reduce the cognitive burden of recalling trips that were expected to be more numerous. The average reporting period was 81 days for the local survey and about 7.5 months for the national survey.

The survey began by asking respondents if they had taken a coastal recreational trip in the last two years to the relevant region, either the six-state region for the local survey or anywhere in the continental U.S. for the national survey. If the date of the trip fell within the respondent's reporting period, he or she was asked about additional trip details (e.g., trip location, party size, duration and, for multi-day trips, main purpose). The survey then worked backward in time, collecting detailed information for up to three of the most recent trips as long as they fell within the reporting period. For any additional trips within the reporting period, destinations and trip counts to individual sites were collected as well as the number that were multi-day trips. The surveys also obtained detailed demographic information from respondents, including for those who did not take coastal recreation trips. In the case of the local subsample, boating trips were identified separately from shoreline trips that did not involve boating. Only the shoreline trips are discussed here.

The survey data were weighted in several different ways to account for the sampling strata and reflect the target populations of the national and local surveys, respectively.[8] First, the weights accounted for sample selection probabilities in the initial sample of addresses for the mail screeners. Sample selection probabilities varied by state throughout most of the country, and by county in areas near the Gulf Coast, primarily to avoid too high a concentration of respondents from densely populated areas and

[8] Table A.1 of Appendix A provides the location of technical documentation for the weights.

to oversample areas with a relatively high rate of participation in Gulf Coast recreation. Second, the weights corrected for nonresponse by selected geographic areas, with completed surveys in each area weighted up to reflect the full sample in each area. Third, weights corrected for the purposeful oversampling of respondents to the mail screener who said they had participated in coastal recreation during the previous year. Fourth, weights were post-stratified to match the number of households in selected geographic areas, which were groups of states in the national survey region and individual states in the local survey region. Fifth, the weights corrected for the subsampling of a single individual within a household during the telephone portion of the survey. Sixth, observations in selected groups were reweighted so that the relative size of each group in the sample matched the relative size of each group in the U.S. Census. This process is called "raking," and the selected groups involved various cross classifications of education level, race, age, sex, and geographic areas (Battaglia et al. 2013; an overview of raking methods is also described in Leggett, 2015). Seventh, using a smaller number of cross classifications to group similar observations, large weights were trimmed by group, and groups were again reweighted to match U.S. Census totals.

The previous steps led to a set of weights that made the sample of respondents representative of the full population. However, each respondent was only asked about their trips for a portion of the year. To create a data set of recreation trips that is representative of annual recreation activity for the full population, monthly weights were created. Specifically, for each month a data set was created using all respondents whose reporting period included that month. Thus, recreation activity in each month was represented using a subset of respondents. Each monthly subset of respondents was then raked to match the full population using raking categories similar to those used in the sixth step described above. The result was twelve data sets with weights, each representing recreation trips by the full population in a given month. Finally, in addition to these "base weights," an additional 120 sets of "replicate weights" for the national and local subsamples was created for use in variance estimation.[9]

Table 1 summarizes the survey sampling outcomes for the local and national subsamples. Across all waves the weighted response rate for the local and national mail components was 44% and 46%, respectively. The weighted response rate of the local and national telephone survey components was 29% and 26%, respectively. The composite response rates were 13% for the local and 12% for the national.

Data were obtained for 41,708 respondents and contain 27,717 trips to the relevant six-state area, with 25,950 of these originating from the local subsample and 1,767 from the national. About 93% of these were used in the demand model; trips with a main purpose other than recreation were excluded, as were those missing important trip details or that could not be precisely geocoded (about 4% overall). Using the monthly sampling weights, we found that about 17% of all trips spanned multiple days, and the mean trip duration was 1.8 days of which 1.5 were recreation days. Most trips originated within 100 miles of the coasts (83%), while 8% originated from over 1000 miles away. For the spill assessment areas from Louisiana to the Florida Keys, survey results suggest about 13% of the trips and 38% of the user days came from the national survey subpopulation.

Beyond the detailed trip data, the survey yielded socio-demographic characteristics for the survey respondent, including age, gender, education and employment status. An important variable in constructing the travel cost of a shoreline recreation trip is household income. A large percentage of respondents – 67 percent – provided an exact amount, and another 23 percent gave bounds (e.g., between $50,000 and $75,000, less than $50,000, greater than $75,000, etc.). When an exact amount was not provided, household income was imputed with a multiple imputation algorithm that closely followed

[9] The use of twelve sets of monthly weights complicates the calculations but is conceptually simple. Data from each respondent is effectively broken out into several distinct observations, one for each month in the respondent's reporting period. Throughout this text, equations using weighted sums omit the summation over months to simplify notation, and one could think of each sum over individuals as a sum over both individuals and their monthly observations. Weighted statistics also refer to the weighted sum over respondents' monthly observations.

Schenker et al. (2006) and used bounds when available. Five sets of income imputations were generated and used in estimation. The resulting weighted average annual income for the sample was $59,383 (2012 dollars). For the most part, other demographic variables had few missing values (< 1 percent).[10] A simple hot decking procedure was used for imputation in these cases (Leggett 2015).

For 98% of the reported trips, destinations were identified with enough specificity to include them in the model. This could include an exact beach, or a nearby town or city. Trip destinations were aggregated into 83 sites that spanned roughly 2,300 miles of coastline from Texas to Georgia. The spill area of the coast line was defined as any part of the Gulf coast, from western Louisiana through the Florida Keys, where visitation rates might plausibly be reduced because of the spill. Of these 83 sites, the spill area contained 54 sites. In our spill scenarios, we divided the spill area into two regions – the North Gulf from Louisiana to Apalachicola, FL (26 sites) and the Florida Peninsula from Apalachicola through the Florida Keys (28 sites). Figure 3 illustrates the extent of coastline covered by these sites.

B. Travel Costs

The cost of traveling to a recreation site is a key component of any travel cost model. For all sample locations one needs exogenous estimates of the cost of getting to all the sites in the choice set, regardless of whether any particular individual visits those sites, or any sites at all. Past researchers typically assume that individuals exclusively drive to recreation sites, which may be reasonable for local sites attracting only regional visitors. However, the assumption would not be correct for the six-state region given our data on trips from throughout the country. Our approach to constructing exogenous travel costs recognizes that some share of visitors fly to their recreation destination. We calculate travel cost as a weighted average of driving and flying costs, where the weights are based on the observed share of respondents who fly versus drive in the telephone survey. Specifically, let C_{ij} represent the roundtrip

[10] Notable exceptions include 2% missing gender in the local subsample and 2.3% missing age in national subsample.

cost to individual *i* of traveling to site *j* and let ρ_{ij} represent the probability of flying to site *j*. Travel cost was calculated as[11]

(1) $$C_{ij} = (1 - \rho_{ij})C_{ij}^D + \rho_{ij}C_{ij}^F$$

where C_{ij}^D is the roundtrip cost of driving from origin *i* to destination *j* and C_{ij}^F denotes the corresponding roundtrip flying cost. The driving cost was calculated as a function of the roundtrip driving distance (d_{ij}), roundtrip driving time (t_{ij}), hotel nights required (h_{ij}), roundtrip cost of tolls (f_{ij}), and party size (n), as follows:

(2) $$C_{ij}^D = [p_d d_{ij} + p_h h_{ij} + f_{ij}]/n + p_t t_{ij}.$$

Driving distances, driving times, and tolls were calculated using *PC*Miler 27*. The per mile out-of-pocket driving cost, p_d, consists of gasoline, per-mile maintenance, and per-mile depreciation for an average passenger vehicle.[12] The per-mile gasoline cost for respondent *i* is the average fuel cost associated with their region of residence (U.S. EIA 2013) divided by the nationwide average fleet fuel economy (U.S. BTS 2013). On average, the per mile out-of-pocket cost was roughly $0.25 in 2012 dollars. The number of hotel nights was calculated as total one-way driving time divided by 12, rounded down to the nearest integer and doubled to reflect roundtrip costs. The per-night cost of hotels, p_h, is $105 (AHLA, 2013). The above costs per vehicle were divided by the average trip party size, $n = 2.7$. The final component in the driving cost reflects the cost of travel time, where p_t was measured as one-third of a respondent's household income divided by 2080 hours worked per year.

To compute flying costs for each respondent-site pair, C_{ij}^F, we identified the four airports nearest to the respondent's trip origin (i.e., residence) and the four airports nearest to the trip destination (site).

---

[11] Each individual's travel costs were tailored to reflect relevant airfare and gasoline prices in their reporting period. For notational simplicity, however, we exclude time subscripts in this section.

[12] Per-mile depreciation was computed using data on an average passenger vehicle from the American Automobile Association (AAA) (see, for example, AAA 2012). Specifically, we take the difference in depreciation between driving 5,000 miles more and 5,000 miles less than AAA's baseline scenario of 15,000 annual miles, and then divide by 10,000 to get a per-mile depreciation of $0.0468. The per-mile maintenance was also derived from AAA for an average passenger vehicle and includes maintenance and tire costs.

