Smart Device Recreation Data: Use and Limitations for a Tank Fire in Texas

Cameron Duff – National Oceanic and Atmospheric Administration, Assessment and Restoration Division, Silver Spring, MD 20910, USA

Chris Giguere – Independent Researcher

Jason Murray – National Oceanic and Atmospheric Administration, Assessment and Restoration Division, Silver Spring, MD 20910, USA

Abstract
Human mobility data (MD) from smartphones may offer a scalable alternative to fieldwork for Natural Resource Damage Assessments (NRDAs) and non-market valuation. We test MD's utility for estimating recreational demand changes after a 2019 Houston tank fire. We apply count regressions and zonal travel cost models to calculate welfare losses. While this MD dataset reflects expected temporal patterns, comparisons with reference data reveal that "coverage rates" vary significantly across sites. This measurement error complicates counterfactual predictions. Economists should use caution when deriving absolute recreational value from MD.

NOAA Disclaimer: The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect the views of NOAA or the Department of Commerce.

Acknowledgements:
The authors wish to thank Bryand Duke, Christopher Plaisted, Susan Snyder, Dan Takaki and participants in the workshop sponsored by the editors of Land Economics, the editors themselves and an anonymous referee for thoughtful comments and suggestions.

1. Introduction

Understanding outdoor recreation is an important objective for natural resource management. In the natural resource damage assessment (NRDA) context, agencies have a special interest in modeling changes in recreation that result from impacts to natural resources. There is a well-developed toolkit of assessment strategies dating back to Hotelling’s famous letter to the National Park Service in 1947, sketching his ideas for modeling consumer surplus as a function of the cost of travel. NRDA cases in general provide motivation for the development of rigorous

doi:10.3368/le.102.3.011626-0005

recreational demand models and techniques (Breffle et al., 1999; Byrd et al., 2001; Chapman & Hanemann 2001; Athos TWG 2007; MacNair & Desvouges 2007; English 2008; Bouchard TWG 2009). Other studies have been inspired directly by the NRDA framework (Deacon & Kolstad 2000; Von Haefen et al., 2004; Parsons et al., 2009). This literature informs the development of the current "gold standard" methods as employed in the Deepwater Horizon NRDA (Horsch et al., 2017; English et al., 2018). But these established methods are resource intensive. They require large up-front expenditures for mobilization of personnel, collection of ephemeral data, and development of survey instruments, typically on a rushed time frame. These methods can be difficult to scale down for smaller NRDA incidents that occur more frequently. The challenges of recreational data collection for NRDA motivate practitioners to keep a watchful eye for new methods and technologies to observe and model recreation.

In this paper, we develop a strategy for quantifying recreational losses using mobility data (MD), with anonymized user locations recorded by apps on smart devices. MD has some promising features that may help facilitate recreation analysis from a remote setting while saving on labor costs. MD offers unlimited spatial coverage of key areas, but even more importantly it offers an opportunity to look backwards in time at recreation that would otherwise have gone unmonitored. If the data are reliable, they could provide valuable and unusual insight[1] into pre-incident baseline use of a resource.

The ITC incident[2] provides a useful test case for MD. The incident involved a large fire at a chemical storage tank farm near Houston, TX, burning from March 17-20 in 2019. As a result, several nearby recreation sites were closed to the public for a period of time while oil products and toxic firefighting chemicals drained into the Houston Ship Channel, prompting response and cleanup efforts. In light of other recent work on MD (Duff et al., 2025) we believe the ITC incident to be a setting that is quite well suited for a MD-based analysis of changes in recreation; the incident occurs in a relatively high-quality era for MD, and some of the larger affected recreation sites have helpful characteristics supporting MD analysis. We used a dataset purchased from Unacast[3] in late 2020.

However, our analysis also highlights some of the troubling limitations of a strategy that relies heavily on MD. A key lesson from this analysis is that MD requires careful consideration of the "coverage rate" - the fraction of actual visitors to a site who are observed in the mobility dataset. This relationship varies over time and across sites of interest, and its use requires reliable "reference data" which measures true visitation. At each step in our analysis, we discuss how features of MD require the researcher to make important assumptions, often with plausible alternatives, based on limited information.

We present methodology and results for calculating lost recreation value at sites affected by the ITC fire, with a focus on three key parameters: the number of lost trips, the expansion factor (inverse coverage rate) relating MD counts to true visitation, and the site-specific value of a lost trip. We use MD counts to estimate lost trips based on the predicted counterfactual trips that would have taken place but for the incident. We used three sources of non-MD reference data to extrapolate from the MD lost trips to true totals. Each source has interesting limitations for its compatibility with MD analysis, which we discuss. For valuation, we use MD-derived home origin location, visible to us only at the level of zip code tabulation area[4] (ZCTA), to estimate site-specific zonal travel cost (ZTC) models. The product of these three key parameters equals the lost consumer surplus from the incident. In Section 4.4, we discuss how each parameter contributes to a range of results following from a set of reasonable alternative assumptions.

We conclude with a discussion of the future role for MD in applications such as NRDA recreational losses. We suggest tempered expectations for researchers hoping to use MD to identify changes in recreation resulting from changes in environmental quality. While the use of MD may be suitable for applications that tolerate high variance results, this particular objective may require an unrealistic level of precision and stability. For NRDA in particular, where results are often highly contested, MD may not yet be appropriate except when some key features of the setting are favorable for analysis, likely a matter of luck. We remain hopeful that the MD industry's transparency and researchers' understanding of these data will improve so that they can be more reliably incorporated into damage assessment and similar analyses.

2. Data

2.1 Mobility data

Our vendor, Unacast[5], provided approximately 5 million device-activities recorded in the 20 months of January 2019 - August 2020, from 484,000 unique devices within the site boundary shape files we provided. Site boundaries were drawn manually, using ArcGIS Pro, in reference to ESRI base maps, publicly-available tax parcel data, and Google Street view. The dataset's unit of observation in the MD is an "activity," an algorithmically-defined (proprietary to Unacast) aggregation of consecutive, clustered location “pings” from a smartphone app. Individual pings from a given device are combined into “activities” based on proximity (pings within a certain distance of the first in the cluster). Our activity dataset consists of rows containing a centroid latitude and longitude, as well as a start and end time, along with a handful of other covariates such as “bump count” (the number of constituent pings in an activity). Unacast classifies activities into categories like “area dwell” and "travel." We filtered out activities like “travel” (where implied speed of movement is relatively high) because we are interested in activities that Unacast believes to be stationary visits to the location, not devices merely passing through. The mobility dataset also includes device-level estimates of home origin[6] ZCTA which we use in the site valuation model. We observe estimated home origin ZCTAs for about 95% of observations. As an example count time series, Figure 1 shows the pattern of daily counts of unique devices at one of the largest sites, Sylvan Beach in La Porte, TX, with the shaded box indicating the period of time when closed to the public. When open, use patterns are consistent with expectations for sites of the same type- higher average use in the summer months, peak use around warm-weather holidays, and recurring higher use on weekends rather than weekdays.

<Figure 1 here>

Timestamps and durations associated with activities in the MD also allow us to visualize the finer scale within-day pattern of use. For example, in Figure 2, Sylvan Beach appears to follow a plausible daily pattern of use over the sample period we observe. We see a higher volume of use on the weekends, peak use in the late afternoon, and very little use outside of daylight hours. Visualizations like this help alleviate some concerns about what was being measured in this particular dataset and do not present any major red flags.

<Figure 2 here>

In Appendix Figure A1, each recorded *activity* observation is an average position of individual location pings aggregated from clusters of pings that are neighbors in space and time. While many providers (including Unacast) share the basic principles of this aggregation approach, the specific details and calculations are unknown to the end user. As with all mobility datasets of this format, data locations are recorded with some degree of imprecision. Figure A1 is a visual characterization of the spatial error as we observe it. This is a plot of MD activities on a pier at Sylvan Beach, over a subset of the time frame. It is possible that some of the activities in the water are actually people swimming or in boats, but it is likely that many of these apparent activities in the water are actually activities occurring on the pier, but recorded with spatial error. Plots like Figure A1 provide visual qualitative evidence for the concern researchers should have about spatial error in MD. This degree of concern should inform the definition of areas of interest (AOIs) for study sites. It seems likely that more caution is required when recreation occurs near significant sources of non-recreation MD activity. Sites near the boundary of high-volume transportation or busy commercial activity are especially susceptible to this spatial error concern.

