

SUPPORTING STATEMENT
NORTH ATLANTIC RECREATIONAL FISHING SURVEY
OMB CONTROL NO. XXXX-XXXX

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

- 1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection method to be used. Data on the number of entities (e.g. establishments, State and local governmental units, households, or persons) in the universe and the corresponding sample are to be provided in tabular form. The tabulation must also include expected response rates for the collection as a whole. If the collection has been conducted before, provide the actual response rate achieved.**

The target population for the NARFS is recreational anglers who have fished for Atlantic cod or haddock in federal waters off the coast of Maine, New Hampshire, or Massachusetts in the past five years. Our sample frame will be drawn from 2018 recreational fishing license/registry databases maintained by the states of Maine, New Hampshire, and Massachusetts. Table 4 displays information about these databases in 2017, as the 2018 registry data is not yet complete. Note that the composition of the 2017 database is unlikely to differ much from that of 2018 database.

Given their lower per-unit cost relative to mail surveys and feedback from NARFS focus group participants, web surveys will account for about three quarters of the total number of surveys distributed.

Table 4. 2017 saltwater recreational fishing licensees by state and contact information availability.

State	Total # of licensees	Phone number	Mailing address	Email address	Mailing and email address
Maine	61,234	60,752	61,234	36,256	36,256
New Hampshire	30,348	29,805	30,348	4,648	4,648
Massachusetts	165,241	162,373	165,241	95,678	95,678

As discussed in our response to question A1, the NARFS is comparable to a survey that was conducted in 2014. The two differ somewhat in sampling methodology and scope, but the response and eligibility rate from the 2014 survey provides the basis for those expected of the NARFS.

In terms of sampling methodology, the 2014 survey was conducted as mail follow-up to NOAA Fisheries' Marine Recreational Information Program (MRIP) Access Point Angler Intercept Survey (APAIS). The sample frame for the 2014 survey consisted of 1,890 anglers who were interviewed for the MRIP APAIS in private boat or party/charter fishing modes in Maine, New Hampshire, and Massachusetts. These anglers were willing to participate in the follow-up survey and subsequently received one in the mail. As displayed in Table 5, the raw response rate for the 2014 survey was 29.3%. In contrast to the 2014 survey, the NARFS sampling methodology

involves sending mail and email invitations containing a web link to participate to a list of potential respondents drawn at random from state license/registration databases.

Table 5. Survey response rates, 2014 survey

Disposition	Number of surveys
Ineligible [†]	35 (1.9%)
Mail returned as undeliverable	232 (12.3%)
No response	1063 (56.2%)
Refusal	6 (0.3%)
Completed	554 (29.3%)
Total	1,890 (100%)

[†]Eligibility criteria defined as having fished for cod, haddock, or pollock in the past 5 years.

In terms of scope, the 2014 survey defined eligibility as having fished for cod, haddock, or pollock in the past five years, while the NARFS excludes pollock from this criteria. Fortunately, the 2014 survey collected the information necessary to give a sense of the likely proportion of anglers across the study region who will have fished for cod or haddock in the past five years. Specifically, one question in the 2014 survey asked anglers to indicate, among cod, haddock, pollock, or none of the three species, “Which of the following species have you personally caught or tried to catch in the **last five years?** (Please mark all that apply)”. Responses to this question are displayed in Table 6, which shows that 67% of the 554 anglers who completed the 2014 survey indicated having fished for cod and/or haddock in the past five years.

Table 6. Species caught or targeted in last five years, 2014 survey.

Response	Number of respondents
Pollock only, none of the 3 species, or not answered	184 (33%)
Cod and/or haddock	370 (67%)

The eligible response rate for the 2014 survey, based on the results in Table 5 and Table 6, is 19.6% (0.293×0.67). With that in mind, a conservative estimate of the eligible response rate for the NARFS is 15%.

