NOAA Technical Memorandum NMFS-NE-226

doi:10.7289/V59021QG

An Overview of the Annual Cost Survey Protocol and Results in the Northeast (2007 to 2009)

US DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Northeast Fisheries Science Center
Woods Hole, Massachusetts
November 2013

NOAA Technical Memorandum NMFS-NE-226

doi:10.7289/V59021QG

An Overview of the Annual Cost Survey Protocol and Results in the Northeast (2007 to 2009)

Chhandita Das

NOAA Fisheries, Northeast Fisheries Science Center/Integrated Statistics,
13 Church Street, Woods Hole, MA 02543

US DEPARTMENT OF COMMERCE
Penny Pritzker, Secretary
National Oceanic and Atmospheric Administration
Kathryn Sullivan, Acting NOAA Administrator
National Marine Fisheries Service
Samuel D. Rauch III, Acting Assistant Administrator for Fisheries
Northeast Fisheries Science Center
Woods Hole, Massachusetts
November 2013

Editorial Notes

Information Quality Act Compliance: In accordance with section 515 of Public Law 106-554, the Northeast Fisheries Science Center completed both technical and policy reviews for this report. These predissemination reviews are on file at the NEFSC Editorial Office.

Species Names: The NEFSC Editorial Office's policy on the use of species names in all technical communications is generally to follow the American Fisheries Society's lists of scientific and common names for fishes, mollusks, and decapod crustaceans and to follow the Society for Marine Mammalogy's guidance on scientific and common names for marine mammals. Exceptions to this policy occur when there are subsequent compelling revisions in the classifications of species, resulting in changes in the names of species.

Statistical Terms: The NEFSC Editorial Office's policy on the use of statistical terms in all technical communications is generally to follow the International Standards Organization's handbook of statistical methods.

Internet Availability: This issue of the NOAA Technical Memorandum NMFS-NE series is being as a paper and Web document in HTML (and thus searchable) and PDF formats and can be accessed at: http://www.nefsc.noaa.gov/nefsc/publications/.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Fisheries vessel cost data are used in virtually every framework, amendment and fishery management plan that requires economic analyses. This document summarizes the results of the annual commercial vessel cost data collection program conducted by the staff of the Social Sciences Branch of the NMFS Northeast Fisheries Science Center in 2007, 2008, and 2009. Information is presented on data collection methodology, data coverage, and data quality. Summary statistics are provided by vessel characteristics. Cost estimation and prediction methodologies are also discussed. The analyses and findings from the 2007-2009 programs are expected to improve future data collection efforts, and to enhance future analyses and evaluations of the economic status of commercial fisheries in the Northeast.

1 INTRODUCTION

The Social Sciences Branch (SSB) of the NMFS Northeast Fisheries Science Center (NEFSC) has collected annual cost information from commercial fishing vessel owners for many years through several initiatives. The latest effort was undertaken in 2007 when cost information was requested from vessel owners via a voluntary mail-in survey. This effort continued for two more years through 2009. This document describes the survey and presents information on the data, analyses, and initial findings from the data collection effort.

2 OVERVIEW OF THE COST SURVEY

2.1 Objective

Economic data on the costs of operating commercial fishing businesses are needed to meet the legislative requirements of the Magnuson-Stevens Fishery Conservation and Management Act, the National Environmental Policy Act, Executive Order 12866, and the Regulatory Flexibility Act. Fisheries cost data are used in the economic analyses in many frameworks and amendments to fishery management plans. Examples include regulatory impact analyses, economic profitability profiles, fleet efficiency and productivity measures estimations, and economic impact evaluations of proposed management measures and regulations. An accurate understanding of the financial costs incurred by commercial fishing businesses is critical for these analyses. The objective of this study was to obtain reliable, timely, and updated information on commercial fishing vessel costs.

2.2 Survey Methodology

2.2.1 Survey Design and Fielding

The Social Sciences Branch at the Northeast Fisheries Science Center has been collecting fishing cost information for many years, using both formal and informal methods. In 2007, SSB initiated a formal data collection process that continued through 2009 in which cost information was acquired via voluntary mail-in surveys. The survey requested vessel owners to report their annual costs for the year preceding the survey year. For example, the survey in year 2007 requested vessel costs incurred during 2006. Reference to survey years in this report pertains to the years in which the costs were incurred (i.e., survey years 2006, 2007 and 2008). A copy of the 2008 survey is presented in the appendix.

The survey was administered by the Permit Office within NOAA's Northeast Regional Office (NERO). Each year during the 3-year study period, SSB provided the NERO Permit Office with a list of hull numbers of the vessels selected to be surveyed. The Permit Office then mailed the surveys to the prospective vessel owners along with their permit renewal forms. A stamped self-addressed return envelope was also included in the survey packet, which contained a cover letter that clearly stated (a) the objective of the survey and (b) that participation in the survey was voluntary.

2.2.2 Population and Sample

The unit of observation of the survey was the fishing vessel, and the survey population comprised all active commercial fishing vessels in each of the three years. In each survey year, an

active fishing vessel was defined as one that held at least one Northeast Federal Fishing Permit and reported landings of at least one pound of fish through the Northeast Seafood Dealer Reporting System. These criteria resulted in a population of 3,055 vessels in 2006, 2,597 vessels in 2007, and 2,879 vessels in 2008. Each year, surveys were sent to all the vessels in the population for that year.

2.3 Response Rates

For year 2006, 630 completed surveys were returned for a response rate of 20.6%. The response rate declined to 16.5% (430 responses) for 2007 and to 8% (232 responses) for 2008. With survey data, analysts are often concerned about obtaining a sufficient response rate in order for the information collected to be scientifically representative and precise. There are different hypotheses in the literature regarding this response rate. For example, Dillman et al. (2009) provides the following formula to determine the sample size needed to make population estimates within a selected level of confidence interval:

$$N_s = \frac{(N_p)(p)(1-p)}{(N_p-1)(B/C)^2+(p)(1-p)} \quad (1)$$

Where: N_s = the completed sample size needed for the desired level of precision.

N_p = the size of the population

p = a measure of population heterogeneity

B = margin of error; (i.e., half of the desired confidence interval width), 0.03=±3%

C= Z score associated with the confidence level (1.96 corresponds to the 95% confidence interval).

Based on this formula, and assuming maximum heterogeneity in the population (p=0.5 implying 50/50 split in population) and a confidence interval of 95%, the number of completed responses needed to be able to estimate the characteristic of the population within ±3% points 95% of the time, was 341 for 2006, 335 in 2007 and 339 for 2008. Hence, the completed survey responses for 2006 and 2007 exceeded the minimum baselines in these years, but fell short in 2008.

Table 1[1] lists the overall survey population sizes and response rates, as well as by gear category, region, vessel length category, and revenue category. Vessel characteristics were obtained from the permit data base, the vessel logbook data base and the dealer data base. The principal gear (and landing region) by a vessel was determined by the gear (or landing state) that accounted for the maximum revenue share for that vessel. Six principal gear categories were

[1] Note, the population sizes in Table 1 are different from the population sizes reported under subsection *"2.2.2 Population and Sample."* This difference is because responses were received from vessel owners who were not included in the original sample. This irregularity might be because some vessel owners received the surveys through their business partner or via a shared address and chose to respond. These additional responses were considered to be part of the population, resulting in a slightly higher population size in Table 1 than in subsection *"2.2.2 Population and Sample."*

used: *Dredge*, *Gillnet*, *Handgear*, *Longline*, *Pot/Trap*, *Trawls*, and *Others*. Two principal landing regions were defined: *Mid-Atlantic* and *New England*. The *Mid-Atlantic* region includes the states of New Jersey, Maryland, Virginia, New York, Delaware, and North Carolina. The *New England* region includes Massachusetts, Rhode Island, Maine, New Hampshire, and Connecticut. Three vessel length categories were established: *Large* (greater than 80 ft), *Medium* (40- 80 ft), and *Small* (less than 40 ft). Total vessel revenues were grouped into five categories: *less than $25 thousand*, *$25 thousand to $100 thousand*, *$100 to $500 thousand*, *$500 to $1 million*, and *over $1 million*. The unclassified category consists of vessels that either had missing information, or lacked a match with the vessel logbook or the dealer data - and therefore characteristic variables could not be defined for these vessels.

A returned survey was regarded complete if it had at least one cost-related question answered. Overall, response rates declined from 2006 through 2008 and across all categories. For example, in 2006, the response rate was 16% for *Dredge* vessels, whereas in 2008, the response rate for *Dredge* vessels was only 8%. A similar pattern in declining annual response rates occurred in every category except for the *Others* principal gear category, where response rates increased from 10% in 2006 to 18% in 2007 but declined in 2008 to 7%.

The rate of response differs largely across categories, which may lead to nonresponse bias. As nonresponse bias may lead to biased inferences, it is extremely important to test for the existence of nonresponse bias and correct for it.

2.4 Nonresponse

The previous section showed high and varied rates of nonresponse in the surveys. The focus of this section is to understand the nature and degree of this nonresponse to improve future data collection efforts. There are two types of nonresponse: one is *unit nonresponse*, where no response is received from a surveyed vessel owner; the other is *item nonresponse*, where a vessel owner responds with some−but not all−information (Lohr 2010). Each type of nonresponse is discussed below.

2.4.1 Item Nonresponse

In the cost survey, receiving an incomplete survey is not unexpected as it is quite possible for a respondent to not have any costs in some expense categories included in the survey. In 2007, only 17 surveys were received in which replies were provided for all of the survey questions; in 2006 and 2008, all returned surveys had some missing information. Although the survey allowed respondents to indicate a zero expense in a category by choosing the option “NA" (Not Applicable), only a few respondents chose to respond in this manner. That is, several respondents neither reported a cost nor chose “NA.” For these responses, it is not clear whether the missing information was due to nonresponse or because no cost was incurred in that particular category. In the absence of further information to make this distinction, no action was taken to correct for these *item nonresponses*.

2.4.2 Unit Nonresponse

Unit nonresponse can arise from several reasons: (a) the respondent may not be willing to provide the information; (b) the respondent may be unable to provide the information; and (c) survey fatigue may have occurred. From an analytical perspective, it is not possible to test for the first two reasons. However, some insight can be gained regarding the third reason by examining

the pattern of responses over time. Because the entire population of vessels was surveyed in each of the three years of the cost collection effort, it was possible for a vessel owner to receive the survey multiple times. Tables 2 and 3 examine if these repetitive survey attempts might have led to survey fatigue and thus nonresponse.

Table 2 the number of vessels owners who received the survey in multiple years. In 2007, surveys were sent to 2,597 vessel owners, of which 2,150 had received the survey in 2006. In 2008, surveys were sent to 2,879 vessel owners, of which 150 had received the survey also in 2006, 323 had received the survey in 2007 as well, and 1,791 had received the survey in both 2006 and 2007.[2] Table 3 reports a summary of the order and frequency at which a respondent received the survey and subsequently responded. Of the 1,491 vessel owners (36%) that received the survey only once during the three years, 107 (7%) responded and 1,384 (93%) did not. Of the 1,791 vessel owners (43%) who received the survey in all three years, only 65 (4%) responded in all three years, 205 (11%) responded in two of the three years, and 336 (19%) responded only once.

The order of responses helps to understand if a respondent's willingness to return the survey depended on whether the owner had received the survey in the previous year or not. Of the 832 vessel owners who received the survey twice, 147 responded only once (Table 3). Of these 147 respondents, 113 responded during the first year in which they received the survey and 34 of them responded in the second year. That is, overall, 158 (113+45) responded in the first year, and 79 (34+45) responded in the second year. Of the 1,791 vessel owners who received the survey in all three years, 336 responded only once. Of these, 209 responded in the first year only, 73 responded in the second year only, and 54 responded in the third year. Among the 205 vessels who responded twice, 157 responded in the 2006 and 2007, 15 responded in both 2007 and 2008, and 33 responded in 2006 and 2008. Overall, this indicates that 464 owners (209+157+33+65) responded in the first year, 310 (157+73+15+65) responded in the second year, and 167 (54+15+33+65) responded in the third year. Response rate did increase with the number of times a vessel owner received the survey suggesting that multiple annual attempts succeeded to some degree in generating new responses. However, the number of new responses gained in subsequent years by multiple attempts was outweighed by the number of responses lost. This might be because vessel owners developed fatigue by receiving multiple surveys and accordingly decided not to respond in subsequent years. The next section specifically tests for existence of nonresponse biases.