All of our shapefiles are drawn with reference to ESRI base maps, administrative parcel shapefiles, and Google Street view. Our initial data queries also include a broader geographic extent than in the strict definitions we end up using. Data points in this buffer zone allow us to visualize the relative volume of potential visitation near the borders of sites. This practice can help identify spots in which it is especially important to carefully exclude observations with questionable validity.

2.2 Suitability of sites for MD analysis

What makes a site a good candidate for application of MD? Our dataset covers several sites subject to closures during the ITC incident, shown here in Table 1. We focus our analysis in this paper on a subset which we consider most appropriate for a MD-based event analysis. We provide a brief overview of the excluded sites in the appendix. In this section, we discuss some characteristics of these sites that indicate a relatively favorable setting for MD analysis.

<Table 1 here>

It is helpful to have consistent and substantial visitation counts for MD analysis so that there is sufficient basis to estimate the count model covariates. Because MD typically only captures a small fraction of the true visitation, it is important for a site to have enough MD observations to fill out a statistical model. The sites we present in this paper all average at least 14 unique MD observations per day, with the exception of Baytown Nature Center (BNC) with 7.3 per day. Because BNC comes with agency data that is similar in quality and format to that of Battleship Texas and San Jacinto, we include it in our analysis despite its relatively low visitation. Small sites with low average visitation may not show enough variation to detect changes in use.

Physical characteristics of the sites can also be important. As mentioned above, one important consideration is potential interference from proximity to sources of spatial error (e.g., homes, businesses). San Jacinto is a particularly promising site where proximity to spatial error sources is concerned, bounded mostly by water and undeveloped public lands. We also consider potential interference from nearby auto traffic. Sites that are only bounded by neighborhood roads or site-specific entrances (such as BNC) are promising, while others are problematic because of their proximity to busy roads. Our understanding of this provider's algorithm[7] is that it attempts to separate "moving" activities from "stationary" activities, yet we observe a considerable number of supposedly stationary events that take place on and around major roads, especially where there are busy intersections. Furthermore, we evaluate sites according to their geographic footprint. Large, roughly circular sites provide for the greatest proportion of area that is interior to the site, and are thus more insulated from false positives due to spatial error. San Jacinto again is strong in this characteristic in particular because its main attraction (a monument with a museum) sits near the center of the very large site, where MD observations are very likely to be generated by true visitors.

Additional factors that influence site suitability for this paper's investigation include the availability of reference data. It is also important to consider unusual patterns such as the presence of incident response workers (firefighters, hazardous materials clean crews, etc.) that may replace lost recreator observations with large numbers of non-recreation observations that are indistinguishable to the analyst.

2.3 Control variables

Following standard approaches to recreation data for NRDA as referenced above, we use standard controls for conditions that influence recreation. These include weather variables (daily maximum temperature bins, average wind, and precipitation) from the National Centers for Environmental Information (NCEI), and indicator variables for day-of-week, year, and holidays. We did not include month indicators in our main specifications because of the confounding timing overlap with the start of the Covid pandemic, which began almost exactly one year after the ITC incident. Following Csomós et al. (2023), we expect an important behavioral response to Covid, especially with respect to the proportion of time spent indoors and outdoors. We use a simple binary indicator for Covid, which started on March 16, 2020.

Zonal travel cost control variables are derived from American Community Survey (ACS) ZCTA-level data for racial composition (proxied by percent white), gender composition (proxied by males per female), and median age. Following Leggett et al. (2018), the travel cost variable is constructed as a function of household income data, also from ACS. We model travel cost at the level of a site-origin ZCTA pair, with a direct mileage cost component and an indirect time cost component (see Section 3.3 for details). Travel distance and travel time are calculated at the ZCTA level using Route Analysis[8] in ArcGIS Pro.

Appendix Table A1 describes the cumulative distribution of travel distances, considering ZCTAs according to bins of one-way driving distance of every affected site. In the full mobility dataset, we observe a substantial portion of visitors appearing to travel from implausibly long distances. As will be discussed in Section 3.3 we do not observe any information about a visitor's primary purpose, therefore we frame the ZTC model in terms of assumptions about valid travel distance cutoffs. Consistent with the ZTC model, the cutoff distances in Table A1 are relative to the set of ZCTAs that are within X miles of *every* site in the analysis. For example, 52% of visits to Battleship TX came from visitors with an origin ZCTA that is within 50 miles of each of the five sites. 24% of visitors to Battleship TX came from a ZCTA that is 150+ miles from at least one of the sites in the analysis. We chose to organize the ZTC analysis in this way so that each distance cutoff contains an identical set of contributing ZCTAs for each site. As we discuss in Section 3.3, we chose 150 miles to represent the farthest likely one-way driving distance for a day trip to one of these sites, and thus it sets an upper bound for which MD observations to consider as valid for the purpose of ZTC valuation analysis.

2.4 Reference data used to derive coverage rates

MD typically only captures an unknown fraction of the true number of visitors. We describe the relationship between MD counts and true visitation counts as a coverage rate. This is a key parameter for any application of MD where total use is an object of interest, because it has a direct multiplicative effect on total use. Our approach to modeling the coverage rate here is to treat it as a static property of the MD that varies across sites but not across time. In reality, coverage rates almost surely vary across time and are complicated by the dynamics of recreational behavior coupled with mobile device/app use, and are further complicated in data processing by the MD vendor. Our static coverage rate assumption allows us to focus on variation in the coverage rate across sites and across reference data sources. As discussed in Duff et al. (2025), these data appear to exhibit increased variability in coverage rates starting sometime after 2020 (likely influenced by structural factors across the MD industry), whereas the time range of this dataset (January 2019- August 2020) seems to be a relatively stable window for MD and its relationship to actual recreation.

The available reference datasets include limited site/incident specific parking lot videography-based counts, existing management agency entrance and parking data, and counts derived from satellite imagery of parking lots. The use of existing agency data is a standard practice in NRDA recreation studies whereas the satellite imagery and videography are more novel approaches. All of these reference datasets share characteristics of lower time frequency (relative to MD) but better documented and understood data generating processes. We expect these will be common characteristics of reference data when working with MD for outdoor recreation. For the purpose of this analysis, we are agnostic as to the relative reliability of these reference data sets and treat each as though it is a potentially unbiased signal of total visitation for the purposes of estimating coverage rates.

In Table 2, we report the estimates of coverage rates for each site by reference dataset. Each calculation reflects the aggregate ratio of MD counts to reference data counts across multiple dates at each particular site. In Bay Area Park and Sylvan Beach, we lack agency data for visitation, so we report the average agency-data estimate from the other three sites (5.8%).

<Table 2 here>

Videography data of parked cars at affected sites were collected shortly after the 2019 incident. Research staff collected videography data at a set of potentially affected sites during the aftermath of the spill and recorded the number of parked cars observed in the video imagery. The timestamps on these manual counts were recorded imprecisely as multi-hour shift windows instead of precise snapshots. Thus, the calculation of the coverage rates from these counts requires a similar assumption to the one made for satellite reference data (see below), that MD durations and timestamps are accurate[9] and representative of the overall visitation.

The second category, which we summarize as “agency data” was derived from daily traffic counters (for the Battleship TX and San Jacinto SHS sites) and daily entrance counts at Baytown Nature Center (BNC). These are both familiar and traditional recreational data sources. For the purposes of calculating a MD coverage rate, it is easy to argue that they are appropriate for the sites that they measure. However, because they only cover a subset of the affected sites, we must make important assumptions about the *portability* of coverage rate calculations across sites. As discussed in Duff et al. (2025), there are numerous factors that cause the coverage rate to vary over space (e.g., visitor demographics or characteristics of the site that affect usage of smart devices). In particular for this analysis, we need to assume that site definitions are established in a way that captures the same spatial error dynamics. Count model results will be particularly sensitive to the shape and extent of AOIs when coverage rates are ported over from another site. This is not a problem for sites with reliable reference data because a more generous site definition (including additional activities near the boundary) will simply result in a higher coverage rate and the total use should be invariant.

Finally, satellite imagery provided an interesting, but ultimately limited picture of visitation due to the low number of images. We acquired Maxar[10] satellite imagery through the USGS satellite imagery aggregation service, which provides access to an existing library of timestamped images. Even though the frequency of collection is low, irregular and unpredictable, satellite imagery is an appealing candidate for calibrating MD. Here, we count cars in parking lots associated with the recreational sites. We calculate coverage rates based on the assumption of 2.5 people per car[11], and extrapolate from snapshots to daily counts based on the number of MD activities that overlapped the time of the snapshot. We use variables from MD indicating timing and duration of the activity to estimate the number of unique visitors that would have been present at the time the satellite image was taken, a subset of “unique daily visitors”. The quality of these coverage rate calculations depends on a key assumption, that the timing and duration variables recorded in the MD need to be accurate and consistent across observations. As discussed in Section 2.1, these timestamp variables in aggregate appear to produce a reasonable time-of-use profile for outdoor recreation sites, but this assumption is ultimately hard to validate. It is likely, for instance, that duration measurements are a conservative[12] estimate of the timeframe of attendance.

