MEMORANDUM | September 1, 2015

TO Craig O'Connor, NOAA

FROM Roger Tourangeau and Eric English

SUBJECT B1– Estimation Procedures for Count Data

This memo describes the estimation procedures used to generate estimates of lost beach ("shoreline") visits, angler visits, and boating trips. For each activity, the memo provides an overview of the sampling plan, and then describes the basic estimator, adjustments for nonresponse and differences in weather, and variance estimation. Each section concludes with an exhibit presenting the loss estimates. Additional details about the sampling plan are provided in Technical Memo BA-1a – Shoreline Use Sample Design, Technical Memo BA-2a – Shore Fishing Sample Design, and Technical Memo BA-3a – Boating Sample Design.

SHORELINE VISITS

SAMPLE SELECTION

The data from which the estimated numbers of beach visits are derived come from three sources — overflight photographs, onsite counts, and onsite interviews. During any given ten-week period (beginning in June of 2010), overflights were scheduled for 20 weekdays (from Monday through Thursday) out of the 40 weekdays in each period and 20 weekend days (Friday through Sunday) out of the 30 possible weekend days. For each overflight, two coders counted the number of people on the beach in the aerial photographs; these photographs cover one out of every five beach "segments." We use the average of these two counts as the basic count data. (When there is a discrepancy between the counts that exceeds a fixed threshold, a third count is made and we used the average of the two closest counts.)

The beach segments were established at the outset of the survey and were designed to include roughly equal numbers of beachgoers. That is, the beach segments were narrower strips of beach in busy areas and wider strips in relatively isolated areas. Systematic samples of one in five segments were counted for each overflight, with each beach segment counted eight times (four times on weekdays, four times on weekends) during any given ten-week period.

The overflights covered seven of the nine target areas for the survey. In the remaining two areas (Grand Isle, Louisiana, and the Florida Keys), all of the counts were done on site. All of the beach segments in Grand Isle and the Keys were counted on a sample of weekend and weekday days and, in addition, onsite interviews were conducted when the counts were collected. The onsite interviews collected information about the visit, including its total expected duration (when the person arrived until when he or she expected to leave). In the overflight areas, onsite counts and interviews were conducted on a subsample of the days on which overflights occurred. Within each ten-week period, onsite count and interview data were collected on one-fourth of the weekend and weekday overflight days.

The sample, thus, was stratified by area (the seven areas with overflights, Grand Isle, and the Florida Keys), type of day (weekend or weekday), and ten-week period. In addition, the ten-week periods were subdivided into two-week periods, a feature that is used in the variance estimation procedure. Every beach segment was represented in each ten-week period. The overflights were randomly assigned to be done in the morning or afternoon and to start at the northernmost /easternmost segment of the area covered by the overflight or at the opposite end of the area (that is, the southernmost/westernmost beach segment). Similarly, the onsite counts were randomly assigned to a morning or afternoon time.

The total period of the day covered by the interviews and overflights varied somewhat by the time of the year and other factors; it was generally either 7.5 hours or 9.5 hours on a given sample day.

BASIC ESTIMATION PROCEDURE

Let c_{pwijk} be the count for a particular beach segment (segment *k*) within a given geographic area (area *j*—for example, the state of Alabama), on a particular day (day *i*) within the two types of day (weekend or weekday, *w*) during a ten-week time period (*p*).[1] Let $\bar{d}_{pwj}$ be the (weighted) mean duration of the visits for the same period-area-day type combination; this mean was the harmonic mean of the durations rather than the arithmetic mean (see Equation 2 below).