Only airports with at least 100,000 annual enplanements were considered, and each set of origin and destination airports was required to have at least one airport with at least 2 million annual enplanements, i.e. a "medium" sized airport as defined by the U.S. Federal Aviation Administration. Roughly 200 airports were used as origin/destination airports. We calculated flying costs for each of the 16 origin/destination pairs, and then to calculate C_{ij}^F, we assumed that respondents chose the least-cost option. Flying costs included five components: (a) round trip driving costs from the respondent's residence to the origin airport using equation (2); (b) parking costs at the origin airport for the duration of the trip; (c) round-trip airfare from the origin to the destination airport plus expected flying time valued as above; (d) round trip driving costs from the destination airport to the site using equation (2); and (e) expected cost of car rental (a per-person weighted average of rental costs by airport size accounting for the share of people from the national survey data that rented cars and the party size). Expected flying time included time spent at airports pre-departure and post-arrival (two hours), in-flight time (based on routes using data from OAG Aviation Solutions Schedules Database), and any layover time for routes (using median layover times by airport size from Sabre Airline Solutions). The out-of-pocket cost for an airline ticket was computed using the 30th percentile fare for round-trip tickets from the Airline Origin and Destination Survey (DB1B), a 10% sample of all airline tickets collected by the Office of Airline Information and the Bureau of Transportation Statistics. Fares and travel times vary quarterly. For all airlines other than Southwest and JetBlue, a $50 baggage fee was added to the fare. Since roughly 40% of fares are for business which are typically higher fares, the 30th percentile fare from the DB1B data can be thought of as representing the median non-business fare. Sensitivity analyses revealed that the fares and flying costs were relatively flat across a broad range of percentiles around the 30th.

The final component of the expected travel cost is the probability of flying, ρ_{ij}. This probability was estimated using data on reported travel mode choices from the national component of the telephone survey. The proportion of trips where respondents chose to fly was calculated for each of several cells, with cells defined by different combinations of one-way driving distance, household income and family size. For parsimony, cells with small sample sizes and similar means were merged. The resulting probabilities for flying are reported in Table 2. In general, the results in Table 2 suggest that the probability of flying rises with driving distance and income and declines with family size. For driving distances under 500 miles, respondents almost always drive, and for driving distances greater than 1,500 miles, respondents fly roughly 85 percent of the time. Figure 4 illustrates the resulting driving, flying and expected travel costs for the average respondent as a function of travel distance. The figure shows that driving travel costs increase monotonically with the individual's distance from the recreation site, ranging from $57 at a 100-mile travel distance to nearly $1,800 at a travel distance of 3,000 miles. Flying costs, on the other hand, rise more slowly with travel distance. With the probability of flying increasing with travel distance, expected travel costs initially match driving travel costs closely, but gradually deviate towards flying travel costs, peaking at roughly $800 when travel distance reaches 3,000 miles.

IV. The Models

The shoreline model was developed to characterize trip taking behavior to the Gulf of Mexico by individuals living in the contiguous United States.[13] Using behavioral data from the general population survey described in Section III above, the model provides the foundation for estimating both the baseline (non-spill) pattern of trips to the region and individuals' responses to changing environmental conditions. In order to characterize recreation demand under spill conditions, the shoreline model is calibrated to reproduce observed reductions in trips to the Gulf during the spill period. Together, the estimated and calibrated shoreline models provide the components needed to construct the value per lost user day stemming from the Deepwater Horizon oil spill. In the remainder of this section, we provide a description of the shoreline model, the calibration procedure used to reflect changes in recreation demand under spill conditions, the computation of the value per lost user day, and the approach used to estimate standard errors for the resulting damage assessments.

[13] The destinations explicitly modelled include the entire U.S. Gulf coast, as well as the Atlantic coasts of Florida and Georgia.

A. The Shoreline Model

The shoreline model is an adaptation of the repeated discrete choice model based on the Random Utility Maximization (RUM) hypothesis (McFadden 1974, 1978, 1981). RUM models assume that individuals, facing a well-defined choice set, select the alternative yielding the highest level of utility. Thus, let U_{ik} denote the conditional utility received by individual i in choosing alternative k ($k = 1,..., J$), so that the individual chooses alternative j (denoted by $y_{ij} = 1$) if $U_{ij} > U_{ik}$ $\forall k \neq j$; i.e.,

(3) $$y_{ij} = \begin{cases} 1 & U_{ij} > U_{ik} \; \forall \; k \neq j \\ 0 & \text{otherwise.} \end{cases}$$

The conditional utilities themselves can depend upon characteristics of both the individual and the available alternatives.

Analysts modeling observed choices in a given setting will not observe all of the factors influencing individual decisions. Instead, they characterize the conditional utilities as a function $V_{ij} = V(X_{ij}; \tau)$ of observable individual/alternative specific attributes (X_{ij}), where τ denotes a vector of parameters to be estimated, and a residual term ϵ_{ij}, implicitly defined as $\epsilon_{ij} = U_{ij} - V_{ij}$, capturing idiosyncratic factors influencing the utility individual i derives from choosing alternative j. Thus, $U_{ij} = V_{ij} + \epsilon_{ij}$. Given assumptions regarding the distribution of the vector $\epsilon_{i\cdot} = (\epsilon_{i1}, \dots, \epsilon_{iJ})$, the analyst can then identify the probability that a specific choice will be made. In general,

(4) $$P_{ij} = Pr(y_{ij}|X_{i\cdot}) = Pr(U_{ij} > U_{ik} \; \forall \; k \neq j) = Pr(\epsilon_{ik} - \epsilon_{ij} < V_{ij} - V_{ik} \; \forall \; k \neq j)$$

where $X_{i\cdot} = (X_{i1}, \dots, X_{iJ})$. These probabilities can, in turn, be used to specify the appropriate log-likelihood function used in maximum likelihood estimation of the model parameters.

The repeated discrete choice model was developed by Morey, Rowe and Watson (1993) in the context of recreation demand. There are two key modifications to the basic RUM model. First, a "stay-at-home" option is added to the choice set, so that j now runs from 0 to J, with $j = 0$ denoting the "stay-at-home" option and $j = 1, ..., J$ denoting the available sites in the choice set. Second, instead of making a single choice, individuals are modeled as choosing from among the $J + 1$ alternatives over a series of T_i

choice occasions. The number of choice occasions varies over individuals in our analysis to the extent that they have different reporting periods for trips. Assuming that the vector $\epsilon_{i.} = (\epsilon_{i0}, \dots, \epsilon_{iJ})$ is identically distributed across choice occasions, individual i's contribution to the log-likelihood function becomes:

$$(5) \qquad \mathcal{L}_i = \sum_{j=1}^{J} n_{ij} ln(P_{ij}) + (T_i - \sum_{j=1}^{J} n_{ij}) ln(P_{i0})$$

where n_{ij} denotes individual i's weighted trips summed over her reporting period, T_i is choice occasions (proportional to the weighted sum of days in her reporting period), and P_{i0} is the probability of not taking a trip on a given choice occasion.[14] Completing the model requires specifications for the V_{ij}'s and distributional assumptions for the error vectors (i.e., the $\epsilon_{i.}$'s).

In the shoreline model, the V_{ij}'s are assumed to have the following structure:

$$(6) \qquad V_{ij} = \begin{cases} \delta Z_i & j = 0 \\ \alpha_j + \beta C_{ij} & j = 1, \dots, J \end{cases}$$

where the Z_i denote individual-specific characteristics impacting the individual's propensity to choose the "stay-at-home" option ($j = 0$) and C_{ij} denotes the roundtrip travel cost for individual i in choosing to visit alternative j on a given choice occasion. The α_j parameters, commonly referred to as alternative-specific constants (ASCs) (e.g., Murdock 2006), capture all site-specific attributes (including environmental conditions), while β captures the impact of travel cost on the propensity to visit a site.

The final step in completing the shoreline model is to choose the distribution for the error vector $\epsilon_{i \cdot t}$. A commonly used assumption is that the $\epsilon_{i \cdot t}$ are drawn from a GEV distribution implying a two-level nesting structure that groups the trip alternatives (i.e., $j = 1,.., J$) into a single nest and the stay-at-home option into a singleton nest. The implied choice probabilities then take the form:

[14] In particular, we constructed n_{ij} by taking the product of individual i's trips to site j in each month and corresponding monthly sampling weight and then summing across the reporting period. Similarly, for T_i, we took the product of the number of days in each month and corresponding monthly weight and summed across the reporting period. Since both n_{ij} and T_i are constructed with monthly weights, the log-likelihood in turn is a function of monthly weights, which is appropriate given our sampling approach (described in Section III). Solon, Haider and Wooldridge (2015) provide a full discussion of when weights should be used in estimation.

(7)
$$P_{ij} = \begin{cases} \dfrac{\exp(V_{i0})}{\exp(V_{i0})+\left[\sum_{k=1}^{J}\exp\left(\frac{V_{ik}}{\theta}\right)\right]^{\theta}} & j = 0 \\ \dfrac{\exp\left(\frac{V_{ij}}{\theta}\right)}{\sum_{k=1}^{J}\exp\left(\frac{V_{ik}}{\theta}\right)} \cdot \dfrac{\left[\sum_{k=1}^{J}\exp\left(\frac{V_{ik}}{\theta}\right)\right]^{\theta}}{\exp(V_{i0})+\left[\sum_{k=1}^{J}\exp\left(\frac{V_{ik}}{\theta}\right)\right]^{\theta}} & j = 1, \dots, J \end{cases}$$
$$= \begin{cases} 1 - P_{i,Trip} & j = 0 \\ P_{ij|Trip} \cdot P_{i,Trip} & j = 1, \dots, J \end{cases}$$

where θ is often referred to in the literature as the dissimilarity coefficient,

(8)
$$P_{i|Trip} = \frac{\left[\sum_{k=1}^{J}\exp\left(\frac{V_{ik}}{\theta}\right)\right]^{\theta}}{\exp(V_{i0})+\left[\sum_{k=1}^{J}\exp\left(\frac{V_{ik}}{\theta}\right)\right]^{\theta}}$$

denotes the probability that individual *i* chooses to take a trip on a given choice occasion and

(9)
$$P_{ij|Trip} = \frac{\exp\left(\frac{V_{ij}}{\theta}\right)}{\sum_{k=1}^{J}\exp\left(\frac{V_{ik}}{\theta}\right)}$$

denotes the probability that they choose to visit site j ($j = 1, ..., J$) conditional on having chosen to take a trip. Effectively, the model implies a system of demand equations with population expected trips to site j equal to:

(10)
$$q_j = \sum_{i=1}^{N} T_i \cdot \frac{\exp\left(\frac{V_{ij}}{\theta}\right)}{\sum_{k=1}^{J}\exp\left(\frac{V_{ik}}{\theta}\right)} \cdot \frac{\left[\sum_{k=1}^{J}\exp\left(\frac{V_{ik}}{\theta}\right)\right]^{\theta}}{\exp(V_{i0})+\left[\sum_{k=1}^{J}\exp\left(\frac{V_{ik}}{\theta}\right)\right]^{\theta}}.$$

The robustness of this model structure is tested in one of the sensitivities we consider below, which extends the shoreline model to use a three-level nesting structure. The three-level structure allows for the possibility that sites are further distinguished into sub-groups of sites such that trip substitution in response to changing conditions is greater within sub-groups than across sub-groups.