3. Analytical Framework

3.1 Damage formula

Natural Resource Damage calculations for recreation typically require two primary inputs, the quantity of trips impacted and the financial measure of the average lost consumer surplus per trip (English et al., 2018). Here, we consider only lost trips and derive a site-specific value per trip so that damages at site i are given by:

$$Lost\ Surplus_i = L_i \times V_i$$

Lost trips, L_i are the difference between baseline trips (those counterfactual trips that would have occurred but for the contamination), B_i, and actual trips, A_i. To relate these counts to the scale of true recreation, a site-specific coverage rate (cr_i) is used to transform the results of the MD counts into trips:

$$L_i = B_i - A_i = \frac{b_i - a_i}{cr_i}$$

As described below, the b_i are predicted counterfactuals from a Poisson count model, at the scale of MD counts. In the following section (3.2), we explore multiple interpretations of the a_i variable, observed MD counts during closure periods, as well. Consumer Surplus values, V_i, are derived from a Zonal Travel Cost Model applied to the MD home origins and destination pairs as detailed below.

3.2 Poisson count model for visitation

For each site we estimate site-specific Poisson count models using MD, with respect to closure status, as a function of covariates such as weather, season, and weekend/weekday/holiday dummies. As discussed above, we use a simple binary indicator for Covid, beginning March 16, 2020. Section 4 reports regression results and discussion. The individual site regression results generally conform to expectations about the effect of common covariates for outdoor recreation.

To estimate lost trips, we use MD alongside covariates to predict a counterfactual baseline, predictions of the MD value during the incident dates and conditions, as if the incident had not occurred. We use a simple approach for extrapolating MD to total visitation. We produce a series of average coverage rates (Table 2) that vary by reference data source and by site. We then divide the resulting MD-based lost trip estimate by its corresponding site-specific coverage rate (or multiply by its expansion factor) to expand to the total.

For practical reasons discussed in Section 2.4, we make the simplifying assumption that the coverage rate is constant over time. As discussed in Duff et al. (2025) this assumption is likely more supportable in the time frame of this ITC incident analysis (2019-2020) than in the 2021-2024 range, but it is still a consequential assumption. In order to extrapolate from MD-based lost trip estimates to total lost trips, we must assume that the available coverage rate calculations specifically are accurate during the closure period when we calculate losses.

A deeper methodological question arises in interpreting MD observations recorded within the boundaries of the park during closure periods; which observations are recreation and which are something else? On many of the closed park days, we observe attendance that is meaningfully greater than zero. We discuss a few possible reasons for this:

- Category 1: People who ignore the closure and recreate at the site;
- Category 2: People who work in an official capacity within the closed parks (e.g., regular staff or non-regular incident response staff);
- Category 3: People appearing at the site due to MD-specific quirks like spatial measurement error.

People in category 1 are generally considered in NRDA practice to be valid recreators and thus should ideally be removed from the loss calculations. People in category 2 should not be considered recreators, but they may appear in a mobility dataset in both closure and non-closure periods. We hypothesize that the prevalence of people from category 2 in MD may be higher during the closure periods because the incident leads to increased presence of incident response staff. Ideally, we would be able to observe the purpose of a visitor's trip to the site, but that is not possible with anonymized MD, and aggregate data about staffing for incident response are generally not available. People in category 3 may also appear in the MD sample in different proportions during or outside of the spill period. MD counts for each site include some proportion of false positives and false negatives during normal times. When a site is physically closed, it is likely that the proportion shifts toward more false positives because visitors are relatively more likely to be on the outside of the boundary rather than the inside[13].

Because the purpose of trips is unknown for those we observe in the MD, we must make an assumption about the relative contributions of these explanations. If sites are not physically possible for general recreators to enter during the closure period (zero people in category 1), then it might be reasonable to assume that any remaining MD observations during the closure period are attributable to non-recreators. However, in order to fully discount these MD observations, we must further assume that the net contribution of people in categories 2 and 3 accounts for the difference. This requires information on the proportion of non-recreators across closure and non-closure periods, which is not generally available. On the other end of the assumption spectrum, an alternative approach would treat the visitation during closure periods as being fully equivalent to real recreation, and thus rightfully subtracted from the counterfactual when calculating lost trips. It is likely that these two extreme assumptions bracket the truth. In Section 4 we report the results from both methods and discuss how they fit into the plausible range of lost trips.

3.3 Zonal travel cost model

Average consumer surplus per trip to a site is based on the data provider's estimate of device-level home origin ZCTAs. Our understanding is that the provider calculates home origin coordinates for most devices in the sample based on the device's most common location at night (see Section 2.1 for more details). ZCTAs are matched to users at the time that the dataset is produced (September 2020) and they do not vary over the time frame, and thus may be less accurate in the early dates. We report valuation estimates based on aggregate trip rates for the entire time frame. The zonal travel cost model for trip rate from ZCTA *j* to site *i* is given by:

$$triprate_{i,j} = \beta\ (travelcost_{i,j}) + \phi\ (demographics_j) + \varepsilon$$

where the travel cost variable[14] is computed for each origin-destination pair:

$$travelcost_{i,j} = 2 * [\ (distance_{i,j} * \$0.2762) + (1/3) * (traveltime_{i,j} * hourlywage_j)\ ]$$

The ZCTA-level zonal travel cost model follows the simple, single site model described in Leggett et al. (2018) and based on Hellerstein (1991) and Hellerstein and Mendolsohn (1993) where we can estimate the value of a single trip as equal to the negative inverse of the coefficient on the travel cost parameter:

$$value = -1/\beta_{cost}$$

We calculate travel distances and driving times using Route Analysis in ArcGIS Pro for ZCTA-origin pairs. As shown in Appendix Table A1, there is a substantial portion of trips that appear to be traveling from distances greater than 150 miles. Because we do not observe any information about the primary purpose of individual trips, we choose a cutoff distance as a conservative assumption that avoids trips originating from ZCTAs that were ‘too far’ away to plausibly have a primary purpose to visit these parks. Because this study is motivated by the NRDA context, where settlement negotiations can involve a wide parameter space, we retain a range of assumptions that can be plausibly supported. Parsons (2017) suggests that 3 to 4 hours of one-way travel distance is a reasonable upper bound for recreational day trips, which is conservatively in line with 150 miles at the top of the range. We consider three versions of the ZTC model with cutoffs at 50, 100, and 150 miles. Only ZCTAs within 50, 100, or 150 miles of *all sites* are included, so the ZTC model for each site contains the same set of origins[15]. By excluding trips beyond a given distance we are effectively setting a value to zero for those longer distance trips.

There are potential complications with this approach. In order to estimate a ZTC model based on MD, we discuss three complications and corresponding assumptions to be made. First, device owners move to a new home, and the timing of moves could lead device-level ZCTAs to be outdated. When a device owner changes origin ZCTAs between a recreation trip and the date their origin ZCTA is recorded, we are ignoring the possibility that a device owner’s travel costs change over the period of analysis. A conservative choice for cutoffs (e.g., 50 miles) helps support this assumption, as people who move to a new home that is over 50 miles from the site they previously visited will not be included in the model. The aggregated nature of the MD home origins does not allow us to examine the realism of this practical assumption. It is possible that vendors could provide additional information on home origin calculations that might help future researchers diagnose the impacts of this assumption.

Second, attribution of travel costs to the value of an individual park trip is questionable given lack of data on purpose of trips. We assume that the distribution of origin ZCTAs for visitors must be representative of the distribution of origin ZCTAs for visitors *whose primary trip purpose was recreation at the site*. The choice of distance cutoffs for alternative ZTC models supports this assumption, reducing the influence of the most important source of outliers, visitors who live far away and who visited these sites as an incidental part of a different trip (e.g., visiting a park during lunch on a work trip from far away). Without using the cutoff assumption, we would be estimating a ZTC model for nearly the entire US, with a non-negligible fraction of “visits” coming from thousands of miles away.