We used the counts, durations, and an estimate of the proportion of beach visits that were recreational ($p_{rec,pwj}$) to estimate the total number of recreational beach visits ($\hat{v}_{pwijk}$) for a beach segment on a specific day:

$$\hat{v}_{pwijk} = \frac{c_{pwijk}}{\bar{d}_{pwj} / s_{pwijk}} p_{rec,pwj} \quad . \qquad (1)$$

The variable s_{pwijk} is the length of the sample day for that segment on that day. Two components of this estimator are based on the onsite interview data — $\bar{d}_{pwj}$ and $p_{rec,pwj}$. These two components are computed using similar procedures. $\bar{d}_{pwj}$ is the weighted harmonic mean of the durations for an area-day type-period combination:

$$\bar{d}_{pwj} = \frac{1}{\dfrac{\sum_{i=1}^{n_i} \sum_{k=1}^{n_j} \sum_{l=1}^{n_{pwijk}} w_{wijk} \, {}^{1}\!/\!{}_{d_{pwijkl}}}{\sum_{i=1}^{n_i} \sum_{k=1}^{n_j} \sum_{l=1}^{n_{pwijk}} w_{pwijk}}} \quad . \qquad (2)$$

[1] Sampling periods were occasionally less than 10 weeks to accommodate a consistent starting point for the sampling plans in each year. In analysis, any periods shorter than 10 weeks were combined with the preceding 10-week period.

The weight (w_{pwijk}) is the inverse of the product of three probabilities—the selection probability for the beach segment, for the day (within a day type and period), and for the individual, person l (for example, one person in three might have been selected for interviews during the two-hour period of data collection).[2] We aggregated to the area-day type-period combination because there were many individual beach segment-day pairs where they were very few or no interviews and hence no durations. Aggregation to the area-day type-period triplet eliminated this problem.

The proportion recreational ($p_{rec,pwj}$) is also a ratio estimate at the level of area-day type-period triplets:

$$p_{rec,pwj} = \frac{\sum_{i=1}^{n_i}\sum_{k=1}^{n_j}\sum_{l=1}^{n_{pwijk}} w_{pwijk} r_{pwijkl}}{\sum_{i=1}^{n_i}\sum_{k=1}^{n_j}\sum_{l=1}^{n_{pwijk}} w_{pwijk}} , \quad (3)$$

where r_{pwijkl} is a flag that takes on a value of 1 when beachgoer l reported visiting the beach for recreational purposes and a value of 0 if he or she reported visiting the beach for some other purpose.

Once we had these basic beach segment-day estimates, we aggregated over them to get estimated totals. For example, to get the total for an entire year for area j, we summed the estimates for all the observations for that area and that year, applying a weight (w_{pwk}) that is the inverse of the selection probability for the beach segment during that time period (reflecting both the segment's selection probability and the selection of days of a given type for that segment).[3]

MISSING COUNTS

There were two forms of missing data that could affect the estimates — missing counts and missing interview data. Regarding the first, sometimes a scheduled overflight did not take place or the ground counters did not make it to their assigned area. In some cases, the overflight was cancelled because of bad weather. Thus, for each beach segment-day pair where a count was scheduled, a disposition code was assigned: 1) the count took place as scheduled; 2) the count did not take place because of bad weather and we infer the count would have been zero; and 3) the count did not take place but a zero count could not be inferred.[4] If both ground and aerial counts were available we used the aerial count. In a few cases, count data were available from ground counters when the overflight data were missing, and we used the ground count. So, we treated the count

[2] Data collection began August 9, 2010 in the Florida Keys. Interview weights for August 2010 in the Florida Keys included a component (*early* weight) that weighted up available interviews to represent the full month. The interview weights also included a minor adjustment that applied when the length of the sample day changed during a sampling period. The adjustment placed a higher weight on interviews conducted during the longer sample days, in proportion the length of the sample day.

[3] The weighting component ***early weight***, described in the previous footnote, was also applied in this calculation.

[4] Sometimes the overflight was cancelled mid-flight so that overflight data were available for some beach segments on a given day but not others. In such cases, we used the aerial counts for the segments where they were available.

data for a given beach segment-day pair as complete if either overflight data were available, ground counts were available, or a zero could be inferred from the bad weather.