2.5 Nonresponse Bias Test

In the previous sections, it was noted that the survey response rate declined in 2007 and 2008, and was below 10% in 2008. This section specifically tests for nonresponse bias in the data. Because an ample amount of information exists on the characteristics of the vessels in the survey populations, it is possible to compare survey non-respondents with survey respondents to determine if there are significant differences between the two groups. To accomplish this, two sample t-tests were conducted in each year comparing the two groups based on vessel length, vessel horsepower, vessel gross tonnage, and total gross revenue. Significant differences were detected between respondents and non-respondents with respect to almost all four vessel characteristics, and this was consistent in all three years except for vessel horse power in 2008 and total gross revenue in 2007 (Table 4). These results suggest that nonresponse is an issue in

[2] 2,150 from cell (1,2) in table 2 includes the figure 1,791 from cell (2,4)

the survey data, and hence, if possible, it would be desirable to adjust the survey responses for this bias.

3 DATA

The survey asked cost-related questions under four broad headings. The cost-related questions included expenses on quota or days-at-sea leasing, vessel haul-out, repair-maintenance, vessel mooring or dockage, business vehicle use, business travel, office, communication, business taxes, permit fees, professional fees, association fees, interest and principal loan payments, catch handling, and non-crew labor services. Crew expenses consisted of crew payments and benefits. Respondents were also allowed to report their total Improvement Costs under six categories and to provide descriptions of the improvement on their fishing vessels. If respondents incurred costs that did not fit under any of the questions asked in the survey, they could report these costs under Other Costs and also describe these costs.

Before using the survey data for analysis, it was necessary to audit the data for possible data entry errors, logical errors, and outliers. The first step of the data auditing process involved a careful review of the data and eliminating any data inconsistencies, mislabeling of answers, and any double counting. In the next step, each individual cost element was plotted to identify outliers, which were subsequently removed from the final analysis. The following section elaborates on the first step in the data auditing process.

3.1 Data Auditing

The data auditing process was conducted after several rules were developed. The rules were crafted after carefully examining each and every response, and were guided by the descriptive responses from the survey. This assessment was then used to revise the existing data to facilitate better analysis and inference. The initial auditing process and rules are explained below:

1. Improvement cost-related responses were examined first. The survey questionnaire allowed respondents to enter up to six types of improvement costs and provide descriptions associated with these costs. Inspection of the descriptions revealed that many respondents were confused between improvement costs and repair maintenance costs, and thus described a repair or maintenance cost as an improvement cost. To correct these inconsistencies, the costs were moved to the correct category. However, if the respondents entered the exact same cost under each category, the entries were considered to be duplicates (i.e., double counting). In such cases, the correctly labeled cost was kept and the other entry deleted.
2. Some respondents also reported costs under “improvement costs" which were asked somewhere else in the survey. In these cases, the costs were moved from “improvement costs" to the correct category. Again, if there was sufficient evidence that double counting occurred (based on the values and description), the extra entry was deleted. If the cost descriptions were not clear enough to make such judgments, no changes were made.
3. A similar approach was taken to edit the “other costs" category. Respondents were allowed to report and describe any fishing expenses that were not specifically covered in the survey

under the label “Other Cost.” Examination of the data revealed that a large number of respondents entered trip costs as “other costs.” As trip-related costs collected by the observer program are not considered part of the annual costs, these were excluded from the analyses. To identify double counting, the approach used for “improvement costs" was followed.

4. If a cost reported under “Other Costs" was not generally considered to be part of the annual costs (such as aquaculture costs), these were also removed from the data set.

3.2 Data Summary

This section provides an overview of the revised data. This revised final data set comprises 1,153 responses: 569 responses from 2006, 383 from 2007, and 201 from 2008. To enable comparisons across years, the 2007 and 2008 cost data were converted into 2006 dollars.[3] Next, the data were weighted to correct for statistically significant nonresponse bias (as discussed in section 2.5). Because the response rates across gear vessel categories varied widely in all three years (i.e., in 2006, *Dredge* gear responses were under-represented, while in 2008 *Trawl* gear responses were over-represented), the data were weighted by gear category and summary statistics calculated based on the weighted data. The weighting procedure is explained in the appendix.[4]

For analytical purposes, individual cost items were grouped in one of three major cost categories: (a) *Repair, Maintenance, and Improvement Cost*, together denoted as *RMI* (Lian 2010); (b) *Crew Cost*; and (c) *Other Annual Cost*. *Crew Cost* is the sum of crew payments and crew benefits, while *Other Annual Cost* is the sum of the individual cost items not included in *Crew Cost* and *RMI Cost*. The *Total Annual Cost* represents the sum of all individual cost items. For each survey year, average (per vessel) costs incurred in each of the three major cost categories are listed in Table 5. Also listed is the average total revenue in each year.

In all three years, *Crew Cost* accounts for half or more of the *Total Annual Cost*, followed by *Other Annual Cost* and *RMI Cost*. In Table 6, the averages of these different types of annual costs are depicted for each of the three vessel length categories. In all cost categories and years, average costs are highest for the *Large* vessels and lowest for the *Small* vessels.

In Figures 1 to 4, graphical representations of the major costs by vessel length category are depicted with boxplots. Boxplots are a convenient way of illustrating the spread of the data and any observations that might be considered outliers. The spacing between different parts of the box helps indicate the degree of dispersion in the data. The bottom and top of the box are 25th and 75th percentiles, respectively, and the band in the middle is the 50th percentile. The bottom and top whiskers are within 1.5 times the inter-quartile range from the lower and upper quartile, respectively. The figures show that for each cost type, the costs are more dispersed for large vessels than for smaller vessels. As well, the dispersion of *Crew Cost* distribution is higher than *Other Annual Cost* and *RMI Cost* distributions. Possible outliers exist mostly for *Medium* and *Small* vessels, as indicated by those observations lying beyond the top and bottom whiskers.

[3] Producer Price Index for unprocessed finfish is used to make this conversion.
[4] The *Unclassified* geargroup was not used for calculating the weights or included in the summary statistics.

Individual cost components within each of the three general cost categories are listed in Tables 7 and 8.[5] These tables summarize individual cost items by vessel length and survey year. Average expenses in all three years are highest for crew payments, improvements, repair-maintenance, and insurance. Mean expenses on crew benefits, interest payments, lease costs, and permit fees vary considerably within years. In 2008, vessels spent a large amount on interest compared to other years.[6]

3.3 Crew Share System

The crew share system is an important element in the commercial fishing business. The 2006 survey contained questions on total crew payments and crew benefits, but did not request vessel owners to provide details on their crew share systems. Questions on the crew share system (also called the lay system) were added in the 2007 and 2008 surveys. In 2007, 88% of respondents answered the questions on the lay system, and 91% responded in 2008 (Table 9). Of the 2007 respondents, 73% reported their lay system to be either "Broken Lay" (gross earning is split between boat and crew; then trip expenses are deducted from the crew's share) or "Clear Lay" (trip expenses are deducted from the gross earning; then split between boat and crew). In 2008, this percentage was 77%. Although most vessels owners used "broken lay" to pay their crew, there are a variety of methods for compensating hired crews.

In 2007, 81% of the respondents to the question on Captain Status indicated that they were the owner-operator of their vessel (Table 9). In 2008, only 71% indicated they were owner-operators. On a percentage basis, the average boat share was higher when a vessel was owner-operated than when it was operated by a hired captain (Table 10). The average percentage boat share for an owner-operated vessel was 52% in 2007 and 60% in 2008. For vessels operated by a hired captain, the mean percentage crew share was higher than the boat share in both 2007 and 2008 (i.e., ≈51% in each year). In Table10, the percentage values listed in the aggregate row also include those vessels for which no information was provided on captain status but other crew pay related answers were provided.

Total crew costs of a vessel obviously depend on the number of hired crew. In 2007, 97% of respondents reported crew size information, and 94% of respondents in 2008 provided such information. The maximum reported crew size was 8 in 2007 and 7 in 2008. The mean crew size in both years was between 2 and 3.

3.4 Data Weighting Procedure

The data summaries presented in this document are weighted by gear category to correct for under-representation and over-representation of certain gear categories. The weight factor (w_{hi}) for respondent i in gear category h is the reciprocal of its inclusion probabilities (π_i), i.e.; $w_{hi}=1/\pi_{hi}$. The inclusion probability of respondent i is calculated as n_h/N_h, where n_h is the number of respondents and N_h is the size of the population in gear category h (Lohr 2010).

[5] Note, since not all respondents entered all cost information, the number of observations are different for each cost type in the detail data summary presented in Tables 7 and 8.

[6] In 2008, the survey added questions on principal payments and catch handling costs. Since these costs were not asked in 2006 and 2007 surveys, they were excluded in this analysis.

4 THE MODELING FRAMEWORK

Although economic information is key to several analyses relevant to commercial fisheries management, it is extremely challenging to obtain this information. Surveys are a useful tool to meet this gap regarding the need for economic data. However, it is often not feasible to survey the entire population of active fishing vessel owners. In rare cases, even if the whole population is surveyed, one cannot guarantee a 100% response rate. Therefore, it is extremely important to build robust statistical models which can estimate and predict costs for unsurveyed vessels with reasonable accuracy. This section discusses the modeling framework for estimating and predicting costs for fishing vessels. Separate models were estimated for *Other Annual Cost*, *RMI Cost* and *Crew Cost*. These separate models allow for more detailed inferences about fisheries cost structure and profitability. The use of separate models also provides more flexibility in addressing specific research questions.

Typically, an ordinary least square method (OLS) is used to estimate fisheries costs. However, using OLS with cost data often leads to negative cost predictions. Therefore researchers use OLS with the log of the dependent variable, and predict cost via exponentiating the predicted cost values in log scales. This retransformation, though frequently used, causes bias (Manning 1998; Jia and Rathi 2008; Manning and Mullahy 2001). The bias is worse if there is heteroscedasticity (nonconstant variance) in the log-transformed model. To correctly predict cost when using the log-transformed linear model estimation, proper adjustment should be applied with anti-log-transformation. However, this adjustment process for unbiasedness can be computationally expensive. An alternative method is to use a generalized linear model (GLM).

A GLM can be viewed as a differentially weighted nonlinear least square estimation method. The advantages in using a GLM approach are that: (a) there is no retransformation bias; (b) no adjustment is needed for anti-log transformation; and (c) GLM does not assume constant variance. GLM is also a preferred method for analyzing skewed data as often encountered with cost and expenditure data. These data are typically characterized by nonnegative measurements of the outcomes, and a positively skewed empirical distribution of nonzero realizations (Manning and Mullahy 2001; Moran et al. 2007). GLMs have been widely applied in analyzing human health expense data where skewness of the distribution is common (Knerer et al. 2005; Wu et al. 2007; Moran et al. 2007).

Fisheries cost data share the same characteristics as health expense data. The skewness of the distributions in each of the major fishery vessel cost categories is evident in Figures 5 and 6. Because of this, all three cost models were estimated via GLM. The estimation method was carried out by specifying a gamma distribution function for the error term and a log-link function for cost using the GENMOD procedure in SAS.[7] For estimation purposes, the data from all three years were combined. As a result, the estimation sample consisted of multiple observations per vessel, as some vessels were surveyed in multiple years and also responded in multiple years. Therefore, it was possible for the responses from the same vessel to be correlated. These within-vessel correlations were taken into account via the *repeated* option within GENMOD.