Third, travel distances from different points within a given ZCTA can vary, which introduces measurement error. We assume that the centroid of the ZCTA is representative of the typical origin point within it. It is reasonable to assume this might lead to a slight overestimate of the average travel distance because recreators prefer nearby sites, all else equal. However, we cannot quantify this because home origins are anonymized and aggregated to the ZCTA level.

4. Results

4.1 Count model

Table 3 shows the results from our main Poisson count model measuring the effect of the site closures on mobility data counts of recreation.

<Table 3 here>

We include a set of covariates that are known to influence outdoor recreation: holidays, day of the week, season, year, precipitation, wind, and maximum daily temperature (Tmax). We model the contribution of temperature in bins (with Tmax <60 degrees Fahrenheit as the omitted bin), reflecting potentially non-linear preferences. As discussed in Section 2.3 we omit a dummy for month in order to avoid an over-weighted comparison between the closure period and the first weeks of Covid in the following year. We report heteroskedasticity-robust standard errors in this table.

The key variable of interest is “closure”, which is strongly negative for all sites as expected, reflecting significantly decreased use during the closure period following the spill. We use site-specific estimates to model counterfactual baseline visitation for each day during the closure period. These estimated loss percentages are the basis for the “Partial Loss” calculations in Section 4.4, whereas they are set to 100% for the “Full Loss” calculations.

4.2 Lost trips

Appendix Table A3 reports the number of lost trips for each site, defined as the difference between the predicted counterfactual and the observed counts. The results in Table A3 are expressed in terms of MD observations, and are later re-scaled using the inverse of the coverage rates as shown in Table 2. As discussed in Section 3.2, this estimate of percent loss must be interpreted carefully with respect to the possible causes of the observed trips in the MD counts. As we discuss in Section 4.4, there is significant variation in the implied percent loss across these five sites.

4.3 Valuation of lost trips

Our zonal travel cost model estimates use a simple distance-based criteria for selection of ZCTAs as discussed in Section 3.3. In Table 4 we show the results for the most inclusive model in the range of cutoffs we considered, with a cutoff of 150 miles. As one might expect, the travel cost parameter is statistically significant and negative for each site. Only the travel cost parameter estimate is used in the derivation of the site-specific trip valuations.

<Table 4 here>

These dollar-value results (in 2019 dollars) are generally comparable to those found in the literature for sites with similar features. Visits to the first two sites are valued at a higher level than the rest, a plausible result considering the unique historical interest. The other three sites are heavily used by people who live nearby, but not as much by people who incur large costs to travel there. As the choice of distance cutoff increases, additional farther visits are included, which increases the calculated site valuation.

4.4 Sensitivity analysis of total lost surplus

The three key parameters in this assessment are 1) number of lost trips measured in MD counts, 2) the expansion factor relating these counts to totals, and 3) the value per trip. The final result for the NRDA context is lost value, the product of all three quantities. In this paper we have identified a range of plausible estimates for each of these parameters, focusing on some of the major assumptions that are required to support a particular calculation. As shown in Section 3.1:

$$Lost\ Surplus = \sum_i \left[L_i \times V_i \right] = \sum_i \left[\left(\frac{b_i - a_i}{cr_i} \right) \times V_i \right]$$

For the first parameter, we consider both ends of the plausible range in Table A3 - either we assume that all observed trips during the closures are valid recreation (lost = predicted - observed; “full loss”) or that all observed trips during closures are to be discounted (lost =

predicted, and a_i defined as 0; “partial loss”). For the second parameter, each site has three plausible options for expansion from raw MD counts (which only observe a small fraction of the likely total) based on different reference datasets. For the third parameter, we consider three zonal travel cost model specifications based on different distance cutoff assumptions - 50, 100, or 150 miles. Taken all together, these options produce a range of 18 different estimates for the total lost surplus across these five sites. Using the range of values shown in Tables 2, A3, and A4, we can produce Figure 3, below.

<Figure 3 here>

Figure 3 shows the range of results for total lost surplus grouped by assumption. The first six columns all assume a 50-mile cutoff for the zonal travel cost model, etc. Within each cluster of six, the first set of two use a coverage rate assumption derived from videography data, the second set of two use agency data, and the third set of two use satellite data. Within each set of two, the first is an estimate using the “full loss” assumption, and the second is the “partial loss” assumption. The largest total lost surplus result was $1,997,954, and the smallest was $346,016, for a ratio of 5.77:1.

The harmonic mean (across the five sites) ratio between the high-end assumption and the low-end assumption for the lost trips parameter is 1.37, compared to 1.68 for the travel cost cutoff assumption, and 2.16 for the coverage rate multiplier assumption. Thus, the greatest degree of variability in this exercise is contributed by the calculation of the coverage rate (Table 2). Ideally, with enough observations of high-quality reference data, there would be a greater degree of convergence across methods. But as discussed in Section 2.4, reference data, as is often the case in outdoor recreation, is unfortunately limited in scope and reliability, so we are left with a relatively broad range of plausible coverage rates. This highlights an important takeaway from our experience using MD to support an NRDA, that the relationship between MD counts and true recreation (as proxied by on-the-ground reference data) is of paramount importance. The coverage rates derived in this project appear to be fairly typical (around 6% in aggregate) of those we have seen elsewhere, although there can be significant variation across sites and over time (as discussed in Duff et al., 2025).

The two assumptions we explore on the lost trips parameter (full vs partial loss) lead to the smallest variation in results (among the three parameter sensitivities we explore). However, the actual values we see in Table A3 are widely dispersed across sites, but it is not immediately clear why. All five sites appear to be able to enforce a closure, either through sliding gates at entrances, fences surrounding the bulk of the site, or both. We might then expect that the true closure effect would be similar across sites. The wide variation in the percent reductions during closures suggests that either these sites are not able to enforce closures at the same level of effectiveness or that many of these MD observations are not recreation related (e.g., facilities or response workers). One potential explanation of the variation across sites is that Battleship TX and San Jacinto were directly adjacent to the actual tank fire site, and thus the closure-period visitation might be influenced by a larger number of non-recreators working in the response/cleanup effort. This explanation would support the “full loss” assumption, but ideally, we would have more supporting information on the timing and location of cleanup efforts before relying heavily on such an assumption.

Finally, the ZTC distance cutoff assumption remains a challenging methodological hurdle for analyses relying on MD. As expected (though not mechanically forced), we see that longer distance cutoffs result in higher values. This difference is most pronounced for Battleship TX and San Jacinto, the two sites with historical museum amenities and the highest calculated trip values. As shown in Table A1, these sites also appear to have the highest proportion of visits coming from farther than the longest cutoff of 150 miles. The challenge in making a distance cutoff assumption is to balance the possible exclusion of true visits that are willing to drive a long distance against the possible inclusion of non-recreation visits that are derived from imprecision in the MD or from the complication of multi-purpose trips. Without additional data (e.g., site-specific visitor surveys collected in person), the best researchers can do is describe the results as they vary across potential assumptions.

5. Conclusion

There are several reasons to be hopeful about MD for NRDA human-use assessments. It offers a high-frequency, spatially detailed sample of site visitation, providing opportunities to measure recreation that has already happened, for a relatively low cost compared to traditional methods. The specific data product we evaluated also contains helpful variables which can be used to estimate travel cost models and understand the time of use profiles at any given site. Our evaluation of the data quality for this setting produced sensible results, and we did not perceive that the MD product demonstrated any major red flags in and of itself.

However, our analysis also highlights some significant drawbacks to the technology which must be considered. We find that the relationship between MD counts and those derived from reference data is highly variable across sites of interest and estimation methods. We also find that using MD to estimate the counterfactual number of lost trips requires making important assumptions (e.g., proportion of non-recreators observed during closures) about what exactly is being measured. Similarly, the opacity of the MD generating process requires caution in the construction and interpretation of valuation model parameters. At each point in the process, we are confronted with fundamental uncertainties about how MD relates to the actual recreation we hope to measure, and we rely on creative but unproven assumptions to estimate changes in recreation. We find that the three major quantities of interest (coverage rate, lost trips, and value of a lost trip) vary significantly with respect to the choice of simplifying assumptions we make about the MD. In each case, there are a variety of available assumptions, and the final result, the product of the three key parameters, is highly sensitive to these assumptions.

Future work on this topic should investigate how MD performs across space, time, and providers. We can apply MD to contexts that are potentially rich in high-quality reference data, such as MLB stadiums as in Duff et al. (2025), to better understand the data generating process. This could better inform some of the important assumptions that are required to rely on MD for practical settings such as the one in this paper where the reference data was sparser. MD researchers should seek out settings in which reliable reference data can be leveraged to improve understanding of dynamics of this data source. In settings like NRDA, MD should be used in parallel with traditional NRDA data collection to study the relationship between MD and traditional data sources like ground-based and aerial counts. Until MD are well understood, they will require companion data whose accuracy and dynamics are better established.