In about 3 percent of the scheduled beach segment-day combinations, count data were missing and a zero count could not be inferred. We used a conventional nonresponse adjustment to compensate for these missing counts. That is, for each area-day type-period triplet, we adjusted the weights based on the proportion of beach segments for which counts were scheduled and count data were available (including inferred zeroes):

$$w'_{pwj} = w_{pwj} \times \frac{n_{spwj}}{n_{cpwj}} \quad , \quad (4)$$

in which n_{spwj} is the number of beach segments where counts were scheduled to be taken in an area-day type-period triplet and n_{cpwj} is the number of beach segments where counts (including inferred zeroes) were actually available. Both n_{spwj} and n_{cpwj} are unweighted (since the sampling probabilities were uniform within a triplet).[5]

MISSING INTERVIEWS

The interview data could also be missing due to person-level nonresponse or item nonresponse. We attempted to create a nonresponse adjustment within each beach segment-day pair but ran into difficulties (e.g., sometimes all the cases at a specific beach segment were nonrespondents). We ultimately decided not to try to adjust for missing interview data at the segment-day level. Any adjustment at the triplet level, analogous to the nonresponse adjustment for counts, would have no effect since the recreation flag and durations were averaged within each triplet. Thus, the final estimator for a given area-day type-period triplet ($\hat{v}_{pwj}$) is:

$$\hat{v}_{pwj} = \sum_{k=1}^{n_{pwj}} w'_{pwj} \hat{v}_{pwijk} \; , \quad (5)$$

where $\hat{v}_{pwj}$ is the estimate for day type w for area j during period p and n_{pwj} is the number of day-beach segment pairs with counts within the area-day type-period triplet. The estimate for a sub-period (such as a month) is the weighted sum (based on w'_{pwj}) for the observations in that sub-period.

RAKING AND WEATHER ADJUSTMENT

The major purpose of the study was to estimate the difference between the number of visits to the beach during the aftermath of the spill (the "spill" period) and the numbers during a baseline period when the Gulf beaches seemed to be back to normal. We have shoreline data beginning in June of 2010 and continuing through May of 2013. As we discuss in greater detail in Technical Memo I3 – Baseline Estimation and the Use of Post-Spill Data, we defined the "baseline" year as the first year for which the data indicated there were no further effects of the spill on beach recreation. In this context "year" is defined with respect to the start of the study, with each year of the study starting June 1. For the North Gulf (that is, Louisiana, Mississippi, Alabama, and the Florida panhandle),

[5] The weighting component, ***early weight***, described in a previous footnote was also applied in this calculation.

the baseline year encompassed the final twelve months (June, 2012, through May, 2013) for which data are available. For the Peninsula (the remaining portions of the Gulf Coast of Florida), the spill period consisted of the first eight months of data collection following the spill (June, 2010, through January, 2011) and the baseline period consisted of the same months one year later (June 2011, through January, 2012). Thus, without any adjustments for differences in weather across years, the estimate of lost visits to the beach is the difference in the $\hat{v}_{mj}$'s for corresponding months in say, 2012 and 2010:

$$\hat{\Delta}_{mj} = \hat{v}_{base,mj} - \hat{v}_{spill,mj} \quad . \quad (6)$$

For example, the estimated number of trips lost in October of 2010 in area j in the North Gulf is the difference between the estimate for that area from October, 2012, ($\hat{v}_{base,mj}$) and the estimate for the same area from October, 2010 ($\hat{v}_{spill,mj}$).

To compensate for variation in weather over time, we created a weather adjustment factor. At the same time, the adjustment ensured that the sample data aligned with the full population data in other important ways. Specifically, we classified each day-beach segment pair (not just the sample pairs but the entire population) by region (North Gulf or Peninsula), three-month period, type of day (weekend or weekday), the amount of rain (up to three categories), the maximum temperature (up to four categories), and the time of day (up to three categories, representing the morning, early afternoon, and late afternoon portions of the day). Then, we "raked" the sample weights to agree with the population figures. First, we raked the spill period weights to align them with the spill period population figures; then we raked the baseline sample weights so they aligned with the corresponding spill period population figures. That is, the adjusted weights estimate the number of beach visits during the baseline period, assuming the same distribution of days by rain and temperature category as in the spill period.