To identify the best predictors, several continuous and categorical variables were constructed based on the information available from the permit database, the vessel logbook database, and the dealer database. All dollar values were converted to 2006 dollars in this

[7] A gamma distribution is often a preferred distribution choice for the error term in cost analysis with GLM (Moran et al. 2007; Knerer et al. 2005)

analysis. These variables used in the model are described in Tables 11 and 12. For this part of the analysis, principal gear categories were further grouped under three major gear types: *Static*, *Mobile*, and *Other* (Table 12). The *Mobile* gear group included *Dredge* and *Trawl*, whereas the *Static* gear group included *Gillnet, Longline, Pot/Trap and Handgear*. All remaining gear categories were included in the *Other* gear group.

Although there were 1,153 vessels in the estimation set, not all characteristics information could be calculated for all these vessels. On average, the vessels included in the estimation sample were 47 ft long, 22 years old, weighed 43 gross tons, and were equipped with 394 horse power (Table 11). In addition, the average vessel had 3 crew members and had an average annual revenue of $212,000. Most of the vessels used *Static* gear (61%), were constructed of fiber glass (60%), and landed their catch primarily in the *New England* region (68%) (Table 12).[8]

The final model specification was chosen based on the quasi-likelihood values, Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). RMSE was calculated by (a) taking the square of the difference between the predicted and known costs; (b) then taking the mean of these squared differences; and (c) then taking the square root of the mean. A low RMSE indicates a better fit (Moran et al. 2007). The following sections elaborate on the three different cost models.

4.1 Repair/Maintenance/Improvement (RMI) Cost

The GLM coefficients of the *RMI Cost* model are listed in Table 13. The positive coefficient for vessel length implies that costs increase with length, but the negative coefficient for length-square indicates that *RMI Cost* eventually decreases as vessel length increases. The results indicate that newer vessels have higher repair, maintenance, and improvement costs as do vessels constructed of steel rather than other types of construction. Variable *gtons* has a negative coefficient implying lower *RMI Cost* for heavier vessels, but this effect is not significant.

4.2 Other Annual Cost

Table 14 reports the estimated GLM coefficients of the *Other Annual Cost* model. All the variables included in the final model are significant. The results indicate that vessels with a higher age have lower *Other Annual Cost*. The positive coefficient on *length* implies higher *Other Annual Cost* for larger vessels. The year coefficients for 2006 and 2007 are both negative and statistically significant implying that the *Other Annual Costs* are lower in these years than in 2008. Positive *fglass* and *gtons* coefficients imply higher costs for vessels with fiber glass construction and for vessels with higher weights.

4.3 Crew Cost

Table 15 lists the GLM coefficients from the *Crew Cost* model. *Crew Cost* increases as vessel length increases. Vessels which primarily use *Static* gear to fish are likely to have a higher *Crew Costs* than do vessels that use mobile or other gear types. *Crew Costs* were also higher for vessels with higher total revenues.

[8] Although vessel's holding capacity could have been an important predictor, it has not been considered for estimation because measures on this variable were missing for about 17% vessels in the estimation set.

4.4 Predicted Cost Summary

The estimated coefficients from the models were used to predict costs for all of the vessels that were not included in the estimation sample. Summary statistics for the continuous variables used in the prediction data set are listed in Table 16, with the frequency distribution of the categorical variables in this data set shown in Table 17. Vessels having a total revenue below $100 were excluded from the prediction dataset, as were a few other vessels which had invalid values for some variables. This resulted in 6,765 vessels in the cost prediction data set. However, because not all variables could be defined for all vessels, it was not possible to predict costs for each of the 6,765 vessels. On average, the vessels included in the prediction data set were 22 years old, 51 feet long, weighed 53 gross tons, had 455 horse power, and earned $308,510 in total annual revenue (Table 16). More than half of the vessels in the prediction data set fished with *Static* gear (54%), were constructed of fiber glass materials (58%), and landed primarily in the New England region (72%) (Table 17).

Table 18 presents the summary statistics of the predicted costs based on the prediction sample, along with the known costs from the estimation sample. To calculate the predicted Total Annual Cost, the *Other Annual Cost* and *RMI Cost* for each vessel were predicted separately based on the model coefficients in Table 14 and 13. The predicted *Crew Costs* were based on *Crew Cost* model estimates in Table 15. These cost components were then summed to derive the *Total Annual Cost* for each vessel.

The greatest discrepancy between the known costs and predicted costs occurred for *Crew Cost* (Table18), and this affected the disparity in *Total Annual Cost* as *Crew Cost* enters in the calculation of the *Total Annual Cost*. The predicted mean *Crew Cost* was $148,176, nearly double the known mean value of $76,917. The higher standard deviations of the predicted costs indicate a larger dispersion in their distributions as well. Because *Crew Cost* increases with vessel revenue, the predicted *Crew Costs* are extremely large for vessels with higher revenue values compared to vessels with lower total revenues. On the other hand, known and predicted average values for *RMI* and *Other Annual Costs* are comparable. The predicted mean values of *RMI* and *Other Annual Cost* were $37,535 and $52,283 compared to the known values of $32,337 and $41,904, respectively.

Table 19 compares known and predicted costs by cost category and vessel length category; the corresponding boxplots are depicted in Figures 7 and 10. Overall, absolute differences between predicted and known values in all cost categories become larger as vessel length increases. These differences among large size vessels are most pronounced for *Crew Cost* and hence *Total Annual Cost*. As noted earlier, these differences are caused by *Crew Cost,* which increases with total revenues and total revenues are higher for the larger vessels resulting in very high predicted *Crew Cost* for the large vessels. The high standard deviations for *Large* vessels also indicate high disparity in predicted *Crew Costs* and *Total Annual Costs* for large vessels. The boxplots for these two cost categories (Figures 9 and 10) also show more potential outliers for *Large* vessels. On the other hand, for the *Medium* vessels, *RMI* and *Other Annual Costs* are slightly more dispersed than the *Large* and *Small* vessels as can be seen in the boxplots as well (Figures 7 and 8).

The predicted cost summaries show that the model coefficients are successful in generating predictions for *RMI* and *Other Annual Costs* that are generally consistent with known values. However, model based predictions for *Crew Costs* yield average values which are highly divergent from the known mean values, and the distribution of the predicted values is highly dispersed as well. *Crew Costs* are expected to vary largely among vessel owners depending on

the crew payment system they use. Better understanding and data on these crew payment systems should potentially improve the *Crew Cost* models. However, in the absence of such additional information, the predicted *Crew Costs* should be carefully interpreted.

5 CONCLUSIONS

This document explains the cost data collection effort that the Social Sciences Branch of the Northeast Science Center had undertaken from years 2006 - 2008. A detailed description of the survey methodology, the data, and the data auditing process is presented here. Several summary statistics of the revised data are also presented. In order to make cost predictions for vessels for which this information is not available, several modeling approaches are discussed. Three different cost models estimated are, *Repair/Maintenance/Improvement (RMI) Cost*, *Other Annual Cost,* and *Crew Cost*. Summary of predicted costs based on model-coefficients in comparison to known costs are also given.

This analysis will facilitate the application of cost data within SSB and among collaborating partners. The cost models will allow analysts to predict vessel level cost estimates rather than using average values. These models will enhance analyses that require cost estimates, such as break-even analyses, profitability profiles and economic performance indicator calculations.

The NEFSC Social Sciences Branch is currently engaged in relaunching the vessel annual cost data collection effort. The new effort will build on the findings and lessons from the 2007-2009 study to improve survey coverage and data quality and to enhance future analyses and evaluations of the economic status of commercial fisheries in the Northeast.

ACKNOWLEDGEMENTS

The author thanks numerous individuals who contributed to this data collection effort. The author especially thanks Andrew Kitts for his efforts in the survey design and implementation of this project. The assistance of the Northeast Regional Permit office is duly appreciated for administering the survey. Thanks are also due to thousands of commercial vessel owners who responded to the survey.

REFERENCES CITED

Dillman DA, Smyth JD, Christian LM. 2009. Internet, mail and mixed-mode surveys: The tailored design method. 3rd ed. John Wiley & Sons, Inc. 512p.

Jia S, Rathi S. 2008, On predicting log-transformed linear models with heteroscedasticity. SAS Publication.

Knerer G, Byford S, Johnson T, Seivewright H, Tyrer P. 2005. The Nottingham study of neurotic disorder: Predictors of 12 year costs. Acta Psychiatrica Scandinavica 112(3):224–232.

Lian CE. 2010. West Coast limited entry groundfish trawl cost earning survey: Protocols and results for 2004. U.S. Department of Commerce., NOAA Technical Memorandum, NMFS-NWFSC-107, 35p.

Lohr SL. 2010. Sampling: design and analysis, 2nd Ed, Cengage Learning, Wadsworth Publishing, 608p.

Manning WG. 1998. The logged dependent variable, heteroscedasticity, and the retransformation problem. Journal of Health Economics 17(3):283–295.

Manning WG, Mullahy J. 2001. Estimating log models: To transform or not to transform? Journal of Health Economics 20(4):461–494.

Moran JL, Solomon PJ, Peisach AR, Martin J. 2007. New models for old questions: Generalized linear models for cost prediction. Journal of Evaluation in Clinical Practice 13(3):381–389.

Wu E Q, Birnbaum HG, Zhang HF, Ivanova JI, Yang E, Mallet D. 2007, Health care costs of adults treated for attention-deficit/hyperactivity disorder who received alternative drug therapies. Journal of Managed Care Pharmacy 13(7):561-569.

Table 1. Survey response by vessel type, revenue, and landing states.

	Survey Year=2006			Survey Year=200 7			Survey Year=2008		
	Survey Popula-tion	Complete Re-sponses	Response Rate	Survey Population	Complete Responses	Response Rate	Survey Population	Complete Responses	Response Rate
Gear Category									
Dredge	692	110	15.92	596	71	11.91	568	45	7.92
Gillnet	256	63	24.61	251	47	18.73	236	23	9.75
Handgear	343	88	25.66	248	46	18.55	259	22	8.49
Longline	56	14	25.00	37	4	10.81	31	1	3.23
Pot/Trap	756	192	25.40	863	160	18.54	1276	81	6.35
Trawl	615	136	22.11	529	87	16.45	458	57	12.45
others	92	9	9.78	60	11	18.33	46	3	6.52
Unclassified	248	18	7.26	18	4	22.22	8	0	0.00
Region									
Mid-Atlantic	842	192	22.80	762	142	18.64	689	70	10.16
New-England	1959	418	21.34	1822	284	15.59	2174	159	7.31
Unclassified	257	20	7.78	18	4	22.22	19	3	15.79
Length Category									
Large (>80 ft)	296	38	12.84	285	27	9.47	285	27	9.47
Medium (40-80 ft)	1452	334	23.00	1304	225	17.25	1325	120	9.06
Small (<40 ft)	1057	256	24.22	978	176	18.00	1272	85	6.68
Unclassified	253	2	0.79	35	2	5.71	0	0	0.00
Revenue Category									
<25k	791	171	21.62	635	104	16.38	620	45	7.26
25k to 100k	596	138	23.15	616	94	15.26	712	44	6.18
100k to 500k	920	210	22.83	861	165	19.16	1034	88	8.51
500k to 1 million	252	59	23.41	216	29	13.43	276	30	10.87
over 1 million	251	34	13.55	256	34	13.28	232	25	10.78
Unclassified	248	18	7.26	18	4	22.22	8	0	0.00
Total	*3058*	*630*		*2602*	*430*		*2882*	*232*	

Table 2. Frequency of multiple survey receipts and responses over three years.

	Survey Year=2006	Survey Year=2007	Survey Year= 2006 & 2007	Survey Year=2008	All Years
Survey Year 2007	2150	447		-	2597
Survey Year 2008	150	323	1791	615	2879

Table 3. Order in which surveys received and returned over the three years.

#Received/ #Responded	Not-Returned	Returned once	Returned Twice	Returned Thrice	Total
Sent Once	1,384	107	-	-	1,491
	(92.82)	(7.18)	-	-	(36.24)
Sent Twice	640	147	45	-	832
	(76.92)	(17.68)	(5.41)	-	(20.22)
Sent Thrice	1,185	336	205	65	1,791
	(66.14)	(18.76)	(11.45)	(3.63)	(43.53)

Table 4. Nonresponse bias test.