Table 7 describes the NARFS’s respondent universe and predicted eligibility and response rates. The first row shows the total number of 2017 saltwater recreational fishing licensees by state. Multiplying these numbers by 0.67, the proportion of anglers who are expected to be eligible for the NARFS, gives an estimate of the size of the respondent universe, which is shown in the second row of Table 7. The third row of Table 7 shows the number of surveys we plan to distribute by state; these values are based on (1) an expected eligible response rate of 15%, (2) the sample stratification procedure, and (3) the total number of surveys completed by eligible anglers we intend to collect; we discuss (2) and (3) in more detail in our response to question B2.

Once the study is completed, we will calculate the final response rate using the appropriate American Association for Public Opinion Research (AAPOR) [Response Rate Calculator](#).

Table 7. NARFS respondent universe, sample frame, and expected response rates by state.

	Maine	New Hampshire	Massachusetts
Total number of 2017 saltwater recreational fishing license holders	61,234	30,348	165,241
Expected number of eligible 2017 saltwater recreational fishing license holders (universe)	41,027	20,333	110,711
Number of saltwater recreational fishing license holder contacted (sample frame)	317	937	2,080
Expected eligible response rate	15%	15%	15%
Expected number of surveys completed by eligible respondents	48	141	312

2. Describe the procedures for the collection, including: the statistical methodology for stratification and sample selection; the estimation procedure; the degree of accuracy needed for the purpose described in the justification; any unusual problems requiring specialized sampling procedures; and any use of periodic (less frequent than annual) data collection cycles to reduce burden.

Statistical methodology for stratification and sample selection

A stratified random sampling strategy will be used to focus efforts on the population of interest and reduce implementation costs. Participants will be drawn from each state license database (g) in proportion to that state’s contribution to the total number of recreational fishing trips taken during 2017 that caught or targeted Atlantic cod or haddock. Using publicly-available MRIP data, we estimated the total number of trips to be 370,243; of these trips, Maine, New Hampshire, and Massachusetts accounted for 9.5%, 28.1%, and 62.4%, respectively. If we denote these proportions as W_g , and the proportions from the realized sample as H_g , we can use stratification weights equal to W_g/H_g during estimation of the economic model to obtain consistent parameter estimates (Louiviere et al. 2000).

Estimation procedure

The most critical input for estimating the angler behavioral model are the data collected from Section B of the NARFS, the discrete choice experiment. Each question in this section presents respondents with three options—two hypothetical North Atlantic cod and haddock fishing trips that vary in catch levels, regulations, and cost, and the option to not go recreational saltwater fishing—and asks them to indicate which of the three options would be their first and second choice if they were presented with these options in the real world. We use these data to estimate a vector of parameters β' that represent marginal utilities of fishing trip attributes included in the discrete choice experiment.

We will analyze the discrete choice experiment data collected by the NARFS using random utility maximization (RUM) models, which decompose utility into its observable and unobservable components (McFadden 1973). RUM models assume that when faced with multiple alternatives, individual n will select alternative i that maximizes utility, U_i .

$$U_i > U_{nj} \quad \forall j \neq i \quad (1)$$

Partitioning U_i into its two component parts, the choice of alternative i is such that

$$V_i + \varepsilon_i > V_{nj} + \varepsilon_{nj} \quad \forall j \neq i, \quad (2)$$

where V_{nj} is a function typically specified to be linear in parameters, $V_{nj} = \beta' x_{nj}$, that relates observed attributes to utility and ε_{nj} captures the utility derived from all other unobservable factors. The utility parameters β' measure the relative importance of the attributes x_{nj} that describe alternative j . Because ε is stochastic, it is not possible to determine absolute levels of utility; however, probabilistic inference about individuals' choices can be made under the standard assumption for logit models that these terms are independently and identically distributed Type I extreme value. From Equation 2, the probability that angler n selects fishing alternative i is

$$P_i = \text{Prob} \left[(\varepsilon_i - \varepsilon_{nj}) < V(\beta' x_{nj}) - V(\beta' x_i) \right] \quad \forall j \neq i \quad (3)$$