For now, we recommend caution for NRDA-style assessments employing MD. High quality reference data is always required to validate, contextualize, and extrapolate the MD to totals. Calibration of MD with limited field data collection may be feasible but the reliability of such approaches is unknown. Even when high quality reference data are available, there may be important confounding factors that limit the usefulness of MD in a complementary role (e.g., unclear portability of MD coverage rates across sites or time frames). We are cautiously optimistic about the usefulness of MD for some less ambitious NRDA tasks in the near term, such as research designs that can accommodate a higher level of MD aggregation. In general, small sites and short time frames are more challenging for MD analyses, and some MD vendors have more recently started to advise us to work with bigger site shapefiles and longer periods of aggregation (e.g., weekly or monthly). Spatial patterns that are predictable across time are a core part of the most common business applications of MD. These repeated patterns may be the most fruitful near-term area to use MD for recreation measurement. Resource management data on visitation is often collected at a high level of spatial aggregation. Repeated MD spatial patterns may provide insight into apportioning existing highly aggregated data across smaller spatial areas. Standardization and transparency in data processing rules by data providers may be an important step in support of wider acceptance of the reliability of these data.

6. References

AAA. 2019. *Your Driving Costs: How Much Are You Really Paying to Drive?* Heathrow, FL: AAA. https://exchange.aaa.com/wp-content/uploads/2019/09/AAA-Your-Driving-Costs-2019.pdf.

Athos/Delaware River Lost Use Technical Working Group. 2007. “Athos/Delaware River Lost Use Valuation Report.” Prepared for the Athos/Delaware River Trustee Council, March 29. Modified September 2008. Available at: https://www.cerc.usgs.gov/orda_docs/DocHandler.ashx?task=get&ID=175

Breffle, William S., Edward R. Morey, Robert D. Rowe, Donald M. Waldman, and Sonya M. Wytinck. "Recreational fishing damages from fish consumption advisories in the waters of Green Bay." Stratus Consulting, Boulder, Colorado (1999).

Byrd, H., E. English, D. Lipton, N. Meade, and T. Tomasi. "Chalk Point Oil Spill: Lost Use Valuation Report." National Oceanic and Atmospheric Administration, US Fish and Wildlife Service, Maryland Department of the Environment, Maryland Department of Natural Resources, Prepared for the Chalk Point Trustee Council (2001).

Bouchard B-120 Oil Spill Lost Use Technical Working Group. 2009. “Bouchard B-120 Oil Spill Buzzards Bay, Massachusetts Lost Use Valuation Report.” March 31. Available at: https://pub-data.diver.orr.noaa.gov/admin-record/6406/Bouchard%20B-120%20Oil%20Spill%20Lost%20Use%20Validation%20Report%2C%20March%2031%2C%202009.pdf.

Chan, Nathan W., and Casey J. Wichman. "Climate change and recreation: evidence from North American cycling." Environmental and Resource Economics 76, no. 1 (2020): 119-151.

Chapman, David J., and W. Michael Hanemann. 15. *Environmental damages in court: the American Trader case*. Edward Elgar Publishing, 2001.

Csomós, György, Endre Márk Borza, and Jenő Zsolt Farkas. "Exploring park visitation trends during the Covid-19 pandemic in Hungary by using mobile device location data." Scientific Reports 13, no. 1 (2023): 11078.

Deacon, Robert T., and Charles D. Kolstad. "Valuing beach recreation lost in environmental accidents." Journal of Water Resources Planning and Management 126, no. 6 (2000): 374-381.

Domanski, Adam, and Jason Murray. 2017. "Texas City Y Oil Spill Natural Resource Damage Assessment: Technical Memorandum: Estimate of Lost Recreational Use Damages." Natural Resource Trustees (NOAA; USFWS; NPS; TPWD; GLO; TCEQ), September 28. Available at: https://www.cerc.usgs.gov/orda_docs/CaseDetails?ID=1071

Duff, Cameron, Alice Antia, Eric Horsch, James Hilger, Jason Murray, Nathaniel Merrill, Samantha Winder, Spencer Wood, Matthew Zafonte. "Evaluating mobility data for recreation monitoring: Diagnostics and implications for best practice." Forthcoming, Land Economics (2026)

English, Eric. "Recreation nonparticipation as choice behavior rather than statistical outcome." American Journal of Agricultural Economics 90, no. 1 (2008): 186-196.

English, Eric, Roger H. von Haefen, Joseph Herriges, Christopher Leggett, Frank Lupi, Kenneth McConnell, Michael Welsh, Adam Domanski, and Norman Meade. "Estimating the value of lost recreation days from the Deepwater Horizon oil spill." Journal of Environmental Economics and Management 91 (2018): 26-45.

Gourieroux, Christian, Alain Monfort, and Alain Trognon. "Pseudo maximum likelihood methods: Applications to Poisson models." Econometrica: Journal of the Econometric Society (1984): 701-720.

Hellerstein, D. 1991. Using Count Data Models in Travel Cost Analysis with Aggregate Data. *American Journal of Agricultural Economics* 73, 860-866.

Hellerstein, D. and R. Mendelsohn. 1993. A Theoretical Foundation for Count Data Models. *American Journal of Agricultural Economics* 75, 604-611.

Horsch, Eric, Michael Welsh, and Jason Price. "Best Practices for Collecting Onsite Data to Assess Recreational Use Impacts from an Oil Spill." (2017).

Lawson, T. V. "The wind content of the built environment." *Journal of Wind Engineering and Industrial Aerodynamics* 3, no. 2-3 (1978): 93-105.

Leggett, Chris, Eric Horsch, and Mark Curry. 2018. “Recreational Camping Damages Due to the Refugio Beach Oil Spill.” *Prepared for Refugio Beach Oil Spill Natural Resource Damage Assessment.* Cambridge, MA: Industrial Economics, Inc.

MacNair, Douglas J., and William H. Desvousges. "The economics of fish consumption advisories: insights from revealed and stated preference data." Land Economics 83, no. 4 (2007): 600-616.

Morey, Edward, Jennifer Thacher, and William Breffle. "Using angler characteristics and attitudinal data to identify environmental preference classes: a latent-class model." Environmental and Resource Economics 34 (2006): 91-115.

Parsons, George R., Ami K. Kang, Christopher G. Leggett, and Kevin J. Boyle. "Valuing beach closures on the Padre Island national seashore." Marine Resource Economics 24, no. 3 (2009): 213-235.

Parsons, George R. “The Travel Cost Model.” A Primer on Nonmarket Valuation, edited by Patricia A. Champ, Kevin J. Boyle, and Thomas C. Brown, pp. 187-233. Dordrecht: Springer, 2017.

Smolak, Kamil, Katarzyna Siła-Nowicka, Jean-Charles Delvenne, Michał Wierzbiński, and Witold Rohm. "The impact of human mobility data scales and processing on movement predictability." Scientific Reports 11, no. 1 (2021): 15177.

Von Haefen, Roger H., Daniel J. Phaneuf, and George R. Parsons. "Estimation and welfare analysis with large demand systems." Journal of Business & Economic Statistics 22, no. 2 (2004): 194-205.

Table 1: Key statistics for sites affected by the ITC Tank Fire

Site	Closure duration (days)	Acres	Average daily MD count
Battleship Texas State Park	52	76	14.6
San Jacinto Battleground State Park	52	1022	16.2
Bay Area Park	14	58	27.5
Sylvan Beach Park	24	62	58.6
Baytown Nature Center (BNC)	4	546	7.3

Notes: Closure duration data comes from state agency and local media sources. Acres are calculated in ArcGIS Pro. Average daily MD count is the number of unique devices observed visiting a site during the duration of the dataset – a sample reflecting a fraction of the total.

Table 2: Coverage rate calculations using mobility and reference datasets

site	videography	agency data	satellite
Battleship TX	8.3%	5.3%	9.5%
San Jacinto	15.9%	4.1%	3.6%
Bay Area	5.3%	5.8%	7.3%
Sylvan Beach	11.3%	5.8%	4.9%
BNC	5.0%	7.9%	2.9%

Notes: Coverage rate is the estimated average proportion of visitors that appear in the mobility dataset, compared to a reference dataset. The average is calculated as the ratio of the sum of all mobility data counts to the sum of all reference data counts over the same set of dates for which data was available. Instances of “5.8%” are imputed from the simple average of other values in the same column due to lack of agency data.