	Survey Year=2006			Survey Year=2007			Survey Year=2008		
Length	N	Mean	Std Dev	N	Mean	Std Dev	N	Mean	Std Dev
Population	2801	50.62	20.90	2563	50.80	20.68	2878	48.53	19.98
Nonrespondent	2173	51.41	21.49	2135	51.34	20.96	2646	48.22	19.92
Respondent	628	47.92	18.49	428	48.07	18.99	232	52.12	20.33
T statistics	4.01				3.20			2.80	
DF	1161				653.27			271.35	
$Pr > \|t\|$	<.0001				0.0014			0.0054	
Gtons	N	Mean	Std Dev	N	Mean	Std Dev	N	Mean	Std Dev
Population	2801	53.14	58.47	2563	53.76	59.19	2878	47.76	56.06
Nonrespondent	2173	55.20	59.65	2135	55.49	60.62	2646	46.726	55.59
Respondent	628	46.01	53.60	428	45.13	50.63	232	59.58	60.10
T statistics	3.69				3.73			3.14	
DF	1115				6696.24			266.82	
$Pr > \|t\|$	0.0002				0.0002			0.0019	
Vhp	N	Mean	Std Dev	N	Mean	Std Dev	N	Mean	Std Dev
Population	2801	438.50	294.50	2563	445.40	302.20	2878	442.40	291.10
Nonrespondent	2173	447.80	307.60	2135	453.60	311.90	2646	440.70	291.50
Respondent	628	406.20	241.10	428	404.50	244.00	232	461.70	286.30
T statistics	3.57				3.62			1.07	
DF	1274.50				736.83			274.68	
$Pr > \|t\|$	0.0004				0.0003			0.2575	
Total Revenue	N	Mean	Std Dev	N	Mean	Std Dev	N	Mean	Std Dev
Population	2810	275,280	432,920	2584	279,970	405,736	2874	273,289	409,184
Nonrespondent	2198	286,078	447,451	2158	285,836	410,901	2642	267,493	405,334
Respondent	612	236,499	373,899	426	250,255	377,547	232	339,295	446,372
T statistics	2.77				1.65			2.37	
DF	1144.8				640.09			265.54	
$Pr > \|t\|$	0.0056				0.0804			0.0187	

Table 5. Summary statistics of weighted annual costs by major cost categories and of total revenue by survey years.

Major Cost Categories	Survey Year=2006				Survey Year=2007				Survey Year=2008			
	N	Mean	Std Dev	Max.	N	Mean	Std Dev	Max.	N	Mean	Std Dev	Max.
RMI	524	34,597	86,918	260,500	372	32,199	106,138	271,163	195	33,202	173,702	395,676
Other Annual Cost	566	39,926	94,144	243,260	382	41,276	108,746	271,950	201	52,224	227,869	334,840
Crew Cost	564	81,423	281,516	686,023	375	84,451	364,255	771,750	196	83,399	499,953	792,209
Total Annual Cost	569	152,187	407,139	923,393	383	155,221	510,886	996,675	201	165,938	768,836	1,125,906
Total Revenue	569	216,112	667,630	1,701,826	383	225,442	804,812	1,683,660	201	279,592	1,477,283	2,820,277

Table 6. Summary statistics of weighted annual costs by major cost categories and of total revenue by survey years and length categories.

Major Cost Categories		Survey Year=2006				Survey Year=2007				Survey Year=2008			
	Length	N	Mean	Std Dev	Max.	N	Mean	Std Dev	Max.	N	Mean	Std Dev	Max.
RMI	L	25	87,927	160,193	260,500	14	93,143	202,164	271,163	17	71,701	182,881	214,178
	M	280	43,585	87,215	203,000	191	42,542	116,302	270,197	102	41,998	187,246	395,676
	S	218	14,372	40,252	138,200	167	11,704	32,943	73,500	76	17,433	124,126	250,860
Other Annual Cost	L	27	127,026	127,051	243,260	15	120,615	215,040	271,950	18	207,013	235,708	334,840
	M	302	48,759	92,555	217,173	197	52,246	107,389	215,411	105	59,656	178,592	265,290
	S	236	15,077	29,098	96,500	170	17,107	34,409	81,275	78	17,569	46,167	62,382
Crew Cost	L	26	340,250	484,133	686,023	15	396,881	711,467	771,750	17	420,575	704,095	792,209
	M	303	104,180	280,197	684,482	194	106,270	369,897	630,096	102	94,008	388,375	522,810
	S	234	12,809	41,443	105,000	166	16,815	54,778	157,141	77	17,371	76,770	85,803
Total Annual Cost	L	27	537,691	588,636	923,393	15	603,617	876,749	996,675	18	666,672	943,630	1,125,906
	M	304	192,128	396,204	902,034	197	198,205	511,296	947,075	105	191,901	597,514	807,452
	S	237	40,919	83,899	261,000	171	44,779	94,994	250,491	78	51,761	196,021	359,141
Total Revenue	L	27	852,005	1,077,008	1,701,826	15	976,209	1,366,064	1,683,660	18	1,253,975	1,523,854	2,055,246
	M	304	270,585	652,507	1,574,525	197	281,645	795,049	1,425,437	105	314,678	1,250,009	2,820,277
	S	237	48,705	129,396	418,717	171	61,551	178,940	440,767	78	74,046	264,966	292,222

L=Large; M=Medium; S= Small

Table 7. Summary statistics of weighted individual cost items by survey years and length categories.

Individual Cost Items		Survey Year=2006				Survey Year=2007				Survey Year=2008			
		N	Mean	Std Dev	Max.	N	Mean	Std Dev	Max.	N	Mean	Std Dev	Max.
Repair/Maintenance/Improvement Costs													
Total Improvement Cost	L	19	49,041	98,061	140,000	11	43,087	81,251	92,984	8	38,607	88,487	74,648
	M	210	32,093	70,999	161,190	139	30,781	87,368	157,500	59	33,661	151,646	255,676
	S	156	12,876	30,393	85,500	107	10,410	26,729	72,450	47	19,010	151,558	250,860
Repair Maintenance	L	25	51,659	106,802	174,386	14	59,627	139,516	178,179	17	55,758	117,459	167,558
	M	273	20,404	53,935	143,000	186	20,692	62,609	139,696	101	23,288	97,389	153,835
	S	210	5,441	22,748	125,000	161	5,189	19,084	63,000	73	6,255	23,364	29,822
Other Annual Costs													
Association Fees	L	25	1,472	5,319	9,300	12	1,185	4,429	5,926	18	2,840	6,448	5,550
	M	280	764	3,621	15,000	180	895	4,832	21,000	101	1,238	6,127	8,371
	S	212	270	1,716	10,000	159	283	2,454	12,285	74	188	1,401	2,220
Communication Cost	L	26	4,313	7,479	15,000	13	4,893	11,486	16,636	18	3,544	7,351	9,759
	M	289	2,216	4,232	17,000	192	2,162	3,403	8,190	101	2,081	6,813	17,094
	S	221	737	1,584	4,000	164	924	2,903	11,865	73	967	3,247	3,996
Haul Cost	L	22	11,638	38,047	65,000	13	10,423	42,845	52,500	16	22,621	87,017	91,214
	M	258	6,380	25,694	100,000	179	6,608	31,153	73,500	90	7,500	41,287	65,490
	S	209	1,469	6,261	40,000	155	1,565	6,649	26,250	73	1,498	8,703	13,320
Insurance	L	26	55,665	47,272	90,000	14	51,422	70,560	88,290	18	68,406	74,110	104,051
	M	289	14,071	36,812	88,000	190	16,389	51,826	86,121	102	14,445	63,007	94,320
	S	224	2,093	4,226	15,000	165	2,238	5,143	13,860	77	2,055	5,314	5,550
Interest	L	25	20,900	62,481	98,996	14	14,007	49,666	52,500	17	55,555	180,254	165,924
	M	269	5,130	19,578	64,863	181	6,553	25,321	46,246	98	8,621	62,621	105,196
	S	210	1,377	6,977	30,000	155	1,234	5,930	15,750	70	1,771	13,620	17,760
Labor Services	L	25	6,247	33,323	57,000	13	2,888	21,813	32,025	16	4,606	23,028	22,200
	M	263	2,057	14,104	40,000	176	512	4,166	10,500	94	1,608	20,073	33,300
	S	202	622	6,056	30,000	158	405	4,474	15,750	74	321	8,432	22,200

L=Large; M=Medium; S= Small

Table 8. Summary statistics of weighted individual cost items by survey years and length categories, continued from Table 7.

Individual Cost Items		Survey Year=2006				Survey Year=2007				Survey Year=2008			
		N	Mean	Std Dev	Max.	N	Mean	Std Dev	Max.	N	Mean	Std Dev	Max.
Repair/Maintenance/Improvement Costs													
Total Improvement Cost	L	19	49,041	98,061	140,000	11	43,087	81,251	92,984	8	38,607	88,487	74,648
	M	210	32,093	70,999	161,190	139	30,781	87,368	157,500	59	33,661	151,646	255,676
	S	156	12,876	30,393	85,500	107	10,410	26,729	72,450	47	19,010	151,558	250,860
Repair Maintenance	L	25	51,659	106,802	174,386	14	59,627	139,516	178,179	17	55,758	117,459	167,558
	M	273	20,404	53,935	143,000	186	20,692	62,609	139,696	101	23,288	97,389	153,835
	S	210	5,441	22,748	125,000	161	5,189	19,084	63,000	73	6,255	23,364	29,822
Other Annual Costs													
Association Fees	L	25	1,472	5,319	9,300	12	1,185	4,429	5,926	18	2,840	6,448	5,550
	M	280	764	3,621	15,000	180	895	4,832	21,000	101	1,238	6,127	8,371
	S	212	270	1,716	10,000	159	283	2,454	12,285	74	188	1,401	2,220
Communication Cost	L	26	4,313	7,479	15,000	13	4,893	11,486	16,636	18	3,544	7,351	9,759
	M	289	2,216	4,232	17,000	192	2,162	3,403	8,190	101	2,081	6,813	17,094
	S	221	737	1,584	4,000	164	924	2,903	11,865	73	967	3,247	3,996
Haul Cost	L	22	11,638	38,047	65,000	13	10,423	42,845	52,500	16	22,621	87,017	91,214
	M	258	6,380	25,694	100,000	179	6,608	31,153	73,500	90	7,500	41,287	65,490
	S	209	1,469	6,261	40,000	155	1,565	6,649	26,250	73	1,498	8,703	13,320
Insurance	L	26	55,665	47,272	90,000	14	51,422	70,560	88,290	18	68,406	74,110	104,051
	M	289	14,071	36,812	88,000	190	16,389	51,826	86,121	102	14,445	63,007	94,320
	S	224	2,093	4,226	15,000	165	2,238	5,143	13,860	77	2,055	5,314	5,550
Interest	L	25	20,900	62,481	98,996	14	14,007	49,666	52,500	17	55,555	180,254	165,924
	M	269	5,130	19,578	64,863	181	6,553	25,321	46,246	98	8,621	62,621	105,196
	S	210	1,377	6,977	30,000	155	1,234	5,930	15,750	70	1,771	13,620	17,760
Labor Services	L	25	6,247	33,323	57,000	13	2,888	21,813	32,025	16	4,606	23,028	22,200
	M	263	2,057	14,104	40,000	176	512	4,166	10,500	94	1,608	20,073	33,300
	S	202	622	6,056	30,000	158	405	4,474	15,750	74	321	8,432	22,200

L=Large; M=Medium; S= Small

Table 9. Frequency and percentages of responses on lay system and captain status.

	Survey Year=2007		Survey Year=2008	
	N	Percentage	N	Percentage
Lay System				
Clear Lay	79	20.95	50	23.70
Broken Lay	197	52.25	112	53.08
Per-trip or hourly	43	11.41	19	9.00
Others	58	15.38	30	14.22
Total	377	14.49	211	7.32
Captain Status				
Hired Captain	74	18.64	60	28.71
Owner Operated	323	81.36	149	71.29
Total	397	15.26	209	7.25

Table 10. Average boat and crew share by captain status.