Train (2003) calculates this probability for a multinomial logit (MNL) model as

$$P_i = \frac{e^{\beta' x_i}}{\sum_j e^{\beta' x_{nj}}} \quad (4)$$

The probability of each individual in the sample choosing the alternative they were observed to actually choose is

$$L(\beta') = \prod_{n=1}^N \prod_i (P_i)^{y_i}, \quad (5)$$

where $y_i = 1$ if an individual is observed to choose alternative n , and zero otherwise. We will obtain, using maximum likelihood procedures, estimates of the parameter vector β' that maximize the log likelihood function

$$\ln L(\beta') = \sum_n \sum_i y_i \ln \left(\frac{e^{\beta' x_i}}{\sum_{i=1}^J e^{\beta' x_{nj}}} \right) \quad (6)$$

Fishing trip utility will be modeled as a function of the number of North Atlantic cod and haddock kept and released, the trip length, the trip cost, and an indicator variable for the “do not go fishing” alternative. To allow for diminishing marginal utility of catch, as is common in the

recreational demand literature, each catch variable enters the model as its square root. Below is our baseline model of fishing trip utility. Preliminary model testing will dictate the control variables derived from Sections A, C, and D of the NARFS that we include the model.

$$\begin{aligned}
 U_{nj} = & \beta_1 \sqrt{Cod_{keepnj}} + \beta_2 \sqrt{Haddock_{keepnj}} \\
 & + \beta_3 \sqrt{Cod_{releasenj}} + \beta_4 \sqrt{Haddock_{releasenj}} \\
 & + \beta_5 Trip_{lengthnj} + \beta_6 Trip_{Costnj} + \beta_7 No_{fishnj} + \varepsilon_{nj}
 \end{aligned} \tag{7}$$

The bioeconomic model is parameterized with the estimated marginal utilities, β_1 through β_7 . It accounts for randomness in catch by simulating trip-level outcomes based on probabilistic catch-per-trip and catch-at-length distributions that are fitted with historical recreational fishing data. On each simulated trip, the model imposes and determine the relative effect of alternative management options for the recreational North Atlantic cod and haddock fisheries on angler welfare and participation. Estimates of angler participation are used to evaluate policy-induced impacts to recreational fishing mortality and future stock levels. A detailed description of how the utility parameters described above are integrated into the bioeconomic model is given in Lee et al. (2017).

The degree of accuracy needed for the purpose described in the justification

The number of completed surveys needed to estimate the behavioral model parameters with adequate precision is based the experimental design of the CE. We evaluated the minimum sample size required for the ensuing economic model using two methods: (1) a general rule of thumb, and (2) an approach based on the statistical power of hypothesis tests for estimated coefficients of interest.

The rule-of-thumb for determining the minimum sample size for CE modelling, given by Orme (2010) is to: set

$$\frac{nta}{c} \geq 500 \tag{8}$$

where n is the number of respondents, t is the number of choice tasks per respondent, a is the number of alternative per choice task excluding the opt-out alternative, and c is the largest number of levels for any one attribute for a main effects model. While the value of 500 refers to the number of times each main effect level of interest should be represented across the design to have ample stability in the ensuing parameter estimates, Orme (2010) notes that this value “was intended to be a minimum threshold when researchers cannot afford to do better. It would be better, when possible, to have 1,000 or more representations per main-effect level”. When its right-hand side set to 500, solving Equation (8) for n based on our main-effects experimental design ($t = 6$, $a = 2$, $c = 6$) yields 250; when the right-hand side of Equation (8) is set to 1,000, $n = 500$. Taken together, the general rule of thumb provided by Orme (2010) suggests that the minimum sample size for the NARFS should be between 250 and 500.