Table 3: Count model results - Poisson specification - mobility data

	Battleship TX	San Jacinto	Bay Area	Sylvan Beach	BNC
closed	-1.042***	-1.181***	-3.550***	-1.781***	-0.712***
precipitation (in)	-0.043	-0.053	-0.399***	-0.174***	-0.075
wind (avg mph)	-0.002	-0.003	-0.009	-0.015*	-0.006
Tmax 60-70 ℉	0.056	0.195	0.338*	-0.066	0.018
Tmax 70-80 ℉	0.051	0.117	0.386**	0.134	0.121
Tmax 80-90 ℉	0.540**	0.512***	0.453**	0.529***	0.256*
Tmax >90 ℉	0.703***	0.476**	0.528***	0.655***	0.194
Monday	-0.741***	-0.535***	-0.503***	-0.654***	-0.668***
Tuesday	-0.834***	-0.622***	-0.547***	-0.689***	-0.764***
Wednesday	-0.872***	-0.689***	-0.488***	-0.684***	-0.719***
Thursday	-0.767***	-0.515***	-0.543***	-0.586***	-0.742***
Friday	-0.474***	-0.378**	-0.656***	-0.591***	-0.550***
Saturday	0.118	0.035	0.015	-0.011	-0.041
holiday	0.099	0.169*	-0.050	0.363**	0.256*
year = 2020	-0.762***	-0.281***	-0.050	0.034	0.107
covid	-0.050	-0.154	0.089	0.337***	0.662***
spring	-0.525***	-0.465***	-0.113	0.065	-0.100
summer	-0.682***	-0.486***	-0.474***	-0.243**	-0.257*
fall	-1.710***	-0.984***	-0.734***	-0.669***	0.052
Constant	3.539***	3.346***	3.589***	4.141***	2.107***

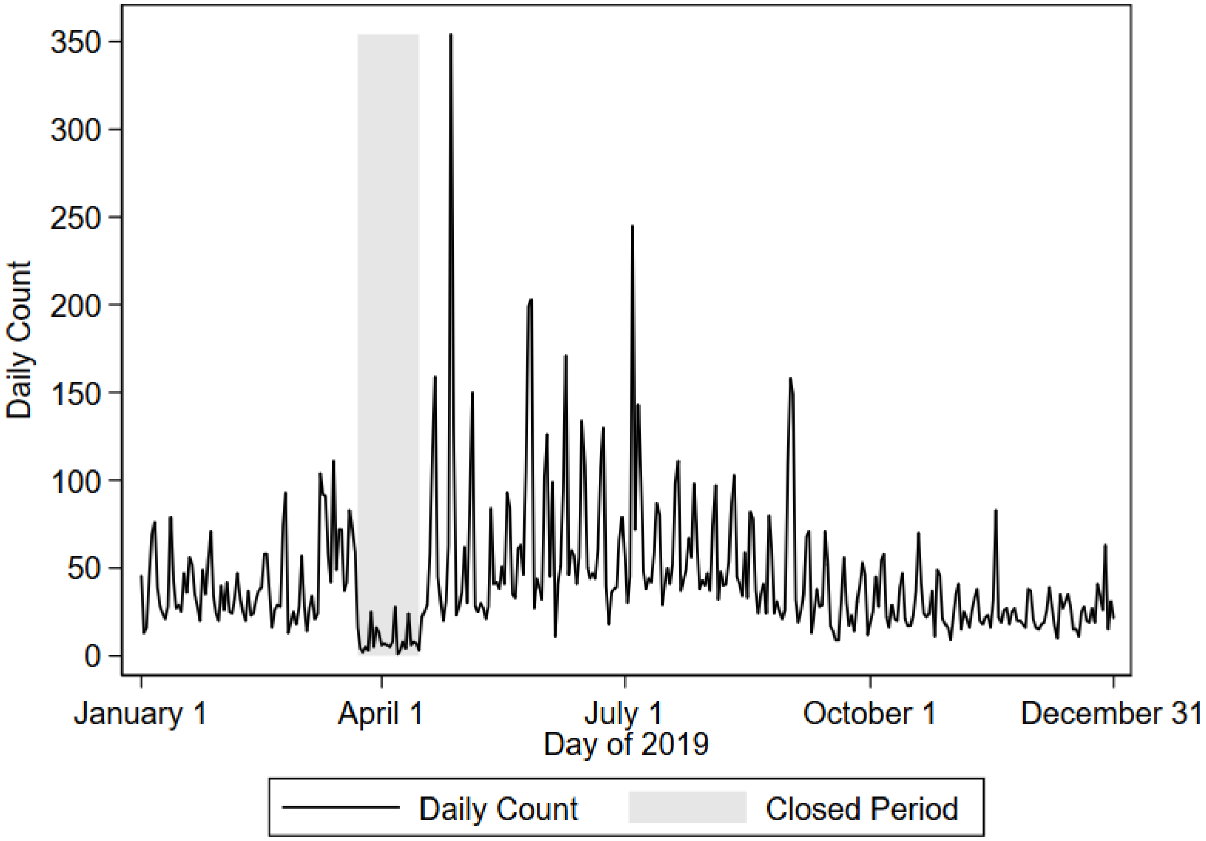
*Notes: N=609 for all specifications. * p<0.05 | ** p<0.01 | *** p<0.001. Temperature bins are relative to Tmax < 60 ℉, day-of-week dummies relative to Sunday, season dummies relative to winter. These specifications use the raw mobility data counts, not extrapolated to totals by the coverage rates.*

Table 4: Zonal travel cost model results using the 150-mile cutoff

	Battleship TX	San Jacinto	Bay Area	Sylvan Beach	BNC
travel cost	-0.022**	-0.025***	-0.118***	-0.070***	-0.085***
gender ratio	0.002	0	0.006	-0.006	-0.005
median age	-0.015	-0.009	0.121***	0.039	0.024
percent white	0.031***	0.035***	0.033***	0.030***	0.017***
Constant	-1.228**	-1.187*	-3.110**	1.025	0.873

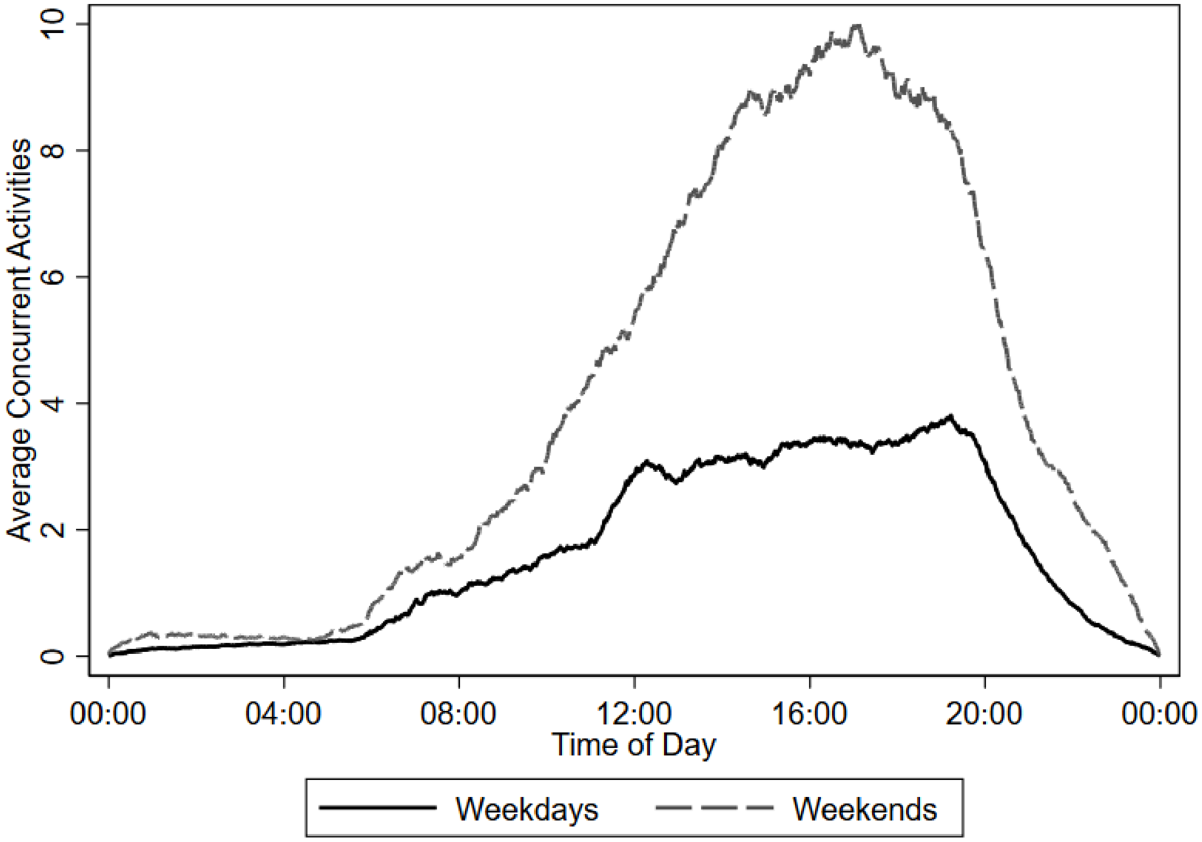
*Notes: N = 370 for all specifications. * p<0.05 | ** p<0.01 | *** p<0.001. Results are shown for the set of ZCTAs that fall within 150 miles of every site in the study. Alternative specifications are shown in the Appendix – 150 miles is the most inclusive cutoff we consider, and thus corresponds to the higher end of the range of per-trip cost estimates.*