The shoreline model is unique in its scope, covering trip taking behavior to the Gulf of Mexico from throughout the contiguous U.S., whereas most studies in the literature focus on regional day trips only. This required two innovations relative to the standard repeated discrete choice model. First, the travel distances involved for a large portion of the country required consideration of alternative travel modes, specifically air travel versus driving, in the construction of travel costs.[15] Our construction of expected travel costs, as described in Section III above, incorporates unique features of airline travel, including individually tailored airport options, the airfares and travel times implied by competing routes for the relevant airports, and any driving costs accompanying air travel.

The second innovation in the shoreline model is the inclusion of both single and multiple-day trips in one model. The literature handles single and multiple-day trips in different ways. Morey, Rowe and Watson (1993) and Hausman, Leonard and McFadden (1995) combined them in a single discrete choice model. More recent literature tends to model these trips separately. The rationale for our approach is detailed in English et al. (2018a) and draws on earlier work by McConnell (1992). The key to understanding the approach is to note that the travel cost, i.e., the exogenous price of getting to a site, is the same regardless of how long one stays at a site. When trip lengths are not exogenously imposed, the consumer endogenously chooses total trips[16] and trip length in response to travel cost. The length of a trip, and any expenditures associated with the number of days on site, are chosen optimally at any given travel cost, which implies a total trip demand function that depends on travel cost and other exogenous factors. Using Roy's identity, this demand function can be consistently linked back to preferences and welfare. A similar link cannot be established for demand models of only single day or multi-day trips. What this result implies is that although consumers may choose to take longer trips as travel cost increases, this behavioral response does not affect welfare at the margin given standard conclusions of the envelope theorem for optimized choice variables. The only marginal effect on welfare as travel cost increases is the change in expenditures represented by the product of total trips and the change in travel cost. This holds regardless of any distinctions that could be drawn between trips whose endogenously chosen length, type of activity, or onsite expenses vary according to an individual's optimizing choices (McConnell 1992).

---

[15] In their model of Alaska residents, Hausman, Leonard and McFadden (1995) also allow for travel costs to vary by travel mode in the context of the Exxon Valdez Spill, though commercial air travel was used in only two percent of trips.
[16] Or in the case of the repeated discrete choice model, whether to visit a site.

B. Calibration

The shoreline model is estimated using trip data from the recreation demand surveys described in Section III. The data cover the post-spill period after the effects of the spill on shoreline recreation had largely dissipated based on onsite counts of visitation (Tourangeau et al. 2017). As such, it characterizes preferences for coastal recreation in the Gulf of Mexico under baseline (or "but for the spill") conditions. In order to assess the welfare impact of the spill itself, we need a similar characterization of preferences under spill conditions; i.e., we need to understand how demand shifted in response to the spill. Traditionally, when one lacks external evidence on this demand shift, one would need to include site characteristics (e.g., water clarity, numbers of tar balls, the presence or absence of oil slicks, etc.) in a recreation demand model, estimate the marginal willingness to pay for these site characteristics, and construct the welfare impact of the spill based on how it shifted each of the relevant site characteristics.[17] There are three significant drawbacks of this traditional approach. First, identifying the impact of spill-related factors on Gulf recreation would require data on these factors and on recreation demand under both spill and non-spill conditions for all sites. Over 80 percent of lost user days occurred in the first five months following the spill, during the spring and summer of 2010. Given the time it takes to field a well-designed recreation demand survey, the available variation in spill conditions would be seriously limited. Second, in the period immediately following the spill, predictions of the spread of oil varied widely, with some concern that it could even cause damage along the Atlantic coast. The impact of the spill on Gulf coast recreation, even in areas where the oil never actually made it to local beaches, is apparent from counts conducted on site (Tourangeau et al. 2017), evidence that individual perceptions of or uncertainty about conditions in the Gulf altered their recreation behavior. Quantifying these effects and their variation across individuals would be difficult. Third, as in any recreation demand model, there is the ever-present risk of omitted variables bias (Murdock, 2006).

[17] See, for example, Alvarez et al. (2104), who include measures of historic catch rates, popularity for each site, and the number of access points at each site, as well as dummy variables to capture seasonality in trip taking and exposure to the spill (determined by federal water closures).

The approach we employ overcomes these challenges by drawing on external information from the on-site counts (Tourangeau et al. 2017) to infer changes to the overall appeal of individual sites induced by the spill, capturing these impacts by changes in the site's alternative specific constants (i.e., the α_j's). In particular, the infield data provides information on the percentage reduction in trips to two broad groups of sites: (1) the North Gulf (from Grand Isle, Louisiana through Apalachicola) and (2) the Florida Peninsula (from Apalachicola through the Florida Keys). The spill impact is divided into two periods that vary regionally. The first period goes from June 2010 through January 2011, and includes both the North Gulf and the Florida Peninsula. The second period runs from February 2011 through November 2011 and includes only the North Gulf. We discuss the methods for calibrating the shoreline model to the first period; calibration in the second period is similar. The percentage reductions are denoted by ζ_g (*g=NG,FP)*[18]. The calibration exercise for the first period involves choosing group level adjustments, δ_g $(g = NG, FP)$, to the alternative specific constants, with the revised constants given by

$$(11) \qquad \alpha_j^1 = \begin{cases} \alpha_j + \delta_{NG} & j \in \mathcal{G}_{NG} \\ \alpha_j + \delta_{FP} & j \in \mathcal{G}_{FP} \\ \alpha_j & j \in \mathcal{G}_{EE} \end{cases}$$

where $\mathcal{G}_g$ $(g = NG, FP, EE)$ denotes the set of sites in the aggregate group of site *g* and *g=EE* denotes the set of modeled sites not in either the North Gulf or Florida Peninsula (i.e., everywhere else). The adjustment comes from a contraction mapping algorithm solved for δ_g $(g = NG, FP)$ such that

$$(12) \qquad \bar{P}_g^1 = (1 - \zeta_g)\bar{P}_g^0 \qquad g = NG, FP,$$

where

$$(13) \qquad \bar{P}_g^s = \sum_{i=1}^N T_i \sum_{j \in \mathcal{G}_g} P_{ij}^s, \qquad g = NG, FP; s = 0,1,$$

[18] The percent reductions are calculated as $\zeta_g = abs(\frac{d_g^1 - d_g^0}{d_g^0})$ where d_g^s is the level of user days in region *g* during *s = 0,1* for baseline and spill.

denotes the predicted number of trips to region g under conditions s, with s=1 denoting spill conditions and s=0 denoting non-spill conditions, with the trip predictions computed using the alternative specific constants (ASCs) from the shoreline model (α_j) for non-spill conditions (i.e., s=0) and using the calibrated ASCs (i.e., α_j^1) for spill conditions (i.e., s=1). With the calibrated model in hand, one can then compute the welfare impact of the spill using the standard log-sum formula.

C. Welfare Calculation

Given the estimated shoreline model, reflecting baseline conditions, and the calibrated model of Gulf recreation under spill conditions, welfare losses are calculated as the product of lost user days per period and the value of a lost user day. The value per lost trip in period t ($\tilde{\mathcal{V}}_t$) can be computed as:

$$(14) \qquad \tilde{\mathcal{V}}_t = \frac{\sum_{i=1}^{N} T_i \frac{1}{\beta}\left[ln\left(\exp(V_{i0})+\left[\sum_{k=1}^{J}\exp\left(\frac{V_{ik}+\delta_k}{\theta}\right)\right]^{\theta}\right)-ln\left(\exp(V_{i0})+\left[\sum_{k=1}^{J}\exp\left(\frac{V_{ik}}{\theta}\right)\right]^{\theta}\right)\right]}{\sum_{i=1}^{N} T_i\left[(P_{iNG}^0+P_{iFP}^0)-(P_{iNG}^1+P_{iFP}^1)\right]} .$$

The numerator is the usual log-sum formula for the welfare loss due to changes in the conditional site utility, where the δ_k are the spill-induced changes in the alternative-specific constants estimated in equations (11) through (13). The denominator is the change in trips to the North Gulf and Florida Peninsula, where:

$$(15) \qquad P_{ig}^s = \sum_{j\in\mathcal{G}_g} P_{ij}^s \qquad g = NG, FP; s = 0,1 .$$

Welfare in equation (14) is expressed in value per lost trip. To convert to value per user day ($\mathcal{V}_t$), we divide by the mean number of recreational days per trip; i.e., $\mathcal{V}_t = \tilde{\mathcal{V}}_t/\bar{r}$ where $\bar{r}$ (=1.7) denotes the average number of recreational user days per trip in the spill area. The damage calculation is completed by multiplying by the estimated number of lost user days Δ_t. That is, the welfare loss L_t in period t is given by

$$L_t = \Delta_t \cdot \mathcal{V}_t$$

where $\Delta_t = d^0 - d^1$, with d^0 and d^1 denotes the estimated user days in baseline and spill periods, respectively.