	N	Mean Boat Share	Mean Crew Share
Survey Year=2007			
Hired Captain	70	48.67	51.33
Owner Operated	176	52.23	47.77
Aggregate	258	51.26	48.74
Survey Year=2008			
Hired Captain	56	49.30	50.70
Owner Operated	80	60.20	39.80
Aggregate	144	55.53	44.47

Table 11. Summary statistics of the continuous independent variables.

Variable	Variable Definition	N	Missing	Mean	Std Dev	Minimum	Maximum
Age	Boat's age	1153	1	22	13	0	82
Crew	Crew Size	1153	1	3	2	1	10
Gtons	Gross Ton	1153	1	43	49	1	201
Length	Vessel's length in ft	1153	1	47	18	18	117
Length-sq	Vessel's length squared	1153	1	2524	2020	324	13689
Vhp	Vessel Horse Power	1153	1	394	207	70	1810
Vhplen	Vessel Horse Power/length	1153	1	8	3	3	25
Totrev_ths	Total Revenue in $1000	1154	0	212	312	0	2541

Table 12. Frequency distribution of the independent categorical variables.

Variable	N	Percentage
Geargroup		
Mobile	426	36.98
Static	706	61.20
Othgr (Other gear types)	21	1.82
Types of Construction		
fglass (Fiber Glass)	680	59.86
Steel	279	24.56
Wood	177	15.58
Principal Landing Region		
Reg_MA (Mid-Atlantic)	363	31.65
Reg_NE (New England)	784	68.35

Table 13. Generalized Linear Model Estimates of the RMI cost model.

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	$Pr > \|Z\|$
Intercept	7.1854	0.338	6.5229	7.8478	21.26	<.0001
age	-0.014	0.0036	-0.021	-0.0069	-3.89	0.0001
length	0.0979	0.0133	0.0718	0.124	7.35	<.0001
length-sq	-0.0005	0.0001	-0.0008	-0.0003	-3.86	0.0001
reg_MA	0.1131	0.0802	-0.044	0.2702	1.41	0.1583
steel	0.2395	0.1385	-0.032	0.511	1.73	0.0839
gtons	-0.0025	0.0023	-0.007	0.002	-1.08	0.2781
No. of observations: 1038; Mean Absolute Error=22260; Root Mean Square Error= 36390						

Table 14. Generalized Linear Model Estimates of the other annual cost model

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	$Pr > \|Z\|$
Intercept	7.0625	0.3052	6.4643	7.6607	23.14	<.0001
age	-0.0073	0.0025	-0.0121	-0.0025	-2.96	0.0031
fglass	0.1806	0.0843	0.0155	0.3458	2.14	0.032
length	0.1027	0.0102	0.0828	0.1226	10.11	<.0001
length-sq	-0.0006	0.0001	-0.0008	-0.0004	-6.07	<.0001
reg_MA	-0.1481	0.0532	-0.2524	-0.0437	-2.78	0.0054
year07	-0.1473	0.0543	-0.2538	-0.0408	-2.71	0.0067
year06	-0.2179	0.0512	-0.3182	-0.1175	-4.26	<.0001
gtons	0.0031	0.0015	0.0001	0.006	2.05	0.04
No. of observations: 1125; Mean Absolute Error=19941; Root Mean Square Error= 31118						

Table 15. Generalized Linear Model Estimates of the crew cost model.

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	$Pr > \|Z\|$
Intercept	7.4567	0.3723	6.727	8.1864	20.03	< .0001
age	-0.0085	0.0027	-0.0138	-0.0032	-3.15	0.0016
length	0.0869	0.0133	0.0608	0.113	6.52	< .0001
length-sq	-0.0005	0.0001	-0.0007	-0.0002	-3.7	0.0002
gtons	-0.0032	0.0019	-0.0069	0.0004	-1.73	0.0842
static	0.114	0.0707	-0.0245	0.2525	1.61	0.1068
year06	0.0022	0.0407	-0.0776	0.082	0.05	0.9568
crew	0.1014	0.0277	0.0472	0.1556	3.66	0.0002
totrev_ths	0.0017	0.0001	0.0014	0.0019	12.46	< .0001
No. of observations: 856; Mean Absolute Error=50029; Root Mean Square Error= 128345						

Table 16. Summary statistics of the continuous variables in the prediction dataset.

Variable	Variable Definition	N	Miss	Mean	Std Dev	Minimum	Maximum
Age	Boat's age	6765	0	22	12	0	89
Crew	Crew Size	6759	6	3	2	0	15
Gtons	Gross Tons	6765	0	53	59	1	496
Length	Vessel's length in ft	6765	0	51	21	19	158
Length-sq	Vessel's length squared	6765	0	3009	2632	361	25122
Vhp	Vessel Horse Power	6765	0	455	302	70	5020
Vhplen	Vessel Horse Power/length	6765	0	9	4	1	62
Totrev_ths	Total Revenue in $1000	6604	161	308	451	0	5880

Table 17. Frequency distribution of the categorical variables in the prediction dataset.

Variable	N	Percentage
Geargroup		
Mobile	2882	43.64
Static	3584	54.27
Othgr (Other gear types)	138	2.09
Types of Construction		
Fiber Glass	3870	57.75
Steel	2055	30.67
Wood	776	11.58
Principal Landing Region		
Reg_MA (Mid-Atlantic)	1817	27.55
Reg_NE (New England)	4779	72.45

Table 18. Summary statistics of the known and predicted costs.

Major Cost Categories	N	Mean	Std Dev	95% Confidence Interval	
RMI Cost					
Known costs	1092	32,337	41,928	29,847	34,826
Predicted costs	6508	37,535	27,240	36,873	38,197
Other Annual Cost					
Known costs	1150	41,904	48,756	39,083	44,725
Predicted costs	6508	52,283	44,648	51,198	53,368
Crew Cost					
Known costs	1136	76,917	129,463	69,380	84,453
Predicted costs	6573	148,176	286,044	141,259	155,092
Total Annual Cost					
Known costs	1154	148,075	191,539	137,013	159,138
Predicted costs	6579	236,889	337,740	228,726	245,052

Table 19. Summary statistics of the known and predicted costs by length categories.

Major Cost Categories		N	Mean	Std Dev	95% Confidence Interval	
RMI Cost						
Large	Known	56	86,675	66,495	68,868	104,483
Large	Predicted	737	81,846	19,606	80,428	83,264
Medium	Known	573	41,635	43,980	38,026	45,244
Medium	Predicted	3,275	44,237	22,834	43,455	45,019
Small	Known	462	14,190	20,891	12,280	16,100
Small	Predicted	2496	15,658	4,465	15,482	15,833
Other Annual Cost						
Large	Known	60	150,252	76,320	130,536	169,968
Large	Predicted	737	139,792	31,962	137,481	142,103
Medium	Known	604	51,675	44,807	48,094	55,256
Medium	Predicted	3,275	57,567	33,671	56,413	58,720
Small	Known	485	16,334	14,226	15,065	17,604
Small	Predicted	2496	19,510	6,068	19,272	19,749
Crew Cost						
Large	Known	58	358,779	219,010	301,194	416,365
Large	Predicted	743	631,990	500,133	595,969	668,010
Medium	Known	599	98,882	125,027	88,850	108,915
Medium	Predicted	3,313	132,576	205,271	125,583	139,568
Small	Known	478	15,142	21,484	13,211	17,073
Small	Predicted	2,517	25,891	9,181	25,532	26,250
Total Annual Cost						
Large	Known	60	577,969	277,548	506,271	649,667
Large	Predicted	743	851,837	512,857	814,901	888,774
Medium	Known	606	188,612	179,776	174,270	202,954
Medium	Predicted	3,316	233,001	242,563	224,742	241,260
Small	Known	487	44,591	43,060	40,757	48,425
Small	Predicted	2,520	60,693	18,655	59,964	61,422

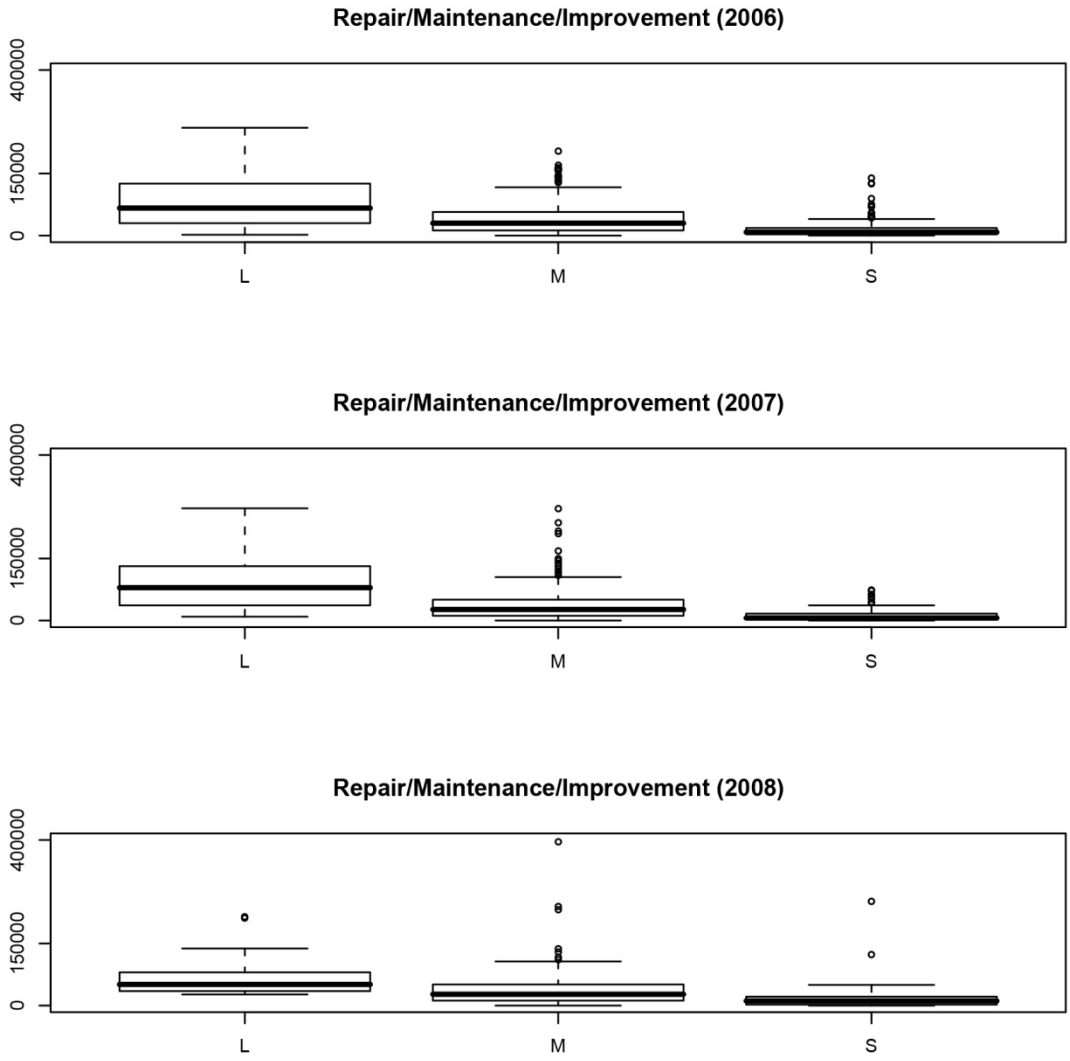

Figure 1. Distributions of RMI costs by year and length categories.