Figure 1: Daily mobility data counts at Sylvan Beach for each day in 2019



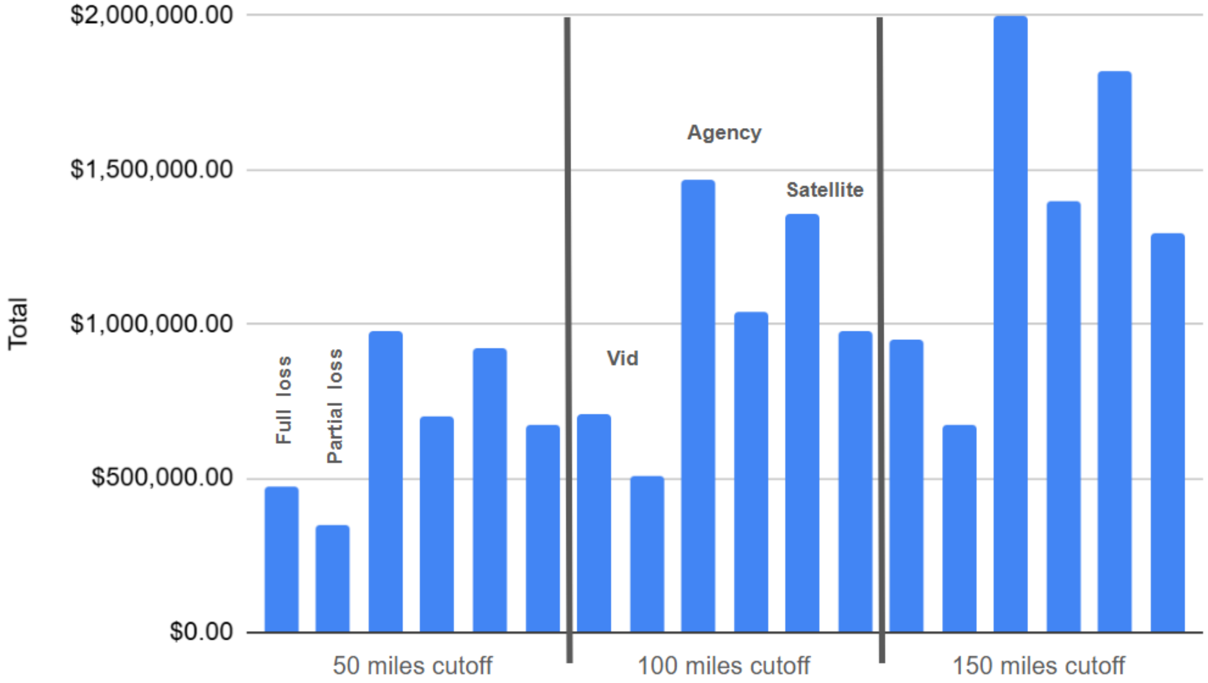
Notes: Line graph depicting variation in total count of unique daily visitors to Sylvan Beach for each day in 2019. The gray section indicates the site's closure period, (March 23 – April 15, 2019).

Figure 2: Unique devices present at Sylvan Beach by time of day



Notes: Average timing of visitation at Sylvan Beach derived from activity timestamps (start and end) in the Unacast dataset, averaged across all days in the dataset. Each activity was converted to a sequence of minute-by-minute indicators, and this sequence is averaged across days for each minute of the day.

Figure 3: Range of estimates for total lost surplus across five sites



Notes: Histogram showing 18 dollar-value total lost surplus estimates. Each estimate is grouped visually according to the three dimensions discussed in Section 4.4.

---

1 Data that are collected after the event is potentially affected by behavioral responses for some unknown period of time into the future. Data collected before the event is often unavailable, but it would be especially valuable for understanding baseline use of the site before the incident.

2 Intercontinental Terminals Company (ITC) 2nd 80s Tank Fire. https://darrp.noaa.gov/oil-spills/itc-tank-fire

3 See unacast.com

4 A ZCTA is essentially the projection of its corresponding zip code onto a polygon.

5 In 2020, NOAA purchased this particular Unacast dataset as a one-time proof of concept project, where the spatial and temporal extents were defined explicitly in the contract language. It is unclear to what extent the data product we obtained in 2020 shares characteristics with datasets that might be available after the time of publication.

6 Origin ZCTAs are calculated by the vendor in reference to the device's most common evening location as determined by a proprietary algorithm. In our dataset, each device is assigned a single ZCTA. If a device owner has multiple true origin ZCTAs (e.g., if they moved homes) during the time window of the dataset, this will introduce error into the variable.

7 For example, an activity might be classified as "moving" if its constituent pings are close in time, but far in space, implying that the device was moving through a site rather than staying for a while. We conservatively use only "stationary" type activities, those most likely to reflect true recreation.

8 Route Analysis is a premium function in ArcGIS Pro, available to the authors through an enterprise agreement between NOAA and ESRI. Route Analysis functions similarly to software like PC Miler™. We calculated the centroids of each ZCTA and each site, and Route Analysis uses them to determine the

---

nearest road, and calculates both driving time and mileage. Route Analysis allows users to make calculations relative to a particular time of day to account for typical traffic, so to be conservative we chose the “midnight” option, reflecting minimal traffic.

[9] In this case, another assumption is required, that the midpoint of the shift window is an appropriate interpretation of the videography timestamp (e.g., a shift from 8am to noon is reformatted as a snapshot count at 10am). This new assumption may be violated if, for instance, a site’s count is recorded systematically earlier or later than the midpoint so that the MD timestamp of reference is disconnected from the reference data.

[10] See maxar.com and earthexplorer.usgs.gov

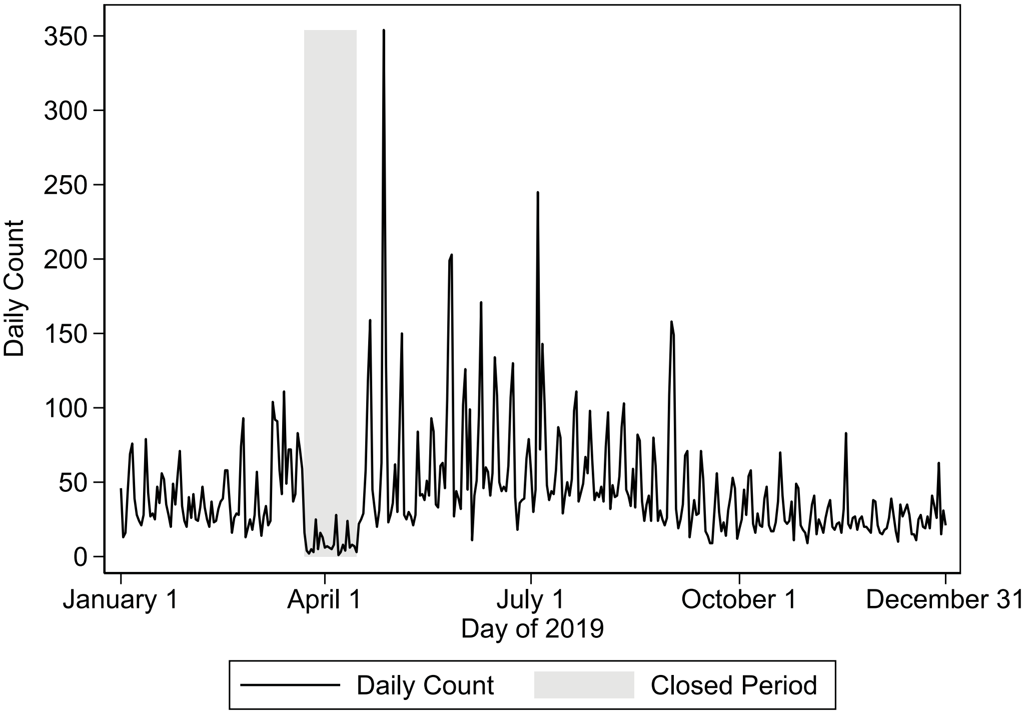
[11] This is a common placeholder assumption used by NRDA analysts and resource managers in lieu of a primary study. See Domanski et al., (2017).