D. Precision

There are two key outputs from the shoreline model: (a) the estimated value per lost user day $\mathcal{V}_t$ and (b) the estimated total loss $L_t = \mathcal{V}_t \cdot \Delta_t$. Computing variances for these outputs requires taking into account three sources of uncertainty: (1) uncertainty in the estimated lost user days for period *t* based upon the infield counts; (2) uncertainty due to sampling variability in the valuation survey; and (3) uncertainty due to the imputation of income for those individuals in the valuation survey who did not provide household income or who provided only income bounds. Typical recreation studies do not account for all of these sources of variation, implying that reported variance estimates are biased downward. Details of the variance computations are provided in Appendix B. In brief, the first source of uncertainty is summarized by $\sigma_{\Delta t}^2 = Var[\Delta_t]$, which is constructed using a jackknife variance estimator. The other two sources of uncertainty affect the estimated value per lost user day ($\mathcal{V}_t$). The overall variance $\sigma_{\mathcal{V}t}^2$ is constructed using a combination of a jackknife variance estimation procedure (Rust 1985) and Rubin's multiple imputation method with ignorable nonresponse (Rubin and Schenker 1986; Steimetz and Brownstone 2005).

Treating the uncertainty stemming from the counts as independent of the uncertainty stemming from the valuation survey, we can compute the variance for the overall loss as:

(16) $$\widehat{\sigma}_{Lt}^2 = \widehat{\sigma}_{\mathcal{V}t}^2 \cdot \sigma_{\Delta t}^2 + \widehat{\sigma}_{\mathcal{V}t}^2 \cdot \Delta_t^2 + \sigma_{\Delta t}^2 \mathcal{V}_t^2.$$

The assumption that $\mathcal{V}_t$ and Δ_t are independent ignores the fact that the counts are used to calibrate the model to reflect the percentage change in user days by zone as a result of the spill, which is in turn used to compute the value per lost user day. However, the assumption of independence is a reasonable approximation given that sensitivity analysis indicates that the value per lost user day is relatively insensitive to the percentage change in trips used in the calibration.[19]

[19] As a check on the importance of the independence assumption, we implemented as a sensitivity check an alternative bootstrap estimator for $\widehat{\sigma}_{Lt}^2$ that allows for correlation between $\mathcal{V}_t$ and Δ_t. Details of the bootstrap procedure are provided in a technical memo available online. See Appendix A for the location of the "Precision Estimation" and all other technical reports.

V. Estimation Results

We estimated the parameters entering the shoreline valuation model in *GAUSS*.[20] Table 3 describes the variables used, and Table 4 reports maximum likelihood parameter estimates and associated *t*-statistics. Given our multiple imputation approach to addressing missing household income, the reported parameter estimates are the mean values across five imputations, and the *t*-statistics are generated with jackknifed variances analogous to those described in Section IV.D for the overall value of a lost user day.[21] For brevity, the 83 site constants are not reported.

As expected, our travel cost coefficient is negative and highly significant. Consistent with economic theory, the nested logit dissimilarity coefficient falls between zero and one and implies that recreation sites are closer substitutes with each other than with the no-trip alternative. Demographic interactions with the no-trip constant suggest that respondents who have lower incomes, reside in rural areas, have kids and are not employed full-time are less likely to take shoreline trips. Moreover, retired, white respondents who are either part-time employed, students or disabled have a higher probability of taking trips.

To assess the natural resource damages from the Deepwater Horizon oil spill we combine the estimated parameters and model structure with two spill scenarios that span the 18-month spill period. As summarized in Table 5, the first spill period involved a 45.5 percent reduction in trips[22] in the North Gulf and a 22.9 percent reduction in trips in the Florida Peninsula (see Figure 3 for affected sites) that arose during the initial eight months. Consistent with equations (11) through (13) in the previous section, the ASCs in both regions were adjusted to replicate these proportional trip reductions. In the second period spanning the later ten months, the spill impacts were spatially limited to the North Gulf only and involved

[20] A replication exercise in *Matlab* produced parameter and welfare estimates that matched those reported in this section to five digits.
[21] Details of the jackknife procedures are provided on page 6 of the technical memo entitled Precision Estimation as listed Appendix A. While the advantage of the jackknifed procedure used is that it accounts for uncertainty stemming from *both* survey sampling and income imputation sources, as a practical matter, the resulting *t*-statistics were very similar to those obtained directly from maximum likelihood estimation when the standard errors were clustered at the individual level.
[22] We assume that the percent reduction in user days and trips are equal.

a 10.1 percent reduction in trips. Therefore, for the second scenario, only the ASCs in the North Gulf were adjusted. For both periods, the economic damages per lost user day were monetized using equation (14).

We find a lost user day value of $37.23 for the first spill period and $40.41 (2015 dollars) for the second period.[23] By construction, the first period scenario assumes that 67.5 percent of baseline trips to coastal areas from Louisiana to the Florida Keys (i.e., the first period spill area) still occurred, although the utility of these trips was diminished. Moreover, of the 32.5 percent of remaining trips, the model predicts that about 39 percent of these trips were substituted to coastal areas in Texas, the Atlantic side of Florida and Georgia, and the remaining 61 percent were lost to the six-state area. Similarly, for the second spill period scenario, 89.9 percent of baseline trips to coastal areas in the North Gulf (i.e., Louisiana to Apalachicola, FL) still occurred with the spill and were diminished. Of the remaining 10.1 percent of trips, 32.5 percent were substituted to Texas, the Florida Peninsula, the Atlantic side of Florida or Georgia, whereas the balance were lost to the six-state area. These different percentages of diminished, substitute and lost trips across the two scenarios help to explain the (modest) differences in the estimated values of a lost user day, which capture all three components.

VI. Sensitivities

To assess the robustness of our damage estimate to various assumptions, we conducted a series of sensitivity analyses. In general, these analyses implied estimates that were either quantitatively similar to those reported in the previous section or different in intuitive ways. Table 6 reports a subset of our findings. To benchmark these estimates, we use the baseline period one lost user day value of $37.23 and report percent deviations from that estimate.[24]

The first sensitivity in Table 6 adjusts the value of time used in travel cost construction from one-third to one-half the wage rate as recommended by U.S. DOT (2014). The resulting lost user day value

[23] The associated standard errors on the lost user day values are 0.99 for the first spill period and 1.06 for the second.
[24] Our sensitivity results do not qualitatively change when we use lost user day value for the second spill period as the baseline.

rises by almost 22 percent.[25] The next three sensitivities consider alternative approaches to imputing missing incomes, including an approach that drops all cases where income data was unavailable. In all three sensitivities the lost user day value changes by less than four percent. In the fifth sensitivity we replace expected travel cost with driving travel costs as Alvarez et al. (2015) and Glasgow and Train (2017) do. This alternative specification raises the lost user day value by 13.2 percent. Next we employ the minimum of flying and driving travel costs in place of expected travel costs, which reduces the lost user day value by four percent. In our seventh sensitivity, rather than use percentage declines in trips for the calibration, we first calibrate both the baseline and the spill scenarios to match the trips we observe with the infield counts and then solve the spill effect by matching the change in trips. With this alternative calibration approach, we find a one percent change in the lost user day value. We then consider two sensitivities that involve a more and less aggregated site definition, which move our lost user day value by less than 1.5 percent. Finally, we consider a three-level nested logit model which allows for greater correlations among neighboring sites. Again, we find virtually no effect on our lost user day value.[26]

VII. Conclusions and Lessons Learned

The Deepwater Horizon oil spill of 2010 was the largest spill ever in U.S. waters. The spill prompted a correspondingly large effort to estimate lost recreational use value as part of the federal-state

[25] Based on empirical results from a route choice model for recreational trips in Italy, Fezzi et al. (2015) argue that an even higher value of time of roughly three-quarters of the wage rate is appropriate. Their approach to measuring the value of time, however, is relative to a respondent's self-reported wages, not his or her annual household income like we and many other recreation studies typically do. The difference is subtle but important; in most countries, household income is significantly larger than individual wages, implying that a value of time based on a fraction of household income-based wage rate will be much higher than a value of time based on the same fraction applied to an individual's wage rate. For example, in the United States the ratio of average individual wages to annual household income is roughly 1.6, which implies that the Fezzi et al. value of time relative to annual household income is closer to one-half, which is what we consider in this sensitivity.

[26] For a variety of reasons, we do not report a random coefficient model sensitivity. As Klaiber and von Haefen (2018) show, panel versions of these models often fit the data better than simpler fixed coefficient models but predict poorly in-sample. These prediction errors propagate through to welfare measures and thus call their credibility into question. In addition, our data is poorly suited for panel random coefficient models. Recall from the previous section that we employ weights in estimation to correct for oversampling of likely coastal recreators. Because reporting periods vary across respondents, we developed separate weights by month. Within a panel random coefficient model where coefficients are fixed across the full reporting period, these time-varying monthly weights enter the log-likelihood function nonlinearly, raising significant numerical challenges which prevented us from recovering robust and stable parameter estimates.

claim for natural resource damages under the Oil Pollution Act of 1990. This paper presents the shoreline model of demand for Gulf recreation, a central component of the damage assessment. The construction of the model, which included a variety of innovations relative to the standard travel cost/recreation demand methodology, was driven by the size of the spill. The spill was large enough to threaten contamination of coastal resources throughout the Gulf, with some potential for harm beyond the Gulf. Given this scale of impact on resources, it was anticipated that demand for Gulf recreation from across the 48 contiguous states would likely be affected. To accommodate demand throughout the US, our shoreline model is a repeated discrete choice model of recreational trips from any of the lower 48 states to 83 sites in the Gulf coast and nearby areas.