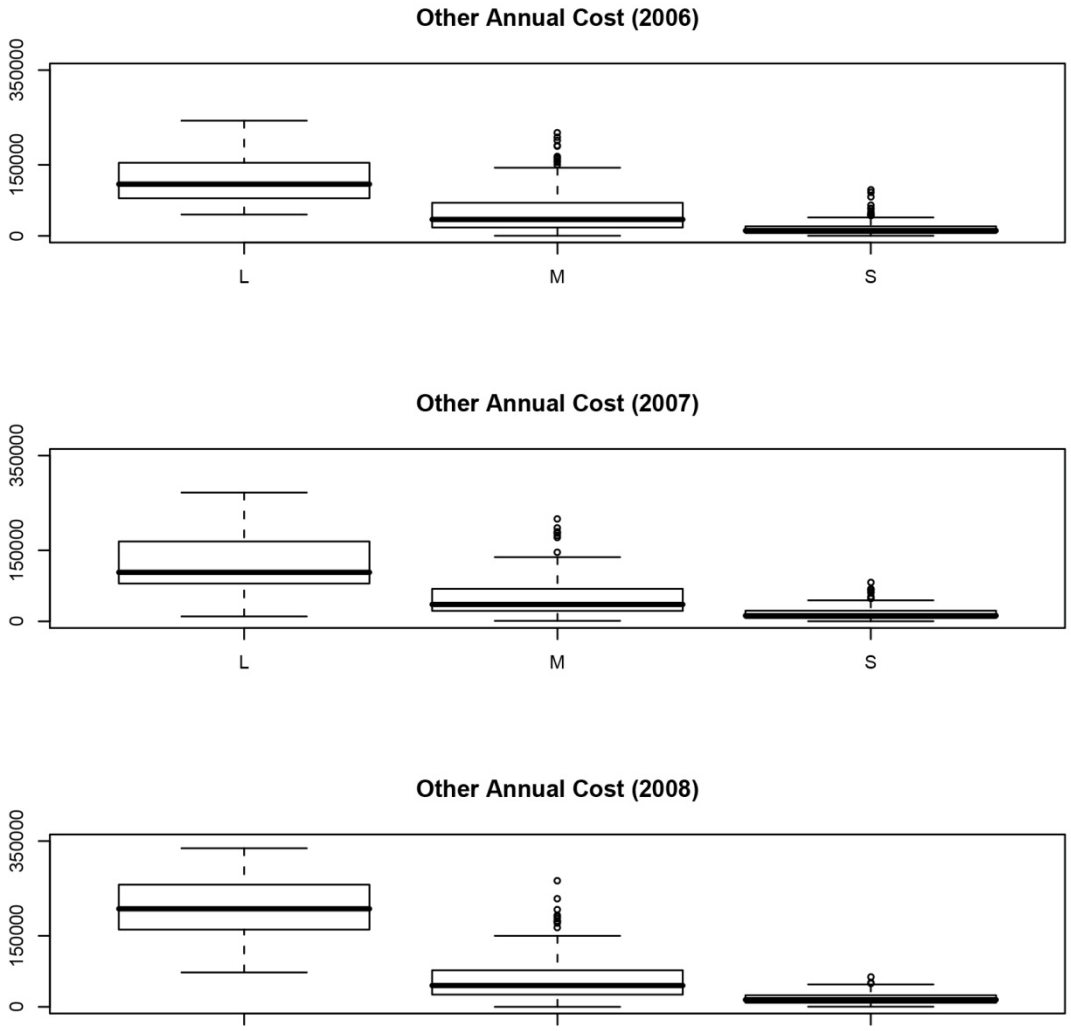

Figure 2. Distributions of other annual costs by year and length categories.

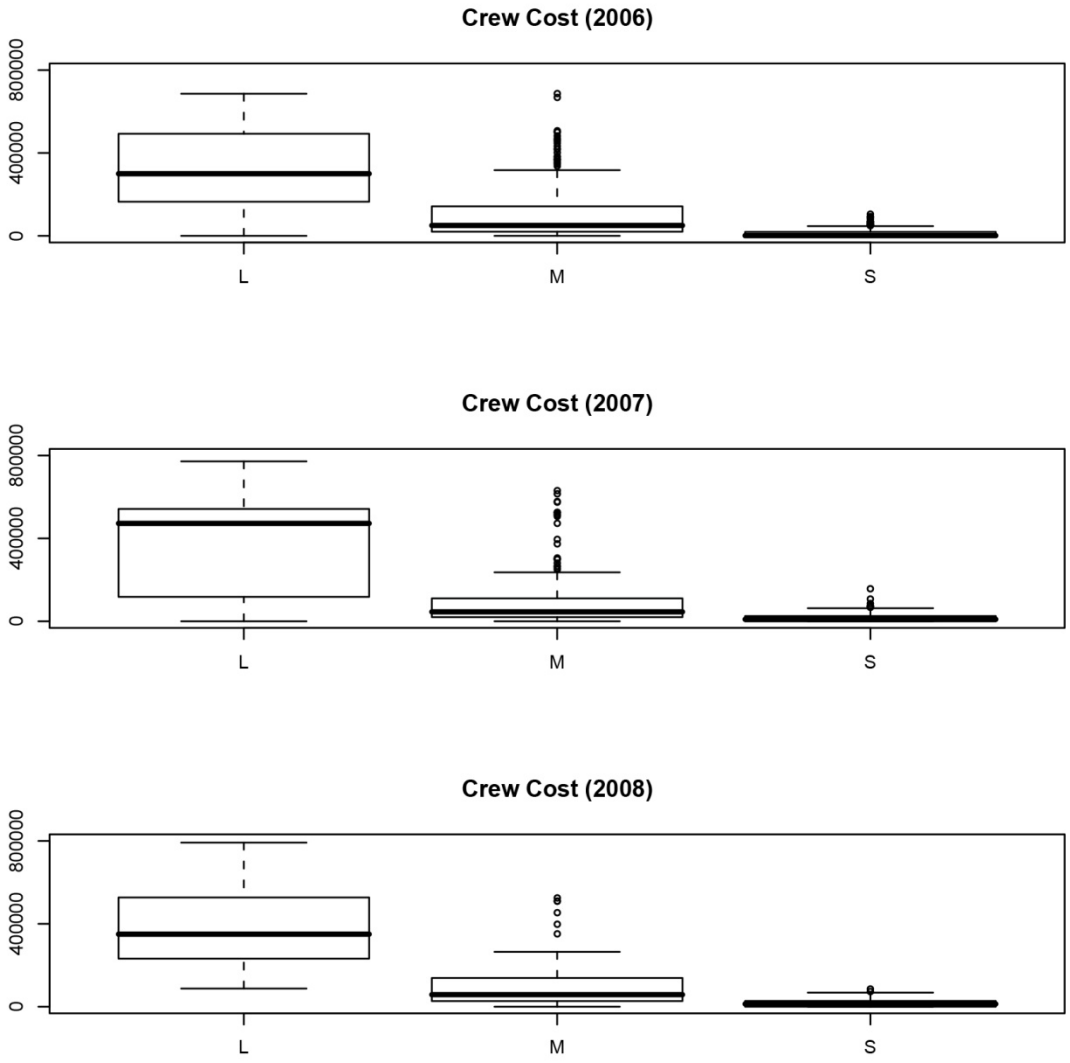

Figure 3. Distributions of crew costs by year and length categories.

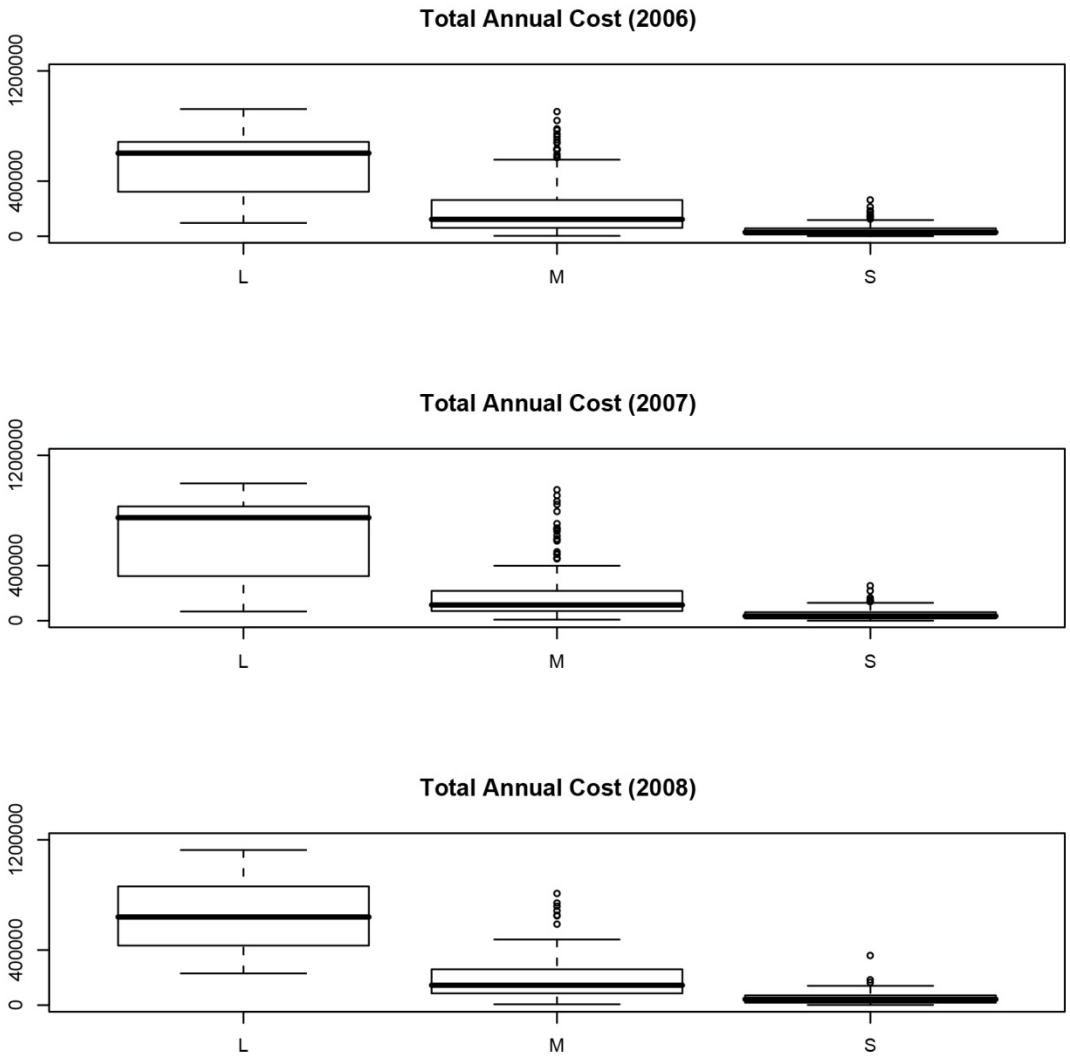

Figure 4. Distributions of total annual costs by year and length categories.