EXHIBIT 1. RAKING SAMPLE PERIODS AND POPULATION TARGETS

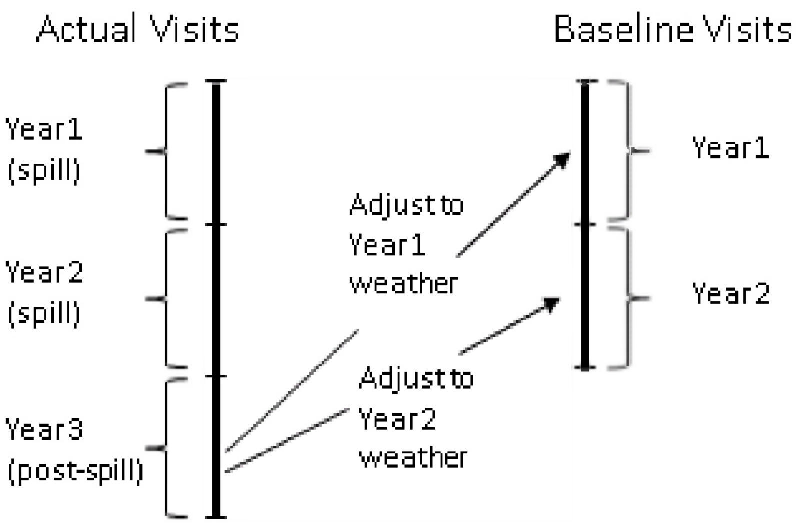
REGION/PERIOD	DATES	RAKING TARGETS (POPULATION FIGURES)
North Gulf		
Spill Period,Year 1	June, 2010—May, 2011	June, 2010—May, 2011
Spill Period,Year 2	June, 2011—May, 2012	June, 2011—May, 2012
Baseline Period	June, 2012—May, 2013	June, 2010—May, 2011 June, 2011—May, 2012
Peninsula		
Spill period	June, 2010—January, 2011	June, 2010—January, 2011
Baseline	June, 2011—January, 2012	June, 2010—January, 2011

Exhibit 1 provides more detail about how we raked the sample weights to account for differences in weather when estimating baseline use for both the North Gulf and the Peninsula. Exhibit 2 displays the process for the North Gulf graphically. In effect, we created two baseline estimates for the North Gulf, one with the same distribution of weather as the first spill year and the second with the same distribution of weather as the second spill year.

In both regions, we raked the weights to the population targets for three dimensions. The first dimension was the cross-classification of each site by region, type of day, three-month period, and rain category. The second dimension was the cross-classification of each site by region, type of day, three-month period, and temperature category. The third dimension was the cross-classification of each site by region, type of day, three-month period, and time category of the observation. This last variable adjusted the weights to impose a uniform distribution of observations across the three portions of the day.

Raking was carried out for both the combined aerial-ground counts data and the onsite interview data. The adjusted weights from the former raking were applied to the estimation of $\hat{v}_{pwj}$; the adjusted weights from the latter raking were applied to the estimation of $\bar{d}_{pwj}$ and $p_{rec,pwj}$. Raking cell sizes were constrained to a minimum of 20 cases and raking adjustment factors were constrained to a range of 0.1 to 10.[6]

EXHIBIT 2 RAKING AND WEATHER ADJUSTMENT FOR NORTH GULF



Variance estimates are computed via jackknifing (using the method that Rust and Rao, 1996, described on page 289; see also Chapter 4 in Wolter, 1985). For variance computation purposes, we grouped all the observations into strata by sampling plan region,[7] type of day, and two-week period. The strata were split into sub-strata (or variance computation strata) containing two days of data or, in a few cases, three days of

[6] These constraints were maintained by collapsing two or more cells together, ensuring either the minimum cell size or the maximum raking adjustment factor. Collapsing was done such that the numbers of cases in adjacent cells were combined.