In the two-step procedure for estimating recreation damages, the shoreline model provided the value of a lost user day, while the infield surveys estimated the number of lost user days. We illustrate damages for the first spill impact period, the eight months from June 2010 through January 2011, by multiplying the lost user day value ($37.23) by the lost user days (10.17 million in Table 5), which equals $379 million. In the second period, February through November 2011, the value of a lost user day was $40.41, which yields losses of $87 million when multiplied by the second period lost user days (2.154 million, Table 5). These two damage estimates illustrate how the shoreline model is combined with lost user days, but they are underestimates because they are not compounded from 2010 to July 2015, the end date used for damage calculations. In practice, damages were calculated on a monthly basis and compounded forward to July 2015.[27] When fully compounded, the primary study damage estimates for shoreline activity are $520 million, with a 95% confidence interval of $354-$685 million.

The assessment of the full recreational damages entailed calculating and compounding the monthly lost user days for both periods for the shoreline and three months in the North Gulf for boating.[28]

[27] Damages are calculated as $\sum_{m=1}^{M} D_m(1+r_m)^{T-m}$ where D_m is monthly damages, r_m is the monthly compound factor based on an annual discount rate of 3 percent, m is the number of months past the beginning of spill impact and T is number of months between the spill impact beginning and July 2015. When appropriately compounded, damages from shoreline activity applied to primary sampling data equal $520 million.
[28] The full recreational damages are described in the technical memo called Overview of Recreation Assessment—see Appendix A for the location of this memo and further supporting documentation.

These damage estimates were combined with estimates of lost user days from supplemental studies that were designed to address coverage limitations with the user day surveys. These supplemental studies include estimates for activities such as recreational boating, shoreline activities before and after the infield counting periods, charter or head boat fishing, and barrier island uses. Adding these to our shoreline valuation model damage estimates, we estimated total recreation damages from the Deepwater Horizon oil spill equal to $661 million.[29]

Estimating a recreational demand model for a spill with nationwide impacts required innovations addressing a series of issues central to the recreation demand literature. To appropriately account for the nationwide scope of demand, we have pooled long trips from as far away as Washington State with short trips originating nearby on the Gulf Coast. This approach differs from many studies in the recreation literature, which tend to separate single and multiple day trips or focus solely on single day trips. As discussed in English et al. (2018a), pooling long and short trips in a single demand model reveals the surplus of recreational trips and is consistent with an economic model of endogenous trip length choice.

The sampling plan for gathering data for the national model was a general population, address-based phone survey of adults in the lower 48 states. For efficiency reasons, the sampling plan called for oversampling various strata, such as recreators or residents near the Gulf coast. Over- and under-sampling is corrected by the sampling weights. When weighted up by the sampling weights, the respondent behavior represents the behavior of the population of adults in the lower 48 states. Our complex, multi-stage sampling weights are an appropriate and critical component of the study, and represent a departure from common practice in past recreation studies.

For our general population demand model, it was necessary to calculate each person's exogenous costs of travel to each of the sites, regardless of whether they visited sites. Consequently, we used the expected costs, that is, the cost that one would predict with information on the distance to the sites and household characteristics. As distance from the sites increased, the potential for flying becomes greater.

[29] By comparison, estimated income losses throughout the Gulf seafood industry ranges from $21 to $310 million (Carroll et al. 2016) and lost non-use values have been estimated to be $17 billion nationwide (Bishop et al. 2017).

Hence our expected cost included the probability of flying versus driving. The flying costs included a detailed assessment of the time and travel costs of different airports and routes. Our sensitivity results show that failure to account for the potential for flying would have overstated welfare losses.

The estimated repeated discrete choice model was specified with alternative-specific constants rather than site characteristics. Using ASCs was an essential element of the research design and enabled our linkage between the model based on the telephone data and the lost user days measured by the infield counts. We were able to launch the infield surveys shortly after the spill began. Their purpose was to provide the best possible estimates of user days during the spill and baseline periods. The valuation surveys for the modelling data were begun and completed during the baseline period, thereby accurately capturing the "but for" the spill site choice behaviors. The ASCs of the shoreline model enabled our innovative calibration of the model to spill period conditions as estimated from the actual declines from infield surveys. Further, the adjustment of the ASCs, which simulated the impact of the spill in the demand model, gave a direct means of calculating welfare losses without any confounding unmeasured site characteristics. Moreover, our comprehensive confidence intervals for the overall welfare losses account for uncertainty in the on-site counts as well as uncertainty in the valuation model due to sampling and missing income.

Finally, the estimation of the value of a lost user day was shown to be robust to a variety of model specifications. We learned that despite our extensive efforts to deal with the ten percent of observations for which we had no income data, the issue had little impact on the estimated values. Further, even though there are numerous articles and conventional wisdom suggesting that nesting structures and site aggregation affect estimated values, we found these had little effect on our estimated values, perhaps due to our ASCs and damage estimation approach. In fact, one insight into why the value of a lost user days was quite robust comes from the recognition that factors that affect the estimated welfare measure in the value per lost user day numerator also affect the estimated change in days in the denominator. The robustness of the value per lost user day measure highlights another advantage of our two-part strategy to derive the lost value per day from the recreational demand model and apply it to the best available external data (i.e., the onsite counts) for the measurement of the demand shift.

VIII. References

AAA (American Automobile Association), 2012. Your driving costs. Retrieved from http://newsroom.aaa.com/wp-content/uploads/2012/04/YourDrivingCosts2012.pdf, accessed September 30, 2013.

Alvarez, S., Larkin, S., Whitehead, J. and T. Haab, 2014. A revealed preference approach to valuing non-market recreational fishing losses from the Deepwater Horizon Oil Spill. *Journal of Environmental Management*, 145:199-209.

AHLA (American Hotel & Lodging Association), 2013. Lodging industry profile. Retrieved from https://www.ahla.com/content.aspx?id=35603, accessed February 6, 2014.

Battaglia, Michael P., David C. Hoaglin, and Martin R. Frankel, 2013. Practical considerations in raking survey data. *Survey Practice,* 2(5).

Berry, S., L. Levinsohn, and A. Pakes, 1995. Automobile prices in market equilibrium. *Econometrica,* 63(4): 841-890.

Bishop, Richard C., Kevin J. Boyle, Richard T. Carson, David Chapman, W. Michael Hanemann, Barbara Kanninen, Raymond J. Kopp, Jon A. Krosnick, John List, Norman Meade, Robert Paterson, Stanley Presser, V. Kerry Smith, Roger Tourangeau, Michael Welsh, Jeffrey M. Wooldridge, Matthew De Bell, Colleen Donovan, Matthew Konopka, and Nora Scherer, 2017. "Putting a Value on Injuries to Natural Assets: The BP Oil Spill. *Science* 356(6335): 253-254.

Carroll, M., B. Gentner, S. Larkin, K. Quigley, N. Perlot, L. Dehner and A. Kroetz. 2016. An analysis of the impacts of the Deepwater Horizon oil spill on the Gulf of Mexico seafood industry. U.S. Dept. of the Interior, Bureau of Ocean Energy Management, New Orleans, LA.

Carson, R.T., R.C. Mitchell, W.M. Hanemann, R.J. Kopp, S. Presser, and P.A. Ruud, 1992. A contingent valuation study of lost passive use values resulting from the Exxon Valdez Oil Spill. Report to the Attorney General of the State of Alaska.

Carson, R. T., R. C. Mitchell, W.M Hanemann, R.J. Kopp, S. Presser, and P.A. Ruud, 2003. Contingent valuation and lost passive use: Damages from the Exxon Valdez Oil Spill, *Environmental and Resource Economics*, 25:257-286.

Diamond, P.A. and J.A. Hausman, 1994. Contingent valuation: Is some number better than no number? *Journal of Economic Perspectives*, 8(4):45-64.

English, E., K.E. McConnell, R.H. von Haefen, and F. Lupi, 2018a. Should single day and multiple day trips be pooled when estimating travel cost models? working paper, May 30.

English, E., J. Herriges, F. Lupi, K.E. McConnell, and R.H. von Haefen, 2018b. Fixed costs and recreation value, working paper, March 12.

Fezzi, C., I.J. Bateman and S. Ferrini, 2015. Using revealed preference to estimate the value of travel time to recreation sites. *Journal of Environmental Economics and Management,* 72:123-136.

Glasgow, G. and K. Train, 2018. Lost use value from environmental injury when visitation drops at undamaged sites. *Land Economics,* 94(1):87-96.

Hausman, J., G. Leonard, and D. McFadden, 1995. A utility-consistent combined discrete choice and count data model assessing recreation use losses due to natural resource damage. *Journal of Public Economics* 56:1-30.

Klaiber, H.A. and R.H. von Haefen, 2018. Do random coefficient and alternative specific constants improve policy analysis? An empirical investigation of model fit and prediction. working paper.

Kling, C.L., D. J. Phaneuf, and J. Zhao, 2012. From Exxon to BP: Has some number become better than no number? *Journal of Economic Perspectives*. 26(4): 3-26.

Leggett, C., 2015. Estimating visitation in national parks and other public lands. Final report prepared for the National Park Service, Environmental Quality Division under Award Number P13D02250.

Lew, D.K. and D.M. Larson, 2008. Valuing a beach day with a repeated nested logit model of participation, site choice, and stochastic time value. *Marine Resource Economics,* 23: 233-252.

McConnell, K.E., 1992. Onsite time in the demand for recreation. *American Journal of Agricultural Economics* 74: 918-925.

McFadden, D., 1974. Conditional logit analysis of discrete choice behavior. In *Frontiers of Econometrics*, edited by Paul Zarembka. New York: Academic Press.

McFadden, D., 1978. Modeling the choice of residential location. In *Spatial Interaction and Planning Models*, 75-96, edited by A. Karlqvist. Amsterdam: North Holland.

McFadden, D., 1981. Econometric models of probabilistic choice models. In *Structural Analysis of Discrete Choice Data with Applications*, edited by C. Manski and D. McFadden. Cambridge: MIT Press.