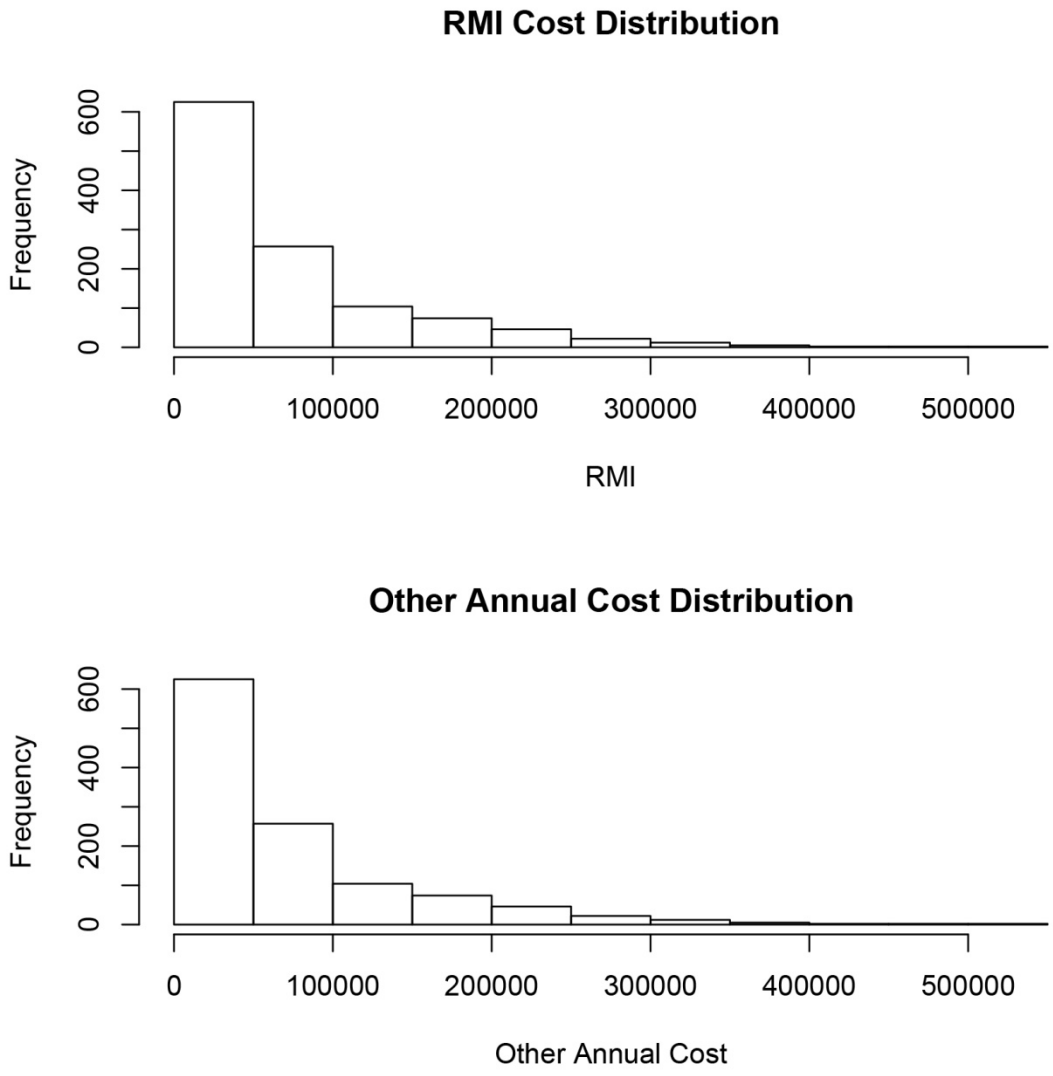

Figure 5. Histograms of RMI costs and other annual costs for three years.

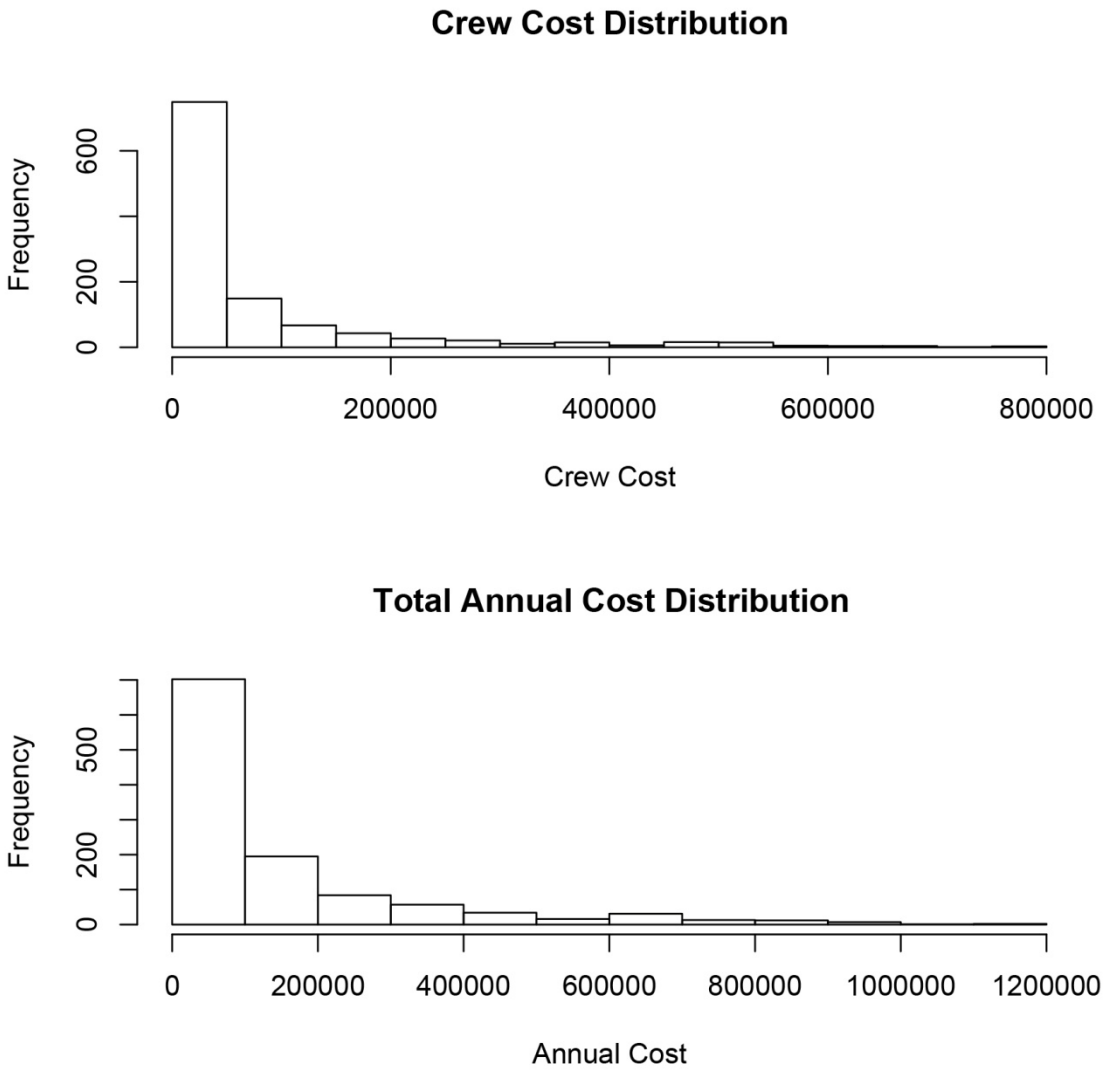

Figure 6. Histograms of crew costs and total annual costs for three years.

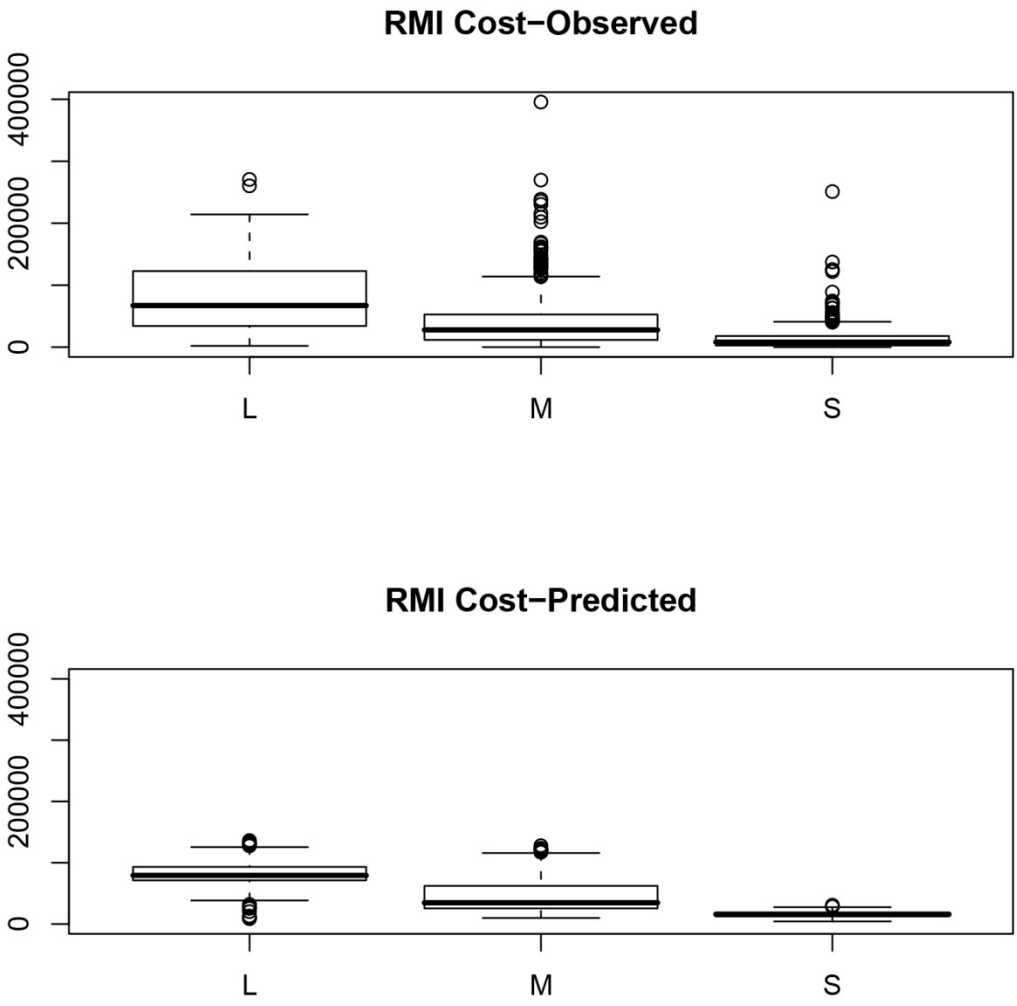

Figure 7. Distributions of known vs predicted RMI costs by length categories.

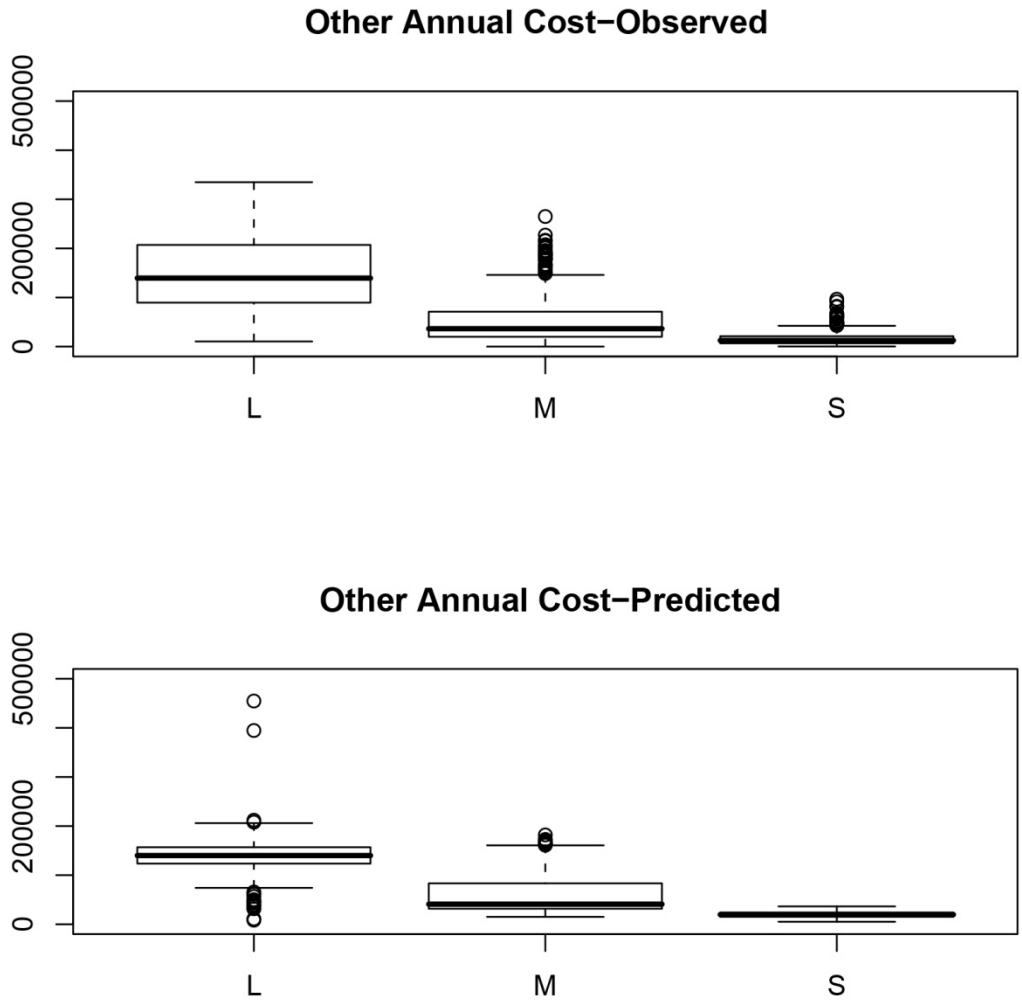

Figure 8. Distributions of know vs. predicted other annual costs by length categories.