[7] There were four sampling plan regions for the shoreline study: Grande Isle, the rest of the North Gulf (Mississippi, Alabama, and the Florida panhandle), the Florida Peninsula, and the Florida Keys.

data. From each of the resulting 911sub-strata, the observations for a randomly chosen day were dropped and the weights for remaining observations were adjusted upward to compensate. The sub-strata were randomly grouped into 200 jackknife replicates for the purpose of computing variances.

LOSS ESTIMATES FOR SHORELINE VISITS

Exhibit 3 below shows the estimates of lost shoreline visits (and their standard errors) with and without the weather adjustment described above. Given the importance of controlling for weather across years, we rely on the adjusted loss estimates and conclude that there were about 12.3 million fewer shoreline visits to the Gulf than there would have been without the spill.

EXHIBIT 3. ESTIMATES OF LOST SHORELINE VISITS (AND STANDARD ERRORS), BY REGION

REGION	UNADJUSTED LOSS ESTIMATE	ADJUSTED LOSS ESTIMATE
North Gulf (18 months)	**7,980,855** (1,590,285)	**8,455,336** (1,622,462)
Peninsula (8 months)	**3,859,294** (1,230,666)	**3,870,176** (1,014,982)
Overall	**11,840,149** (2,055,065)	**12,325,512** (1,893,531)

ANGLER VISITS

SAMPLE SELECTION

The sample of angler visits was selected in three stages. At the first stage, we selected a sample of angling sites from the Marine Recreation Information Program (MRIP) database. There were 406 MRIP angling sites in the target area. Sample sites were selected using systematic probability-proportional-to-size (PPS) sampling, where the size measure was the MRIP estimate of the annual site visits by anglers. There were relatively few sites in Alabama and Mississippi, and we selected those sites with certainty. The selection of the remaining sites was done systematically, with sites sorted by state, county, and size. Subsequently, some of the angling sites were dropped from the sample because they were already covered in the shoreline sample. In total, the final first-stage sample included 49 sites in the North Gulf and 68 sites in the Peninsula, or a total of 117 sites. The final frame consisted of 323 sites.

At the second stage of sampling, nearby angling sites samples were grouped together and a systematic sample of weekdays and weekend days (again defined as Fridays, Saturdays, or Sundays) were selected for each group. During each twelve-week sampling period, every group of sample sites was slated for data collection on three weekdays and three weekend days. (At the end of the sampling year, a four-week sampling period was used, with one weekday and one weekend sampled for each group.) The grouping of sites was done so that a team of interviewers could cover the entire group in a single day. On the sample days, a randomly selected site within the group was the start point for data collection, with the remaining sites visited in a fixed order (e.g., on one sample day the order might have been BCDEFGA and on the next, FGABCDE). In addition, the group was assigned a randomly selected start time (between 7 a.m. and 7 p.m. in summer and between 6:30 a.m. and 5:30 p.m. in winter).

In the final stage of sampling, individual anglers were selected for interviewing; the anglers were selected systematically with a random start; the sampling rate was determined by the total number of anglers at the site.

ESTIMATION PROCEDURE

The estimation procedure for the angler visits closely paralleled that for the shoreline visits. In the first step, we used the counts, durations, and an estimate of the proportion of visits to the angling site that were recreational to estimate the total number of recreational angling visits at that site on that day (as in Equation 1). Both the harmonic mean of visit duration and the proportion of visits that were recreational were calculated for the area-day type-sampling period triplet and then applied to all sample sites and days within that triplet (cf. Equations 2 and 3). Second, the weights were adjusted for site-level nonresponse (cf. Equation 4).[8] Then, we calculated an unadjusted estimate of the lost angler visits, using the analogue to Equation 6. Finally, we calculated a weather adjustment (and adjustment to the frame) using the same raking dimensions as with the shoreline visits. In the North Gulf, the period from June, 2010, through March, 2011, constituted the spill period and the period from June, 2011, through March, 2012, constituted the baseline period. Therefore, North Gulf sample observations from the spill period in the first year and the baseline period one year later were both raked to the population totals of the first analysis year. Under the methods outlined in Technical Memo I3 – Baseline Estimation and the Use of Post-Spill Data, there did not appear to be a spill effect on recreational fishing in the Peninsula.