The second method for determining the minimum sample size incorporates prior knowledge about the parameters we wish to estimate using the NARFS data. We follow the step-by-step approach outlined in de Bekker-Grob et al. (2015), which requires five elements:

- 1) Significance level (α): the probability for an incorrect rejection of a true null hypothesis.
- 2) Statistical power level ($1 - \beta$): β indicates the probability of failing to reject a null hypothesis when the null hypothesis is actually false, with its chosen value related to the statistical power of the test ($1 - \beta$).
- 3) Statistical model used in the DCE analysis [e.g., multinomial logit (MNL) model, mixed logit (MIXL) model, generalized multinomial logit (G-MNL) model]
- 4) Initial belief about the parameter values
- 5) The DCE design

Our selection of the elements in 3), 4), and 5) are guided by the statistical analysis of the 2014 angler survey data that employed a MNL model specification. The estimated parameters from this model form our initial beliefs about the parameter values, which we subsequently use to generate an efficient experimental design containing 72 choice questions. With these ingredients, we calculate the variance-covariance matrix of the estimated parameters, $\Sigma_{\hat{\gamma}}$, and use the diagonal elements of this matrix for hypothesis tests on individual coefficients. The minimum sample size (N) for the estimated coefficients in a DCE can be determined using following equation.

$$N > \frac{z_{1-\beta}^2 + z_{1-\alpha}^2}{\delta^2} \tag{2}$$

where the values of α and β are used to determine the corresponding quantile of the Normal distribution $z_{1-\beta}$ and $z_{1-\alpha}$ required for sample size calculations, respectively, and δ is the effect size of attribute k , which is set to the parameter estimate obtained from the 2014 angler survey analysis.

Table 8 displays equation (2) calculations based on several values of α and $1-\beta$ and across all choice experiment attributes. The largest minimum sample size displayed in Table 8 is 238. This value indicates that, with a statistical power of 0.9 and assuming $\alpha = 0.01$, a minimum sample size of 238 respondents is needed to discern whether the marginal utility associated with a one-hour increase in the length of a North Atlantic cod and haddock fishing trip is significantly different than zero.

Given results of both approaches used to determine the minimum sample size required for the NARFS, we seek to obtain at least 500 surveys completed by eligible anglers. Meeting this objective conditional on an expected 15% eligible response rate requires distributing a minimum of 3,333 surveys across the three states. This information is incorporated with our sample stratification procedure and shown in the third row of Table 7.

Table 8. Minimum sample size required to obtain the desired power level $1-\beta$ for finding an effect when testing at a specific confidence level α .

α	$1-\beta$	Choice experiment attributes				Trip length	Trip Cost	No Fish
		$\sqrt{Cod\ kept}$	$\sqrt{Had.\ kept}$	$\sqrt{Cod\ released}$	$\sqrt{Had.\ released}$			
0.1	0.6	3	3	3	28	43	39	43
0.05	0.6	4	4	4	43	66	59	66
0.025	0.6	6	5	6	59	90	81	89
0.01	0.6	8	7	8	80	122	110	121
0.1	0.7	4	4	4	39	60	54	59
0.05	0.7	6	5	5	56	86	78	86
0.025	0.7	8	7	7	74	113	102	112
0.01	0.7	10	9	9	97	149	134	148
0.1	0.8	6	5	5	54	83	74	82
0.05	0.8	8	7	7	74	113	102	112
0.025	0.8	10	9	9	94	144	129	143
0.01	0.8	12	11	12	120	184	166	183
0.1	0.9	8	7	8	79	120	108	120
0.05	0.9	11	9	10	103	157	141	156
0.025	0.9	13	11	12	126	192	173	191
0.01	0.9	16	14	15	156	238	215	237

3. Describe the methods used to maximize response rates and to deal with nonresponse. The accuracy and reliability of the information collected must be shown to be adequate for the intended uses. For collections based on sampling, a special justification must be provided if they will not yield "reliable" data that can be generalized to the universe studied.