Moeltner, K. and J. Englin, 2004. "Choice behavior under time-variant quality: State dependence versus "play-it-by-ear" in selecting ski resorts," *Journal of Business and Economic Statistics*, 22(2): 214-224.

Moeltner, K., and R. von Haefen, 2011. Microeconometric strategies for dealing with unobservables and endogenous variables in recreation demand models. *Annual Review of Resource Economics*, 3 (1): 375-396.

Morey, E.R., 1999. Two RUMs Uncloaked: Nested-logit models of site choice and nested-logit models of participation and site choice. In *Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice*, edited by Joseph A. Herriges and Catherine L. Kling Cheltenham: Edward Elgar.

Morey, E.R., R.D. Rowe, and M. Watson, 1993. A repeated nested-logit model of Atlantic salmon fishing. *American Journal of Agricultural Economics,* 75(3): 578-592.

Murdock, Jennifer, 2006. Handling unobserved site characteristics in random utility models of recreation demand. *Journal of Environmental Economics and Management*, 51(1): 125.

Parsons, G., A. Kang, C. Leggett, and K. Boyle, 2009. Valuing beach closures on the Padre Island National Seashore. *Marine Resource Economics* 24:213-235.

Portney, P.R., 1994. The contingent valuation debate: Why should economists care? *Journal of Economic Perspectives*, 8(4), 3-17.

Pew Research Center, 2010. *Top stories of 2010: Haiti earthquake, Gulf Oil Spill*. Pew Research Center news release. Retrieved from http://www.people-press.org/2010/12/21/top-stories-of-2010-haiti-earthquake-gulf-oil-spill/, accessed on December 21, 2010.

Provencher, B., and R. Bishop, 1997. An estimable dynamic model of recreation behavior with an application to Great Lakes angling. *Journal of Environmental Economics and Management*, 33(2): 107-127.

Rubin, D., and N. Schenker, 1986. Multiple imputations for interval estimation from simple random samples with ignorable nonresponse. *Journal of the American Statistical Association,* 81: 366-374.

Rust, K., 1985. Variance estimation for complex estimators in sample surveys. *Journal of Official Statistics*, 1(4):381-397.

Schenker, N., Raghunathan, T.E., Chiu, P.L., Makuc, D.M., Zhang, G. and A.J. Cohen, 2006. Multiple imputation of missing income data in the National Health Interview Survey. *Journal of American Statistical Association*, 101:924-933.

Solon, G., S.J. Haider and J.M. Wooldridge, 2015. What are we weighting for? *Journal of Human Resources*, 50(2):301-316.

Steimetz, S.C., and D. Brownstone, 2005. Estimating commuters' 'value of time' with noisy data: A multiple imputation approach. *Transportation Research Part B,* 39:865-889.

Stratus Consulting, Incorporated, 2010. *Damage Estimate for Shoreline Recreation*. Appendix J to Cosco Busan Oil Spill: Final Damage Assessment and Restoration Plan/Environmental Assessment (February 2012), prepared by California Department of Fish and Game, California State Lands Commission, National Oceanic and Atmospheric Administration, United States Fish and Wildlife Service, National Park Service, and Bureau of Land Management. https://casedocuments.darrp.noaa.gov/southwest/cosco/pdf/CB%20Final%20DARP%20revised%202-28-12.pdf

Thiene, M, G. Swait and R. Scarpa, 2017. "Choice set formation for outdoor destinations: The role of motivations and preference discrimination in site selection for the management of public expenditures on protected areas. *Journal of Environmental Economics and Management*, 81(C): 152-173.

Tourangeau, R., E. English, K.E. McConnell, D. Chapman, I.F. Cervantes, E. Horsch, N. Meade, A. Domanski and M. Welsh, 2017. The Gulf recreation study: Assessing lost recreational trips from the 2010 Gulf Oil Spill. *Journal of Survey Statistical Methodology,* 5(3):281-309.

Tourangeau, R. and E. English, 2015. Estimation procedures for count data, Technical Memorandum B1, Deepwater Horizon Oil Spill Administrative Record, https://www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0026633.pdf.

Train, K., 2016. Comment on "A revealed preference approach to valuing non-market recreational fishing losses from the Deepwater Horizon Oil Spill" and its "Corrigendum" by Alvarez et al. *Journal of Environmental Management,* 167:259-261.

U.S. BTS (Bureau of Transportation Statistics), 2013. *National Transportation Statistics,* 2013, Table 4-23.

U.S. DOT (Department of Transportation), 2014. *Revised Departmental Guidance on Valuation of Travel Time in Economic Analysis*, Internal memo from Roberto Ayala to Peter Rogoff dated July 9, 2014.

U.S. EIA (Energy Information Administration), 2013. Weekly retail gasoline and diesel prices. Retrieved from http://www.eia.gov/dnav/pet/pet_pri_gnd_a_epmr_pte_dpgal_m.htm, accessed October 10, 2013.

U.S. v. BP et al. (United States of America v. BP Exploration & Production, Inc., et al.), 2015. Findings of fact and conclusions of law: Phase Two trial. In re: Oil spill by the oil rig "Deepwater Horizon" in the Gulf of Mexico, on April 20, 2010, No. MDL 2179, 2015 WL 225421 (LA. E.D. Jan. 15, 2015). (Doc. 14021). U.S. District Court for the Eastern District of Louisiana.

von Haefen, R., 2008. Latent consideration sets and continuous demand systems. *Environmental and Resource Economics*, 41 (3): 363–379.

Whitehead, J. C., T. Haab, S. Larkin, J. B. Loomis, S. Alvarez and A. Ropicki, 2018. Estimating lost recreational values of visitors to northwest Florida due to the Deepwater Horizon oil spill using cancelled trip data. *Marine Resource Economics,* 33(2):119-132.

Welsh, M., 2015. Media tracking and public awareness, Technical Memorandum I10, Deepwater Horizon Oil Spill Administrative Record, www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0079605.pdf

IV. Figures and Tables

Figure 1
Location of Survey Respondents from the Local and National Survey Components