VARIANCE ESTIMATION

As with the beach visits estimates, we computed variance estimates for the angler visits via jackknifing. For variance computation purposes, we grouped observations by region[9], type of day, and four-week subperiod. These strata were split into sub-strata (or variance computation strata) containing two days of data or in a few cases three days of data. From each of the resulting 806 sub-strata, the observations for a randomly chosen day were dropped and the weights for remaining observations were adjusted upward to compensate. The sub-strata were randomly grouped into 250 jackknife replicates in computing the variance estimates.

LOSS ESTIMATES FOR ANGLER VISITS

The unadjusted estimate of lost angler visits (which were confined to the North Gulf) was 207,502, with a standard error of 61,681; the adjusted estimate was 144,050, with a standard error of 46,348. The combined shoreline and angler losses are shown in Exhibit 4.

[8] Data collection in the North Gulf began June 4, 2010, and in the Florida Peninsula data collection began July 2, 2010. Interview weights for June 2010 in the North Gulf and July 2010 in the Florida Peninsula included a component that weighted up available interviews to represent the full month.

[9] There were two sampling plan regions for the fishing study: the North Gulf (Louisiana, Mississippi, Alabama, and the Florida panhandle) and the South Gulf (Florida Peninsula and the Florida Keys).

EXHIBIT 4. ADJUSTED ESTIMATES OF COMBINED SHORELINE AND ANGLER VISITS AND LOSSES (AND STANDARD ERRORS) AND PERCENT DECLINE, BY REGION AND PERIOD

PERIOD /REGION	ADJUSTED LOSS ESTIMATE	ADJUSTED SPILL ESTIMATE	ADJUSTED BASELINE ESTIMATE	PERCENT DECLINE
Months 1-8				
North Gulf	**6,425,237** (944,623)	**7,782,270** (565,853)	**14,207,507** (737,483)	**45.2%**
Peninsula	**3,870,176** (1,014,982)	**13,601,695** (701,037)	**17,471,871** (701,090)	**22.2%**
Overall	**10,295,412** (1,388,797)	**21,383,964** (910,558)	**31,679,377** (1,029,779)	**32.5%**
Months 9-18				
North Gulf	**2,174,149** (1,068,929)	**19,580,582** (639,215)	**21,754,732** (873,894)	**10.0%**
Total (Months 1-18)	**12,469,562** (1,894,098)	**40,964,547** (1,109,725)	**53,434,109** (1,582,834)	**23.3%**

BOATING TRIPS

SAMPLE SELECTION

The boating sample was also selected in three stages. In the first stage, boating sites were selected using PPS sampling. Prior to sampling, the list of sites had been sorted by state, county, and size. The frame for the boating sites was the MRIP list of public and private boat launches and marinas in the target area and the measure of size was the MRIP estimate of the number of boats launched from each site over the course of a year. One hundred ninety three sites were selected (out of 534 sites in the frame), 32 of them with certainty. At the second stage of sampling, samples of weekdays and weekend days (again defined as Fridays, Saturdays, or Sundays) were selected for each site. During each twelve-week sampling period, each sample site was slated for data collection on three weekdays and three weekend days. For each sample site, a sample of days of each type was selected systematically (with a random start). Some of the certainty selections were very large (approximately twice the selection interval for first-stage sampling) and were scheduled for data collection on six weekdays and six weekend days during each sampling period rather than three. (At the end of the sampling year, a four-week sampling period was used. One day of each type was selected for each site during this abbreviated period; for the large certainty sites, two days of each type were selected.)