Implementation of the NARFS will follow procedures suggested by Dillman et al. (2009) that have been used extensively in survey research to increase response rates for mail and web surveys. Repeated contacts with respondents through one or more mediums (phone, email, and mail), which allow flexibility regarding how potential respondents choose to respond, and small participation incentives both serve to maximize the overall survey response rate. While subject to minor changes, the implementation schedule for the NARFS based on recommendations from Dillman et al. (2009) is shown in Figure 2. Each step from Figure 2 is described in more detail below. Materials for all tracks and contacts, as well as the NARFS survey instrument, are included in the accompanying supporting documents file.

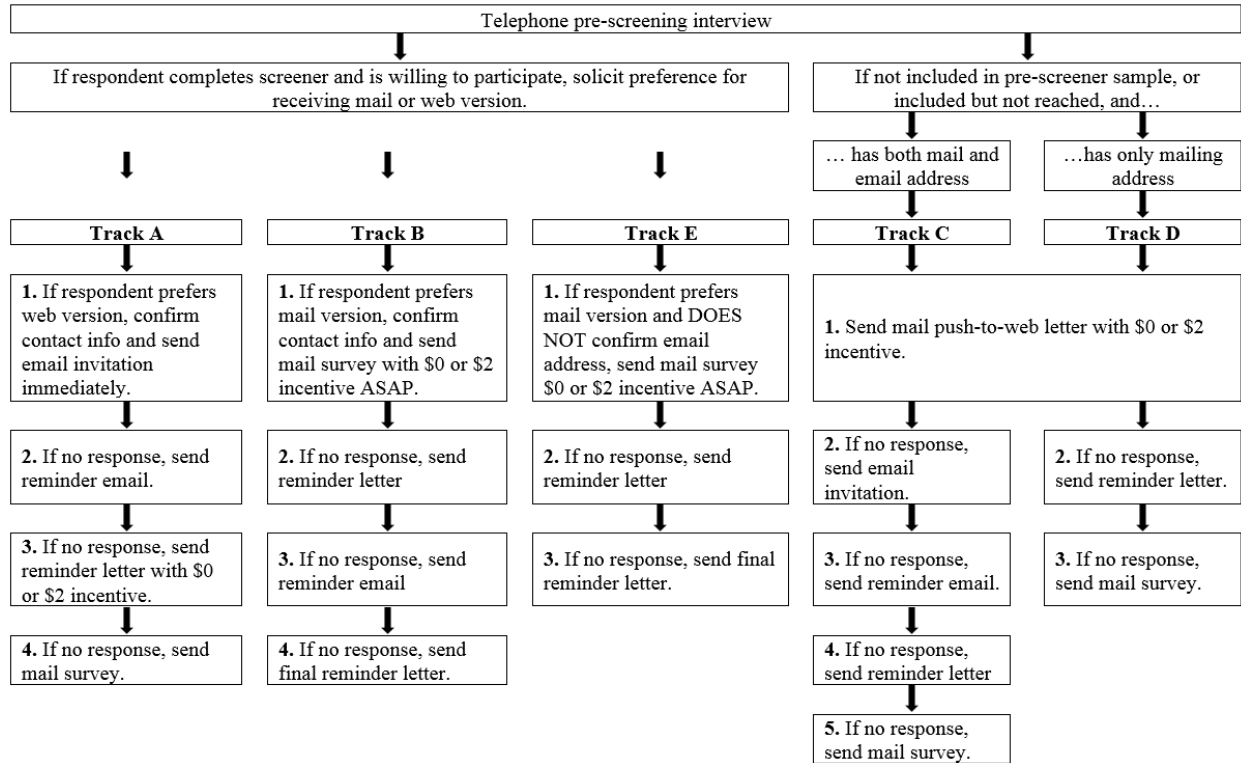


Figure 2. Flow chart of NARFS sampling procedure.

Contact 1: Telephone Screening Interview

The first point of contact with potential respondents for whom telephone contact information is available, a telephone pre-screening interview, will ensure that the NARFS is distributed to individuals who have fished for cod and haddock within the past five years. The NEFSC will attempt telephone pre-screening interviews with 2,000 license holders among of the initial sample of 4,000. Eligible respondents who complete the telephone pre-screening interview will be asked whether they are willing to participate in the follow-up NARFS survey and if so, whether they prefer a mail and/or web version of the NARFS.