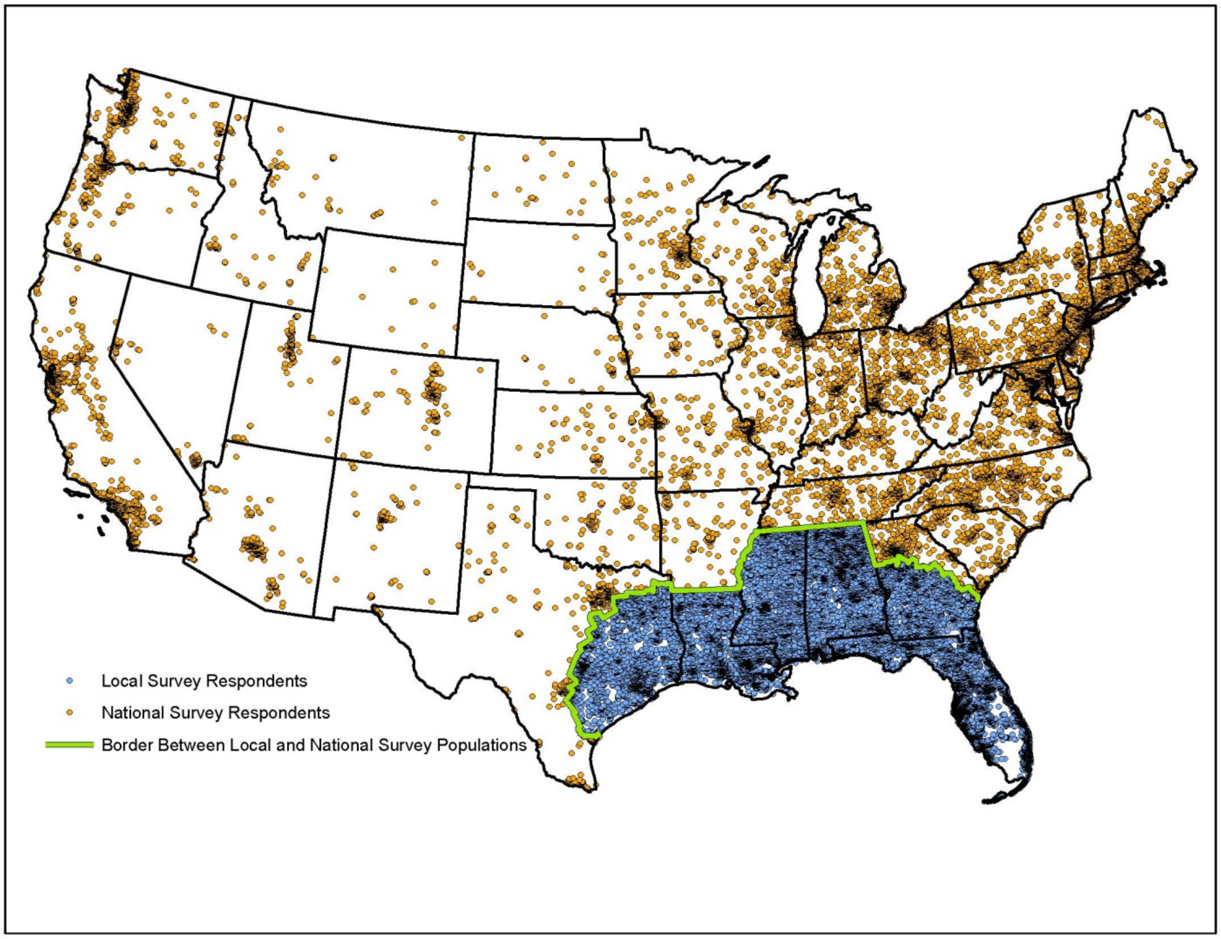

Figure 2
Local and National Sampling Waves

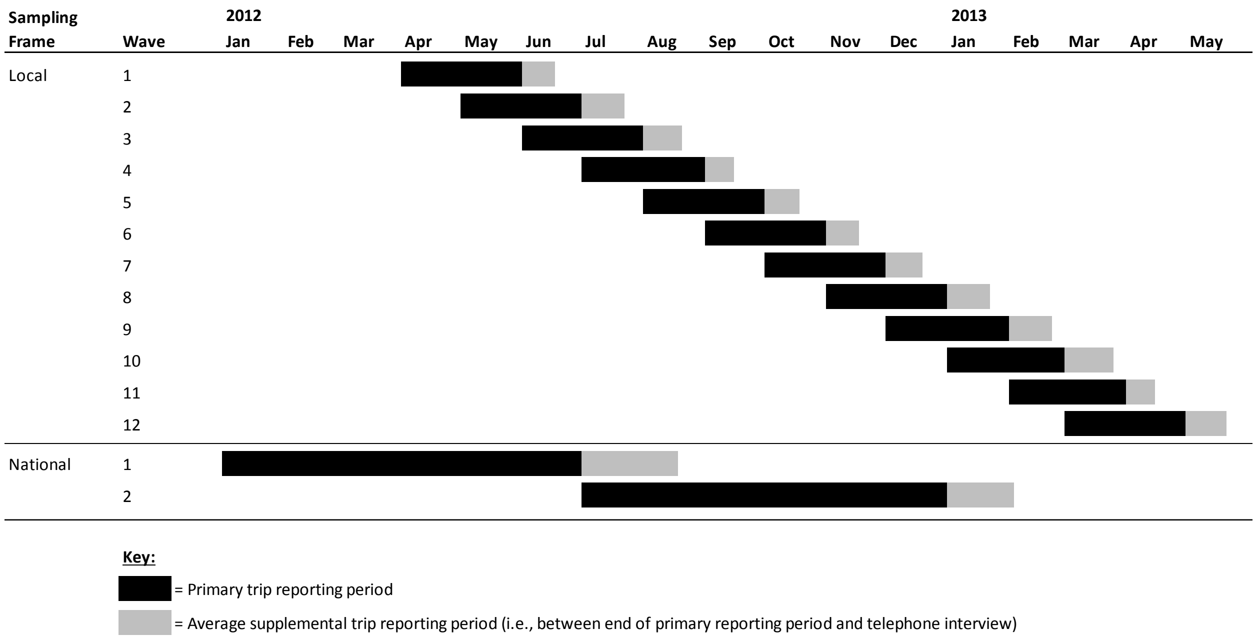

Figure 3
Coastal Area Covered by Shoreline Sites

Figure 4
Average Driving, Flying and Expected Travel Cost by One-Way Driving Distance

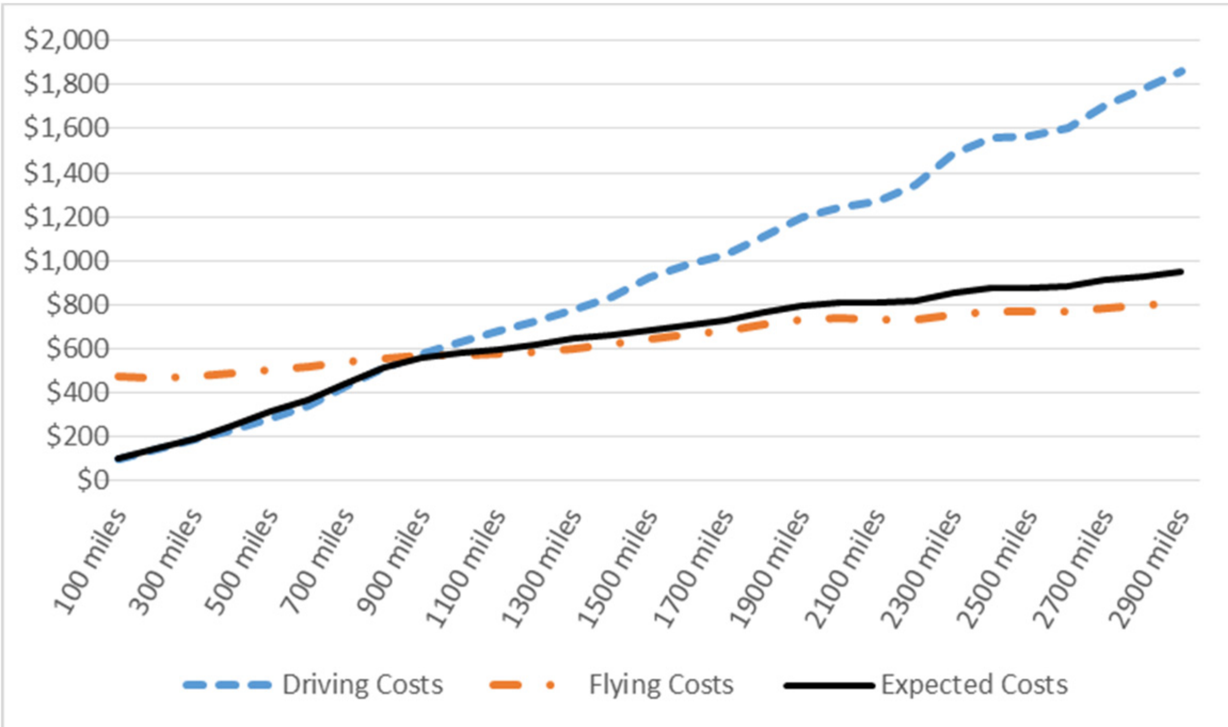

Figure constructed by averaging all travel costs falling within separate 100 mile one-way driving distance bins across all observations and trips in the estimation sample.

Table 1
Survey Sampling

	Local Survey		**National Survey**
	Address Sample	Boater Sample	Address Sample
Screeners Mailed	325,285	67,072	248,001
Screeners Returned Eligible for Telephone Survey	115,161	29,368	97,714
Screeners Returned Not Eligible for Telephone Survey	32,591	2,729	19,279
Weighted Screener Response rate	41.5%	46.5%	45.9%
Telephone Survey Sample	72,135	24,506	51,519
Complete Telephone Interviews	22,774	7,104	13,457
Weighted Telephone Response Rate	29.6%	28.4%	26.1%

Table 2
Probability of Flying

One-way driving distance	*HH Inc ≤ \$70k, Family Size ≤ 2*	*HH Inc > \$70k, Family Size ≤ 2*	*HH Inc ≤ \$70k, Family Size > 2*	*HH Inc > \$70k, Family size > 2*
≤ 250 Miles	0.000	0.000	0.000	0.000
> 250 Miles & ≤ 500 Miles	0.000	0.030	0.000	0.000
> 500 Miles & ≤ 1000 Miles	0.168	0.338	0.056	0.201
> 1000 Miles & ≤ 1500 Miles	0.736	0.788	0.443	0.784
> 1500 Miles	0.842	0.880	0.842	0.880

HH Inc = household income; *Family Size* = total number of adults and children in the household.

Table 3
Socio-Demographic Variables Interacted with the No-Trip Constant

Variable	*Description*
\$25k < Income ≤ \$50k	Dummy variable equal to one if respondent's household income between \$25,000 and \$50,000, 0 otherwise
\$50k < Income ≤ \$75k	Dummy variable equal to one if respondent's household income between \$50,000 and \$75,000, 0 otherwise
\$75k < Income ≤ \$100k	Dummy variable equal to one if respondent's household income between \$75,000 and \$100,000, 0 otherwise
\$100k < Income ≤ \$150k	Dummy variable equal to one if respondent's household income between \$100,000 and \$150,000, 0 otherwise
\$150k < Income	Dummy variable equal to one if respondent's household income greater than \$150,000, 0 otherwise
% Urban	Urban population divided by total population for the respondent's zip code (Source: American Community Survey)
Age	Respondent's age normalized by 100
High school diploma	Dummy variable equal to one if respondent has no more than a high school diploma, 0 otherwise
College degree	Dummy variable equal to one if respondent has a 4-year college degree, 0 otherwise
Full-time	Dummy variable equal to one if respondent is employed full-time, 0 otherwise
Part-time / student / disabled	Dummy variable equal to one if respondent is employed part-time disabled or a full-time student, 0 otherwise
Retired	Dummy variable equal to one if respondent is retired, 0 otherwise
White	Dummy variable equal to one if respondent is white, 0 otherwise
Male	Dummy variable equal to one if respondent is male, 0 otherwise
HH members ≥ 18	Count of adults (18 years and older) in respondent's household

Table 4
Parameter Estimates for Shoreline Valuation Model

Variable	*Estimate*	*t-stat*
Travel cost/100	-1.119***	-25.289
Dissimilarity coefficient	0.226***	20.064
No-trip constant interacted with:		
$25k < Income ≤ $50k	-0.767***	-5.055
$50k < Income ≤ $75k	-0.838***	-4.158
$75k < Income ≤ $100k	-1.026***	-4.108
$100k < Income ≤ $150k	-0.989***	-4.002
$150k < Income	-2.330***	-3.336
% Urban	-0.708***	-5.932
Age	-3.350*	-1.801
Age^2	4.939***	2.885
High school diploma	0.462**	2.405
College degree	-0.202	-1.661
Full-time	-0.461*	-1.911
Part-time / student / disabled	-0.667***	-3.416
Retired	-0.617**	-2.266
White	-0.605**	-2.519
Male	0.002	0.005
HH members ≥ 18	-0.017	-0.063
HH members < 18	0.185**	1.991
Observations	41,708	

Table 5
Lost User Days from Deepwater Horizon Oil Spill

	Baseline Level	Spill Period	Loss	Percent Loss
Period 1 - Months 1-8				
North Gulf	13,838,623	7,538,278	6,300,345	45.5%
Peninsula	16,895,848	13,025,672	3,870,176	22.9%
Period 2 - Months 9-18				
North Gulf	21,243,632	19,088,641	2,154,991	10.1%

Source: Tourganeau *et al.* (2017), Table 4.

