SUPPORTING STATEMENT – PART B

U.S. Army Corps of Engineers

Pacific Northwest Households Recreation Use Survey - OMB 0720-XXXX

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

If the collection of information employs statistical methods, it should be indicated in Item 17 of OMB Form 83-I, and the following information should be provided in this Supporting Statement:

1. Description of the Activity

The target population for this collection is adult residents of Washington, Oregon, Idaho, and western Montana. A stratified random sample of households will be selected using address-based sampling (ABS) and the United States Postal Service's Computerized Delivery Sequence File (CDSF). A single adult will be randomly selected from each household.

Four primary strata will be defined as follows (Exhibit 1):

- **Stratum 1 (Lower Snake River):** Counties bordering the Lower Snake River from Lewiston, Idaho to Pasco, Washington.
- **Stratum 2 (Columbia River East of Cascades):** Counties bordering the Columbia River from Hood River, OR to the U.S./Canadian border.
- **Stratum 3 (West of Cascades):** Counties in Washington and Oregon west of the Cascades.
- **Stratum 4 (Other Counties East of Cascades):** All other counties within the sampling frame but outside of strata 1, 2, and 3.

Initial sample sizes within each stratum are presented in Exhibit 2. The sampling rate is highest in strata 1 and 2, areas near the Columbia and Lower Snake rivers east of the Cascades. These areas have a high potential for impacts associated with changes to CRSO. To reflect our expectations about potential for CRSO-related impacts, the sampling rate will be greatest for stratum 1, followed by stratum 2, stratum 4, and stratum 3 (sampling rates by stratum shown in Exhibit 2). Within each stratum, the sample will be allocated to counties in proportion to the square root of each county's population. This allocation ensures that the sample will be well distributed geographically within each stratum.