The telephone pre-screening interview will collect socioeconomic and fishing-related information, Z , such that comparisons between survey respondents and non-respondents can be made. These comparisons will inform the subsequent use of weighting class adjustments to mitigate the biasing effects of response propensities that are a function of Z (Groves 2006). The telephone pre-screening interview will gather the following information:

- Number of recreational fishing trips taken in the past 12 months
- Number of recreational fishing trips for cod or haddock taken in the past 12 months
- Likelihood of cod or haddock fishing in the next 12 months
- Age
- Income

Additionally, recent experience suggests that a telephone pre-screening interview will boost the overall survey response rate. In a 2016 striped bass angler survey directed by members of the

proposed research project, telephone pre-screened license holders were more likely to complete the survey (55%) than those who were not screened (29%). The completion rate for screened anglers and other telephone pre-screening interview results for the 2016 striped bass angler survey are displayed in Table 9.

Table 9. Telephone pre-screening survey results, 2016 striped bass survey.

Final screener survey disposition	# of licensees	# licensees that subsequently completed full survey [†]
Answering machine	816	236
Bad/disconnected number	197	33
Busy signal	36	8
Completed screening	318	175 (55% of 318)
Language barrier	53	3
No Answer	92	21
Ineligible: Has not fished for Striped Bass in past 3 years	252	0
Ineligible: Respondent deceased	5	0
Refusal: eligible for mail/web survey	3	0
Respondent not available during dialing period	25	10
Refusal: Completed screening, but refused mail/web survey	7	0
Refusal: eligibility unknown	191	35
Wrong Number	90	13
Total	2085	534

[†]The overall response rate cannot be inferred from this table, as the survey was sent to a supplemental web-only sample.

Track A: Web survey to screened anglers who request web version of the NARFS

Those who indicate a preference for a web version will receive an email invitation to participate in the NARFS immediately after confirming their email contact information. The email will introduce the survey, demonstrate its relevance, encourage participation, and assure confidentiality invitation. It will include the follow information:

- The survey’s purpose and how the results will be used;
- A request for the respondent’s participation and a statement of promised confidentiality or anonymity;
- Detailed instructions for accessing the survey using a URL and unique passcode;
- A phone number and e-mail address to use for technical support or if a respondent has questions about the study’s validity.

Subsequent contacts with Track A participants who do not respond to the initial email invitation are as follows:

- A reminder email, which reinforces the importance of the survey and reminds those who did not respond about the survey. The reminder email will contain a web link to the survey.
- A reminder letter, which reinforces the importance of the survey and reminds those who did not respond about the survey. The reminder letters will contain a web address to access the survey; some will also include a \$2 incentive.
- A mail package that includes a paper version of the NARFS and a cover letter explaining the importance of the survey.

Track B: Mail survey to screened anglers who request mail version of the NARFS

Eligible respondents who complete the telephone pre-screening interview and request mail version of the survey will receive a NARFS survey in the mail; some of these respondents will also receive a \$2 incentive with this survey package. Subsequent contacts with Track A participants who do not complete initial survey mailing are as follows:

- A reminder letter, which reinforces the importance of the survey and reminds those who did not respond about the survey. The reminder letter will contain a web address to access the survey online.
- A reminder email, which reinforces the importance of the survey and reminds those who did not respond about the survey. The reminder email will contain a web link to the survey.
- A final reminder letter.

Track E: Mail survey to screened anglers who request mail version of the NARFS and do not confirm email address

Eligible respondents who complete the telephone pre-screening interview, do not confirm their email address, and request mail version of the survey will receive a NARFS survey mail; some of the mailings will include a \$2 incentive. Subsequent contacts with Track E participants who do not complete initial survey mailing are similar to those in Track B, with the exception that these participants will not receive a reminder email.