Table 6 - Sensitivities

Baseline Model			
Lost User Day Value – Period 1			$37.23
Sensitivity			
	Baseline Assumption	***Alternative Assumption***	***Percent Change***
1.	Opportunity Cost of Time = 33% of hourly wage	Opportunity Cost of Time = 50% of hourly wage	+21.9
2.	Impute missing incomes; no upper or lower limits	Drop cases with no income data; use midpoint if range provided and use $150k if income >= $150k	+2.7
3.	Impute missing incomes; no upper or lower limits	Impute missing incomes, then impose maximum = $1m	+4.0
4.	Impute missing incomes; no upper or lower limits	Complete cases only (no imputation of any variables)	+2.2
5.	Travel cost = expected travel cost across flying and driving	Travel cost = driving travel cost	+13.2
6.	Travel cost = expected travel cost across flying and driving	Travel cost = minimum of flying and driving travel cost	+4.0
7.	Calibrate to approximate percentage change in trips (baseline versus spill)	Also calibrate to approximate count levels of trips under baseline conditions	+1.0
8.	83 recreation sites	31 recreation sites	+1.5
9.	83 recreation sites	185 recreation sites	+0.2
10.	Two-level nested logit model	Three-level nested logit model with 4 sub-nests: 1) Texas, 2) North Gulf, 3) Florida Peninsula, 4) Atlantic Coast (common dissimilarity coefficients)	-0.5

Appendix A: Chief Technical Memos for the Shoreline Model

Data and methods used in the analyses of recreational losses due to the Deepwater Horizon oil spill are described on the public archive for the case located at: www.doi.gov/deepwaterhorizon/adminrecord under the heading "5.10 Lost Human Use." That folder includes the following subfolders containing documents and data for the case:

5.10.2 Study Protocols: Contains detailed study protocols for the counts of recreational users and aerial photographs, including specific site definitions for every shoreline segment.

5.10.3 Surveys: Contains 20 reports and scripts for the various survey interviews.

5.10.4 Technical Reports: Contains 64 technical reports for with exhaustive details of the methods used to develop the data, models and results in support of the findings summarized here.

5.10.5 Data: Contains documents listing all original datasets developed in support of the work and pre-existing datasets utilized in the work. In particular, the file www.fws.gov/doiddata/dwh-ar-documents/941/DWH-AR0305129.pdf includes download instructions and links to each original dataset available at NOAA's public data archive: www.diver.orr.noaa.gov/. In addition, the exact datasets and code used to estimate the shoreline model is available at www.diver.orr.noaa.gov/deepwater-horizon-nrda-data by clicking on "Download Data Packages," selecting "Recreational Use" from the Topic dropdown and choosing "Valuation Model Datasets."

Of the many technical reports, Table A.1 summarizes key reports in supporting the shoreline recreation modelling.

Table A.1: Key supporting reports and their location

Report/Description	Link
General Documents	
Overview of Recreation Assessment	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0021412.pdf
Summary of the findings and methods along with details of the economic damages by time periods, spatial areas, and types of activities.	
Estimation Procedures for Count Data	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0026633.pdf
Methods and results for the on-site counts of visitors using on-site intercept counts and interviews as well as aerial overflights of sandy beaches.	
Surveys	
National Valuation Survey	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0061973.pdf
Details of the national general population survey that collected data for the demand models; covers residents of all contiguous US except for those in "local" survey.	
Local Valuation Survey	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0066343.pdf
Details of the local general population survey that collected data for the demand models; covers residents of all Louisiana, Mississippi, Alabama, Florida, and parts of Texas and Georgia.	
Weights for National Valuation Survey	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0061982.pdf
Description of the national survey weights that adjust for sample periods, strata, deviations from census variables, and other factors.	
Weights for Local Valuation Survey	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0066354.pdf
Description of the local survey weights that adjust for sample periods, strata, deviations from census variables, and other factors.	
Valuation Model	
Model Structure	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0045972.pdf
Description of the econometric specification of the multi-site demand system models for deriving demand responses and valuation of trips.	
Parameter Estimates	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0262432.pdf
Econometric estimation results of demand model parameters.	
Calibration Methods	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0045983.pdf
Methods and implementation of the approach for calibrating the demand system to the lost trips from the on-site counts to derive the demand effect and lost value due to the spill.	
Precision Estimation	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0045986.pdf
Methods for estimating the precision of overall assessment results by accounting from the dual sources of estimation uncertainty from both the on-site counts and the demand models.	
Travel Cost Computation	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0056724.pdf
Details of the computation of each sample member's expected travel costs from their origin to each possible destination and accounting for the chances of driving and flying.	

Table A.1: Key supporting reports and their location (cont.)

Report/Description	Link
Air Travel Data Integration	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0056788.pdf
Details of the data sources and methods used to determine possible airline routes and airfares for the travel cost computations.	
Value of Time and Income Imputation	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0056732.pdf
Details of income multiple imputation method for the subset of cases with incomes missing or reported in a range, along with derivation of value of travel time form income.	
Monthly Allocations	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0056749.pdf
Details and methods for determining the months for all trips made by each survey respondent.	
Multiple-Day Trips	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0056826.pdf
Overview of rationale for combining single and multiple day trips when trip length is endogenous.	
Site Definitions	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0056785.pdf
Description of the process for mapping the trip destinations from the survey data into the sites used in the demand models.	
Model Sensitivities	www.fws.gov/doiddata/dwh-ar-documents/940/DWH-AR0056807.pdf
Description and results of the suite of sensitivities performed on the demand models and the implications for the valuation of lost trips.	

Appendix B: Variance Construction for the Estimated Total Loss and Value per Lost User Day

The two key outputs from the shoreline model are: (a) the estimated value per lost user day $\mathcal{V}_t$ and (b) the estimated total loss $L_t = \mathcal{V}_t \cdot \Delta_t$ where Δ_t denotes the estimated lost user days for period *t*. Computing variances for these outputs requires taking into account three sources of uncertainty:

- Uncertainty in the estimated lost user days for period *t* based upon the infield counts;
- Uncertainty due to sampling variability in the valuation survey; and
- Uncertainty due to the imputation of income for those individuals in the valuation survey who did not provide household income or who provided only income bounds.

The first source of uncertainty is summarized by $\sigma^2_{\Delta t} = Var[\Delta_t]$, which is constructed using a jackknife variance estimator. The other two sources of uncertainty both affect the estimated value per lost user day ($\mathcal{V}_t$).

The overall variance $\sigma^2_{\mathcal{V}\mathrm{t}}$ is constructed using a combination of a jackknife variance estimation procedure (Rust, 1985) and Rubin's multiple imputation method with ignorable nonresponse (Rubin and Schenker, 1986; Steimetz and Brownstone, 2005). The key inputs to this process are:

- A total of $R = 120$ sets of monthly replicate weights ($\omega_{ir}, \mathrm{i} = 1, \dots, \mathrm{N}; r = 1, \dots, R$) and
- A total of $S = 5$ income imputations ($I_{is}, \mathrm{i} = 1, \dots, N; s = 1, \dots, 5$).

Each set of monthly replicate weights is randomly paired with one of the five income imputations. Let n_s denote the number of replicate weight sets assigned income imputation *s*, with $\sum_s n_s = R$, and s_r denote the income imputation assigned to replicate weight set *r*. For each of these pairings (r, s_r), the recreation demand model is re-estimated using the associated replicate monthly weights ($\omega_{\cdot r}$) in constructing trips and the associated incomes ($I_{\cdot s_r}$) in forming travel costs and the income brackets included in the model's stay-at-home option. The resulting estimated model is then used to form the estimated value per lost user day $\hat{\mathcal{V}}_{\mathrm{t}}^{\mathrm{rs_r}}$ for pairing (r, s_r). Let

(17) $$\bar{\mathcal{V}}_{\mathrm{t}}^{s} = \frac{1}{n_s}\sum_{r=1}^{R} \delta_{k_r s}\, \hat{\mathcal{V}}_{\mathrm{t}}^{\mathrm{rk_r}}$$

denote the mean value per lost user day across estimates using income imputation *s*, where $\delta_{ks} \equiv 1(k = s)$.

The variance within imputation *s* is given by:

(18) $\Omega_{st} = \left(\frac{R-1}{n_s}\right) \sum_{r=1}^{R} \delta_{k_r s} \left(\hat{\mathcal{V}}_t^{rk_r} - \bar{\mathcal{V}}_t^{s}\right)^2,$

with the average *within* imputation variance given by:

(19) $U_t = \left(\frac{1}{S}\right) \sum_{s=1}^{S} \Omega_{st}.$

The *between* imputation variance is given by:

(20) $B_t = \left(\frac{1}{S-1}\right) \sum_{s=1}^{S} \left(\tilde{\mathcal{V}}_t^{s} - \overline{\mathcal{V}_t}\right)^2.$

where

(21) $\bar{\mathcal{V}}_t = \left(\frac{1}{S}\right) \sum_{s=1}^{S} \tilde{\mathcal{V}}_t^{s}.$

The overall variance in the value per lost user day is then given by:

(22) $\hat{\sigma}_{\mathcal{V}t}^2 = U_t + \left(1 + \frac{1}{S}\right) B_t.$

Treating the uncertainty stemming from the counts as independent of the uncertainty stemming from the valuation survey (i.e., sources 2 and 3 mentioned above), we can compute the variance for the overall loss as:

$$\hat{\sigma}_{Lt}^2 = \hat{\sigma}_{\mathcal{V}t}^2 \cdot \sigma_{\Delta t}^2 + \hat{\sigma}_{\mathcal{V}t}^2 \cdot \Delta_t^2 + \sigma_{\Delta t}^2 \mathcal{V}_t^2.$$