At the third stage of selection, interviews and onsite counts were conducted during one of three shifts. The field counters were randomly assigned to one of three start times and were supposed to cover a site for an assigned period, or shift, of between 2.5 and 5 hours. During that time, they counted the number of boating parties and asked a member of each party whether the purpose of the outing was recreational. Counters were assigned to one of the three shifts with a probability of one in three.

BASIC ESTIMATION PROCEDURE

We produced two sets of estimates from these data—an estimate of the total number of boating parties for a given area or period and an estimate of the total number of boaters.

The estimators for parties and boaters were very similar. The estimate for the number of boating parties was based on the counts and the interview data regarding whether each party was recreational or not. Let b_{pwijk} be the count for a particular boating site (site *k*) within geographic area *j* on sample day *i* of specific day type *w* (weekend or weekday) during sampling period *p*. The estimated number of recreational boating parties for the site-day pair was:

$$\hat{p}_{pwijk} = b_{pwijk} \times p_{rec,pwj}\,, \quad (8)$$

where $p_{rec,pwj}$ is the proportion of parties that are recreational. As with the shoreline and angling data, $p_{rec,pwj}$ is an aggregated estimate, calculated for the area-day type-period triplet, since it was often impossible to estimate for a specific day-site pair (owing to missing data):

$$p_{rec,pwj} = \frac{\sum_{i=1}^{n_i}\sum_{k=1}^{n_j}\sum_{l=1}^{n_{pwijk}} w_{pwijk} r_{pwijkl}}{\sum_{i=1}^{n_i}\sum_{k=1}^{n_j}\sum_{l=1}^{n_{pwijk}} w_{pwijk}}\,. \quad (9)$$

Once again, r_{pwijkl} is a flag that takes on a value of 1 if the purpose of the outing was recreational and 0 otherwise, and *l* is the index for a given party. The flag was also set to 0 if the trip was on a charter boat. The weight, w_{pwijk}, is the inverse of the product of the site selection probability, the day selection probability, and shift selection probability (which was a constant—1/3).[10]

The estimated number of boaters for a site-day pair is based on the average size of the boating parties:

$$\hat{m}_{pwijk} = b_{pwijk} \times p_{rec,pwj} \times \bar{s}_{pwj} \quad (10)$$

where $\bar{s}_{pwj}$ is the weighted average of the sizes of the recreation parties for the relevant area-day type-period triplet.

Unadjusted aggregate estimates for a given time period or a given area are weighted sums of the $\hat{p}$'s or $\hat{m}$'s for that period or area. For example, the estimates for year *y* are:

[10] Data collection for boating began June 2, 2010 in Alabama and Mississippi. Data collection in the Florida Peninsula began June 11, 2010. Interview weights for June 2010 in all these areas included a component that weighted up available interviews to represent the full month. On rare occasions, the sampling plan selected a site for data collection twice at the same time on the same day.

$$\hat{p}_y = \sum_1^{n_y} w_{pwijk}\hat{p}_{pwijk}$$
$$\hat{m}_y = \sum_1^{n_y} w_{pwijk}\hat{m}_{pwijk} \quad . \qquad (11)$$

in which n_y is the number of day-site pairs for which data were collected.

COMPENSATING FOR MISSING DATA

As with the shoreline and angling studies, scheduled boating observations sometimes did not take place or were not completed. In some cases, the shift was cancelled due to bad weather and it was reasonable to infer counts of zero. In other cases, the shift was cancelled due to other reasons (e.g., a counter was ill). These latter cases were treated as nonrespondents, and we adjusted the weights to compensate for these missing observations. We used the same method described in Equation 4 to adjust for missing observations, increasing the weights for the non-missing observations within each area-day type-period triplets by the inverse of the weighted proportion of site-day pairs for which data were obtained or zero counts could be inferred. Again, we did not create an adjustment for missing interviews within a site-day pair.