EXHIBIT 1. SAMPLING STRATA

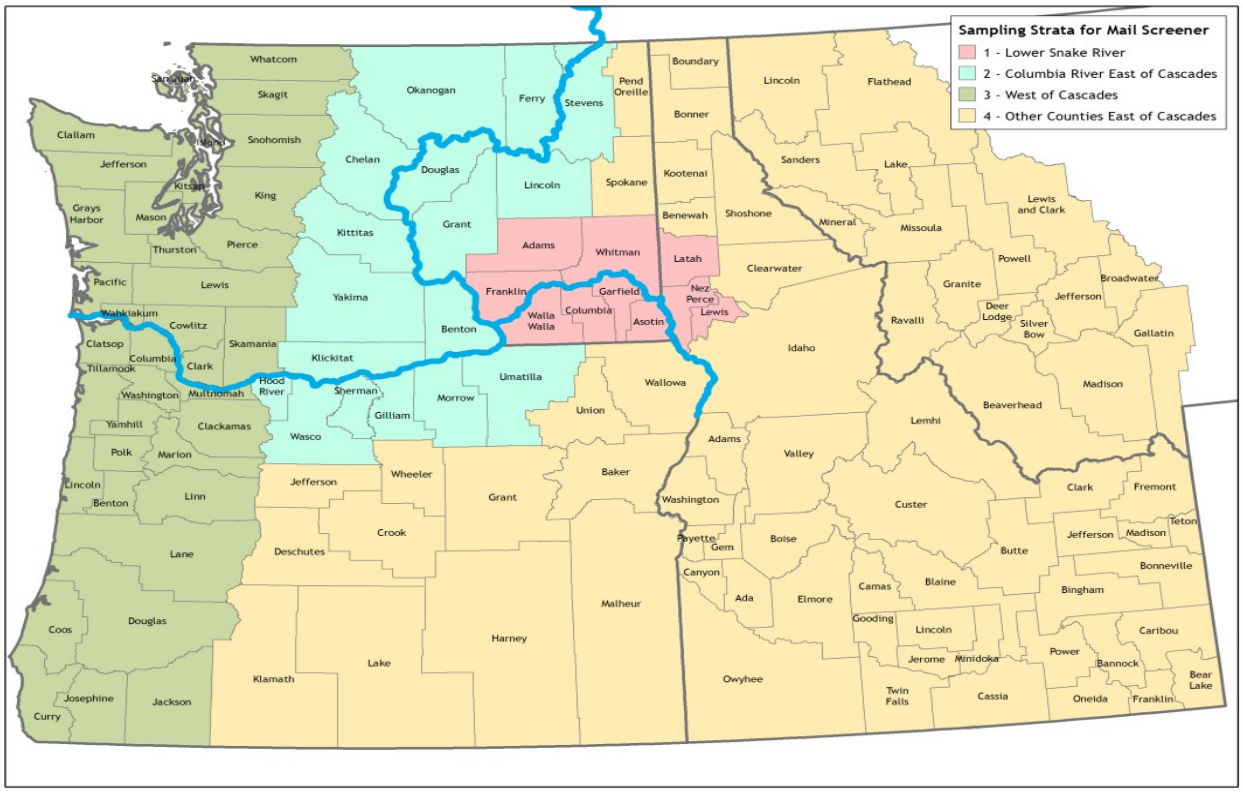

EXHIBIT 2. SAMPLING RATES

STRATUM	TOTAL HOUSEHOLDS	SAMPLE SIZE	PERCENTAGE OF SAMPLE	SAMPLING RATE (PER 1,000 HHS)
1. Lower Snake River	114,833	15,000	30%	131
2. Columbia River East of Cascades	337,005	15,000	30%	45
3. West of Cascades	3,474,377	10,000	20%	3
4. Other Counties East of Cascades	1,143,108	10,000	20%	9

With 50,000 sampled addresses, we anticipate approximately 9,200 survey responses:

50,000	Sampled addresses
x 0.92	Expected proportion valid addresses
x 0.20	Expected survey response rate
9,200	Survey responses

With the sampled addresses allocated equally across the three temporal waves, we anticipate obtaining approximately 3,067 survey responses per wave.

2. Procedures for the Collection of Information

a. Statistical methodologies for stratification and sample selection

The statistical methodology for stratification and sample selection is described in item 1 above. The target population for this collection is adults in Washington, Oregon, Idaho, and western Montana. A stratified random sample of households will be selected using address-based sampling (ABS) and the United States Postal Service's Computerized Delivery Sequence File (CDSF). Four primary strata will be defined (Exhibit 1) with sampling rates varying by strata (Exhibit 2). Within households the adult (age 18 or older) with the most recent birthday will be asked to complete the on-line survey.

b. Estimation procedures

The study will use data on individuals' recreation trips, recreation site attributes, and travel costs to estimate a RUM travel cost model. The model will be used to calculate changes in the economic value under various CRSO management alternatives. The specific type of RUM travel cost model originally described by Morey, Rowe, and Watson (1993). The utility associated with a visit to site *j* by individual *i* (i.e., the "site choice" decision) is given by

$$u_{ij} = \beta_{TC} C_{ij} + \gamma' X_j + \varepsilon_{ij}$$

$$¿ \; v_{ij} + \varepsilon_{ij}$$

where:

β_{TC} = the unknown parameter associated with the cost of traveling to a site

C_{ij} = the cost to individual *i* of traveling to site *j*

γ = a vector of unknown parameters associated with site attributes

X_j = a vector of attributes associated with site *j*

ε_{ij} = an error distributed as generalized extreme value

The utility associated with a decision not to visit a recreation site (i.e., the "participation" decision) is given by

$$u_{i0} = \alpha + \theta' Z_i + \varepsilon_{i0}$$

$$¿ \; v_{i0} + \varepsilon_{i0}$$

where:

α = a constant

θ = a vector of unknown parameters associated with demographic characteristics

Z_i = a vector of demographic characteristics associated with individual i

ε_{i0} = an error distributed as *i.i.d.* extreme value

Given these utilities and assuming the errors are jointly distributed as generalized extreme value, the probability that individual *i* will select site *j* on any given choice occasion is given by (Kling and Thomson 1996):

$$P_{ij} = \frac{\exp\left(\frac{v_{ij}}{\rho}\right)}{\exp\left(I_i\right)} x \frac{\exp\left(\rho I_i\right)}{\exp\left(\rho I_i\right) + \exp\left(v_{i0}\right)}$$

with the probability associated with choosing not to visit a recreation site is given by

$$P_{i0} = \frac{\exp\left(v_{i0}\right)}{\exp\left(\rho I_i\right) + \exp\left(v_{i0}\right)}$$

where I_i represents the "inclusive value" for individual *i* and is defined as:

$$I_i = \ln \sum_{j=1}^{J} \exp\left(\frac{v_{ij}}{\rho}\right)$$

In the inclusive value definition, *J* is the total number of recreation sites and ρ is the "dissimilarity coefficient" that represents the degree of substitution between trip-taking and staying home. Given these probabilities, the model can be estimated by maximizing the following likelihood function

$$L = \prod_{i=1}^{N} \prod_{j=0}^{J} \left(P_{ij}\right)^{T_{ij}}$$

where T_{ij} represents the number of choice occasions where individual *i* selected site *j*. The second product includes *j* = 0, which represents the no-trip alternative, so that T_{i0} is the number of choice occasions where individual *i* chose <u>not</u> to take a trip to a recreation site.