Track C: Mail push-to-web letter will to unscreened anglers and anglers not reached for telephone pre-screening interview who have both mail and email addresses on file.

Respondents not included in the telephone pre-screening sample (2,000) plus those who could not be reached for the telephone pre-screening interview will be sent an initial mail push-to-web letter. The mail push-to-web letter provides an opportunity to explain the survey's purpose and elicit cooperation. The letters will introduce the survey, demonstrate its relevance, encourage web participation, and assure confidentiality. The mail push-to-web letter will include the same information as the initial email invitation used in other tracks to encourage participation; some of these letters will also include a \$2 incentive. Subsequent contacts with Track C participants who do not respond after receiving the mail push-to-web letter are as follows:

- An email invitation, which reinforces the importance of the survey and reminds those who did not respond about the survey. The reminder email will contain a web link to the survey, in the hopes that that will make it easier for participants to respond online.
- A reminder email, which reinforces the importance of the survey and reminds those who did not respond about the survey. The reminder email will contain a web link to the survey.
- A reminder letter, which reinforces the importance of the survey and reminds those who did not respond about the survey. The reminder letter will contain a web address to access the survey online.
- A mail package that includes a paper version of the NARFS and a cover letter explaining the importance of the survey.

Track D: Mail push-to-web letter will to unscreened anglers and anglers not reached for telephone pre-screening interview who have only mail address on file.

Track D is very similar to Track E, with the exception that these participants will not receive email contacts.

4. Describe any tests of procedures or methods to be undertaken. Tests are encouraged as effective means to refine collections, but if ten or more test respondents are involved OMB must give prior approval.

To test the efficacy of the proposed survey instrument, we conducted focus group sessions with recreational cod and haddock anglers in the study region during November of 2018. Using the license databases described in B1, we recruited focus groups participants who differed in terms of gender, age, and Atlantic cod/haddock fishing experience to obtain feedback from a diverse mix of anglers. The first two of these focus groups were held in Braintree, MA, and the second two groups were held in Portland, ME.

We used feedback received from these focus groups to (a) conform survey language to regional differences in dialect and ensure consistent interpretation of survey questions across the study region, and (b) design contextually realistic and straightforward choice experiment questions. As we continue to develop the survey instrument, any additional concerns encountered will be addressed through additional focus group interviews.

5. **Provide the name and telephone number of individuals consulted on the statistical aspects of the design, and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.**

Design, analysis, report

Andrew Carr-Harris (co-P.I.)
Northeast Fisheries Science Center
Phone: 914-330-7881

Scott Steinback (co-P.I.)
Northeast Fisheries Science Center
Phone: 508-495-4701

Sabrina Lovell (co-P.I.)
Office of Science and Technology
Phone: 301-427-8153

Data collection

Gustavo Rubio
ECS Federal, contracting company
Phone: 301-427-8180

REFERENCES

- Bekker-Grob, Esther W. de, Bas Donkers, Marcel F. Jonker, and Elly A. Stolk. 2015. "Sample Size Requirements for Discrete-Choice Experiments in Healthcare: A Practical Guide." *Patient* 8 (5): 373–84.
- Dillman, Donald A., Jolene D. Smyth, and Leah M. Christian. 2009. *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. 3rd ed. New York: Wiley.
- Groves, Robert M. 2006. "Nonresponse Rates and Nonresponse Bias in Household Surveys." *Public Opinion Quarterly* 70 (5): 646–75.
- Lee, Min-Yang, Scott Steinback, and Kristy Wallmo. 2017. "Applying a Bioeconomic Model to Recreational Fisheries Management: Groundfish in the Northeast United States" 32 (2).
- Louviere, Jordan J., David A. Hensher, and Joffre D. Swait. 2000. *Stated Choice Methods: Analysis and Applications*. Cambridge University Press.
- McFadden, Daniel. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, 105–142. New York.
- Train, Kenneth. 2003. *Discrete Choice Methods with Simulation*. New York: Cambridge University Press.