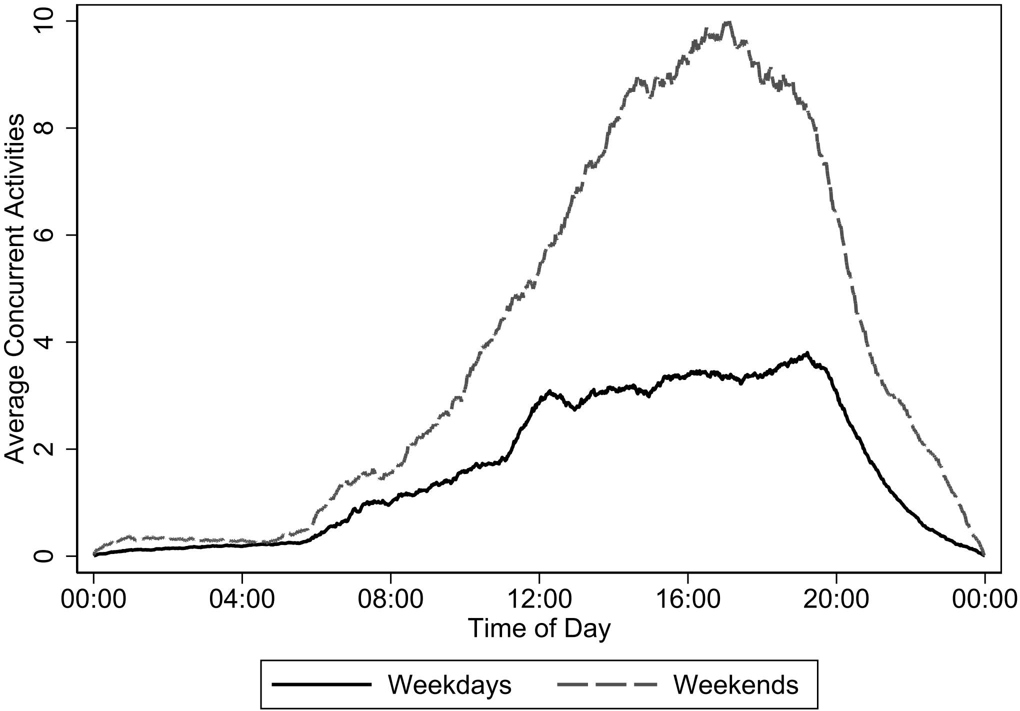
[12] Consider the hypothetical example of a recreator who arrives at a site without using location services until after the satellite image is captured, and then proceeds to record their first MD ping. This means they are captured in the reference data but not in the overlapping mobility data count, as their true activity has begun before their MD-recorded activity.

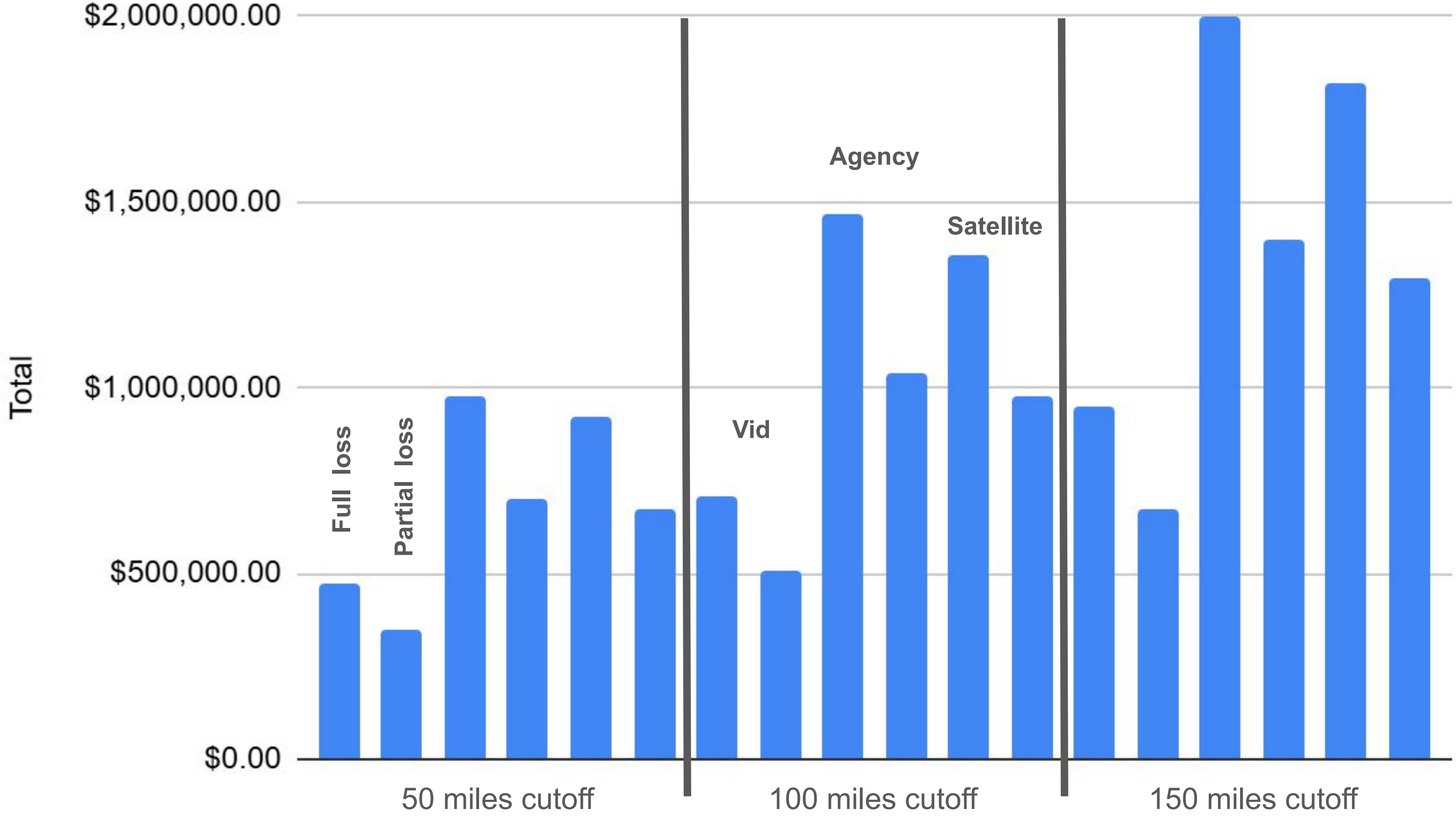
[13] For example, a false positive might be recorded if a potential recreator arrives at the entrance to a closed site and reads signage explaining the closure from the outside. In summary, we strongly expect that a site’s closure will influence the distribution of the types of people represented in the MD.

[14] Using data from “Your Driving Costs” (AAA, 2019) and the methodology described in Leggett et al. (2018), we model the direct mileage cost of driving as the sum of: 1) the per mile costs of fuel, maintenance, repair, and tires (20.54 cents for the average vehicle); and 2) the per-mile increment of vehicle depreciation ($354 increased depreciation for an additional 5000 miles = 7.08 cents per mile). The total direct cost of miles driven is thus 27.62 cents per mile.

[15] The most distant pair of closed sites is about 18 miles apart.


350
300
250
200
150
100
50
0
Daily Count
January 1
April 1
July 1
October 1
December 31
Day of 2019
Daily Count
Closed Period


Average Concurrent Activities
0
2
4
6
8
10
00:00
04:00
08:00
12:00
16:00
20:00
00:00
Time of Day
Weekdays
Weekends


Total
$2,000,000.00
$1,500,000.00
$1,000,000.00
$500,000.00
$0.00
Full loss
Partial loss
Vid
Agency
Satellite
50 miles cutoff
100 miles cutoff
150 miles cutoff