DIFFERENCES AND WEATHER-ADJUSTED DIFFERENCES

Again, the key estimates from the boating study are the changes in the numbers of boating parties and the number of boaters from the spill period to the baseline period:

$$\hat{\Delta}^p_{mj} = \hat{p}_{base,mj} - \hat{p}_{spill,mj}$$
$$\hat{\Delta}^m_{mj} = \hat{m}_{base,mj} - \hat{m}_{spill,mj} \quad , \qquad (12)$$

in which $\hat{\Delta}^p_{mj}$ represents the reduction during month m in the number of boater parties from the base period to the spill period and $\hat{\Delta}^m_{mj}$ represents the reduction in the number of boaters in that month.

As with the shoreline and angling estimates, we created an adjustment factor to compensate for variation in weather across years and to bring the weights into line with frame totals. This was done using the same raking dimensions as with the shoreline and angling estimates, except in boating the three daily sampling shifts were used in place of the time-of-day categories. In the North Gulf, the period from June, 2010 through August, 2010 constituted the spill period and the period from June, 2011 through August, 2011 constituted the baseline period. North Gulf sample observations from these first-year and second-year periods were both raked to the population totals of the first year. Under the methods outlined in Technical Memo I3 – Baseline Estimation and the Use of Post-Spill Data, there did not appear to be a spill effect on recreational boating in the Peninsula. The weather-adjusted base year estimate for a given month and area (e.g., $\hat{m}'_{base,mj}$) incorporated the weather adjustment factor; that is, the base period estimates were adjusted to the weather of the corresponding spill period.

In addition, the spill period estimates were adjusted to the population totals for the period. Thus, the weather adjusted estimate for the difference in the number of boating parties ($\hat{\Delta}'^{p}_{mj}$) was:

$$\hat{\Delta}'^{p}_{mj} = \hat{p}'_{base,mj} - \hat{p}'_{spill,mj} \quad ,$$

in which $\hat{p}'_{base,mj}$ is the weather-adjusted estimate for boating parties in a given month in the base period.

Weather-adjusted estimates for the difference in the total number of boaters (rather than parties) were computed using the same adjustment factors.

VARIANCE ESTIMATION

Variance estimates were again computed via jackknifing. For variance computation purposes, we grouped observations by region,[11] type of day, and four-week subperiod. The strata were split into sub-strata containing with two or, in a few cases, three days of data. From each of the resulting 1,171 sub-strata, the observations for a randomly chosen day were dropped and the weights for remaining observations were adjusted upward to compensate. The sub-strata were randomly grouped into 200 jackknife replicates in computing the variance estimates.

LOSS ESTIMATES FOR BOATING

The estimates of lost boating trips and boating parties (which were confined to a three-month period in the North Gulf) are shown in Exhibit 5. Overall, we estimate that there were 215,374 fewer boaters during the spill period than there would have been had the spill not occurred. This corresponds to 58,622 "parties" or boating trips.

EXHIBIT 5. ADJUSTED ESTIMATES OF BOATING VISITS AND LOSSES (AND STANDARD ERRORS) AND PERCENT DECLINE

	ADJUSTED LOSS ESTIMATE	ADJUSTED SPILL ESTIMATE	ADJUSTED BASELINE ESTIMATE	PERCENT DECLINE
Boaters	**215,374** (72,944)	**544,231** (49,880)	**759,605** (53,556)	**28.4%**
Parties	**58,622** (21,080)	**193,766** (15,122)	**252,387** (14,716)	**23.2%**

[11] There were three sampling plan regions for the boating study: Louisiana and the Florida panhandle, Mississippi and Alabama, and the South Gulf (Florida Peninsula and the Florida Keys).

REFERENCES

Rust, K.F., & Rao, J.N.K. (1996). Variance estimation for complex surveys using replication methods. *Statistical Methods in Medical Research*, *5*, 283-301.

Wolter, K. M. (1985). *Introduction to Variance Estimation*. New York: Springer-Verlag.