The compensating variation per choice occasion associated with a change in the characteristics of one or more sites can be expressed as (Hanemann 1982)

$$CV_i = \frac{\ln\left[\exp\left(v_{i0}^{\square}\right) + \exp\left(\rho I_i^1\right)\right] - \ln\left[\exp\left(v_{i0}^{\square}\right) + \exp\left(\rho I_i^0\right)\right]}{\beta_{TC}}$$

where I_i^0 represents the inclusive value for individual *i* with the original site characteristics and I_i^1 represents the inclusive value for individual *i* with the new site characteristics.

c. Degree of accuracy needed for the Purpose discussed in the justification;

The survey responses will provide recreation trip data that will be used to estimate a random utility maximization model. We will use the model to develop a variety of estimates that are complex functions of model parameters. The confidence levels for these estimates will vary with the type of estimate and with the precision of the associated model parameters. While the precision of these parameters is difficult to predict in advance, based on past experience with similar models, the study team believes that reasonably precise estimates can be obtained with 1,000 or more responses.

d. Unusual problems requiring specialized sampling procedures; and

No specialized sampling procedures will be used. Households will be selected using stratified random sampling of households from four primary strata (see item 1 above). Within households the adult (age 18 or older) with the most recent birthday will be asked to complete the on-line survey.

e. Use of periodic or cyclical data collections to reduce respondent burden.

This is a one-time survey and is therefore the most infrequent collection interval possible.

3. Maximization of Response Rates, Non-response, and Reliability

Several measures will be taken to encourage sampled individuals to respond to the survey, including:

- Branded survey materials with color USACE logos;
- Multiple follow-up reminders after the initial invitation; and,
- Provision of a toll-free number in survey correspondence to address any questions.

Despite these measures, the response rate for the web survey may be as low as 20%, raising potential concerns about non-response bias. Demographic differences between respondents and non-respondents will be addressed by calibrating design weights through iterative proportion fitting, or "raking" (Kolenikov 2014; Battaglia, Hoaglin, and Frankel 2009) to match demographic controls from the American Community Survey (e.g., gender, age, ethnicity, and education) within each of the four sampling strata.

Even after controlling for demographic differences between respondents and non-respondents, avid recreationists may be more likely to respond to the survey, which could lead to overestimates of recreation activity for the population of interest. The potential for this type of non-response bias will be investigated through a targeted non-respondent follow-up survey (NRFU). The NRFU survey will consist of a subset of questions from the main survey, including general questions about participation in outdoor recreation and demographics. The survey will be formatted as an oversized postcard and sent to a sample of 5,000 non-respondents via priority mail. Responses to the NRFU survey will be compared to responses to the main survey to assess the potential for non-response bias.

Finally, recreation trip estimates generated from the survey data will be compared to existing recreation trip estimates generated through onsite counts by federal and state agencies at various sites within the region (e.g., creel surveys of fishing effort or recreation counts at USACE reservoirs). If large differences are observed, the survey data may be calibrated so that trip estimates align more closely with these onsite counts.

4. Tests of Procedures

In early 2018, an initial round of pre-testing was conducted with seven USACE and Bonneville Power Administration (BPA) employees identified as recreators within the region and who were not involved in survey development. A second round of pre-testing was conducted with three additional USACE employees and with six volunteer participants identified by USACE as recreators within the region with no specific background or training in survey research methods or analysis (i.e., members of the general public). Individuals were administered the questionnaires verbally and then individually de-briefed to elicit feedback on questionnaire language, organization and clarity. Comments were incorporated into the revised instruments included with this submission.

5. Statistical Consultation and Information Analysis

a. Provide names and telephone number of individual(s) consulted on statistical aspects of the design.

Chris Leggett, Ph.D.

Director, Resource Systems Group

(802) 359-6418

Robert Paterson

Principal, Industrial Economics, Incorporated

(617) 354-0074

Frank Lupi, Ph.D.

Professor, Department of Agricultural, Food, and Resource Economics

Michigan State University

(517) 432-3883

b. Provide name and organization of person(s) who will actually collect and analyze the collected information.

Robert Paterson and Eric Horsch (Industrial Economics, Incorporated)

Chris Leggett (Resource Systems Group)

References

Battaglia, M. P., D. C. Hoaglin, and M. R. Frankel. 2009. "Practical Considerations in Raking Survey Data," *Survey Practice* 2(5).

Hanemann, W.M. 1982. Applied Welfare Analysis with Qualitative Response Models. Working Paper No. 241, Giannini Foundation of Agricultural Economics, University of California, Berkeley. 26pp

Kolenikov, S. 2014. "Calibrating Survey Data Using Iterative Proportional Fitting (Raking)," The Stata Journal, 14(1): 22–59.

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

Morey, E.R., R.D. Rowe, and M. Watson. 1993. A repeated nested-logit model of Atlantic salmon fishing. *American Journal of Agricultural Economics* 75(3): 78-92.

Rizzo, L., J.M. Brick, and I. Park. 2004. A minimally intrusive method for sampling persons in random digit dial surveys. *Public Opinion Quarterly* 68(2): 267-274.