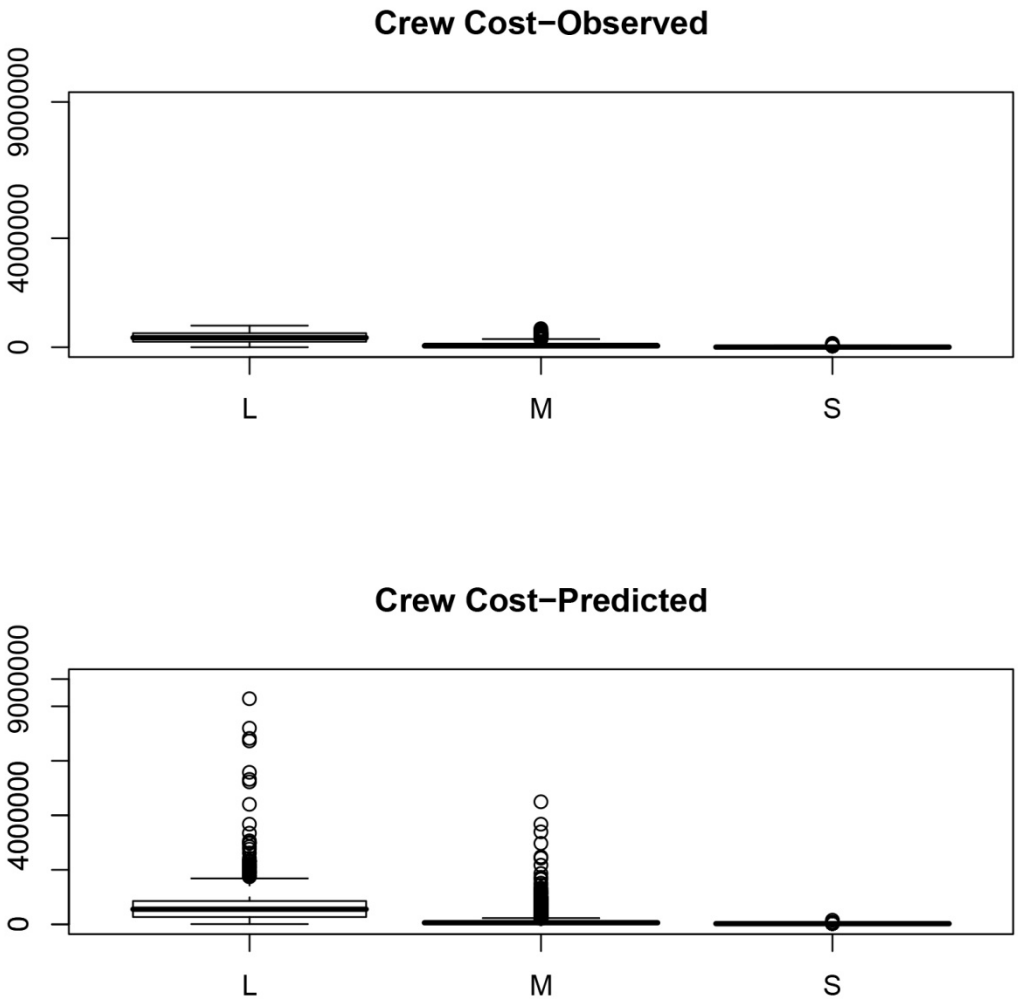

Figure 9. Distributions of known vs. predicted crew costs by length categories.

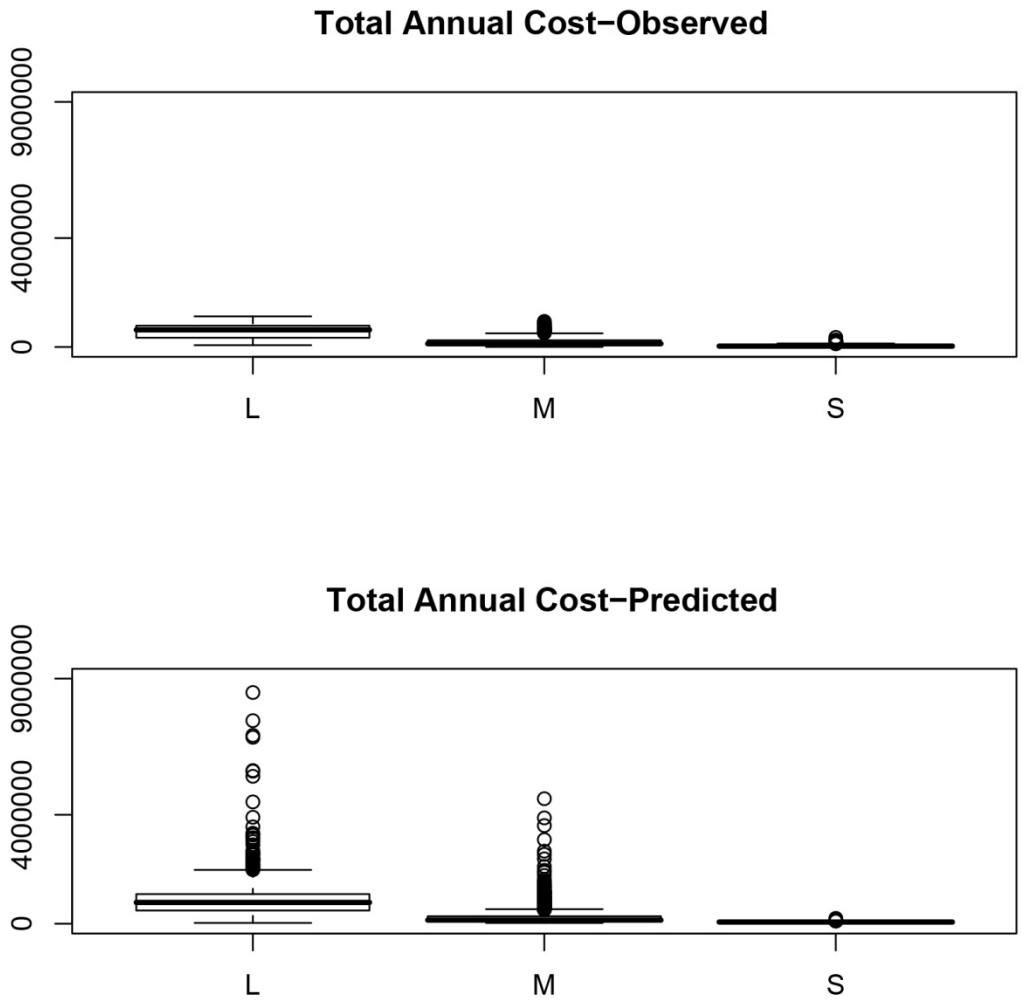

Figure 10. Distributions of known vs. predicted total annual costs by length categories.

APPENDIX

OMB Control No. 0648-0369
Expires: 07/31/2009

Northeast Fishing Vessel Annual Cost Survey

United State Department of Commerce
National Oceanic & Atmospheric Administration
National Marine Fisheries Service
Permit Office
55 Great Republic Drive
Gloucester, MA 01930-2276
Tel: (978) 281-9370

Instructions: Please record the annual costs associated with only the vessel identified below. If you own more than one vessel, certain costs may need to be divided among vessels (for example, divide office expenses by the number of vessels owned). Record the combined annual cost for all fisheries you may have participated in this fiscal year. This survey does not have questions about trip costs such as fuel, ice, bait, and supplies. This type of information is collected by observers at sea. **IMPORTANT:** *if you do not know the cost of a particular item (but an expense was incurred), please leave the question blank. If this expense does not apply to your vessel, please check the "not applicable" box.*

Please return completed surveys to the Permit Office

SECTION A - Vessel Information

Coast Guard Documentation or State Registration Number: 12345678 (one survey per vessel)

Fiscal year that corresponds with the annual costs you will provide below (use the most recent year for which you have complete records). Please provide information for *one* year only. Format: (mm/dd/yyyy)

Start date: |__|__| / |__|__| / |__|__|__|__| End date: |__|__| / |__|__| / |__|__|__|__|

Vessel Ownership Type (*check one*):

A. Sole proprietorship ☐
B. General partnership ☐
C. Limited partnership ☐
D. Corporation ☐
E. Other ☐ |__|__|__|__|__|

If you checked "D" (Corporation), please check which type:

C corporation ☐
S corporation ☐
Limited Liability Corporation ☐

Please list the number of owners: |__|__|

Was the vessel purchased from a previous owner or was it bought new? Previous owner ☐ New ☐

In what calendar year did you acquire the vessel? |__|__|__|__|

Please estimate the market value of your vessel (including all equipment, fishing gear, permits, and fishing history).

$ |__|__| , |__|__|__| , |__|__|__|

NMFS Permit Number: 123456 Hull Number: 12345678

SECTION B -
Improvements, Quota Transfer/Lease, Repair/Maintenance, Crew Compensation

What improvements (new or replacement gear, equipment, electronics, etc.) were made to the vessel this fiscal year? Please use the table to list the improvements. Also in this table, **please include the cost of buying PERMANENT quota** (surf clam/ocean quahog ITQ shares, for example). If you leased quota or days-at-sea, please provide those costs in the next question.

Description of improvement or quota transfer	Cost of improvement or quota transfer
|___|___|___|___|___|___|___|___|___|___|___|___|___|	$ |___| , |___|___|___| , |___|___|___|
|___|___|___|___|___|___|___|___|___|___|___|___|___|	$ |___| , |___|___|___| , |___|___|___|
|___|___|___|___|___|___|___|___|___|___|___|___|___|	$ |___| , |___|___|___| , |___|___|___|
|___|___|___|___|___|___|___|___|___|___|___|___|___|	$ |___| , |___|___|___| , |___|___|___|
|___|___|___|___|___|___|___|___|___|___|___|___|___|	$ |___| , |___|___|___| , |___|___|___|
|___|___|___|___|___|___|___|___|___|___|___|___|___|	$ |___| , |___|___|___| , |___|___|___|

What was the cost of **LEASING** quota or days-at-sea for use this fiscal year (include one-year leasing of surf clam/ocean quahog ITQ, days-at-sea in the multispecies fishery, or sector quota)?

$ |___|___|___| , |___|___|___| ☐ *not applicable – did not lease quota or days-at-sea this fiscal year*

Was the vessel hauled out this fiscal year? Yes ☐ No ☐

If yes, what was the cost of the haul-out (not including the cost of vessel improvements listed above)?

$ |___|___|___| , |___|___|___| ☐ *not applicable – vessel was not hauled out this fiscal year*

What is the typical number of years between haul-outs for this vessel? |___|___|

What was the cost of all other repair/maintenance for this fiscal year (not including haul-out and improvement costs)? $ |___|___|___| , |___|___|___| ☐ *not applicable – no repair/maintenance costs this fiscal year*

Please record the total payments to crew for the fiscal year (include hired captain):

$ |___|___| , |___|___|___| , |___|___|___| ☐ *not applicable - no crew payments this fiscal year*

Please record the annual cost of the benefits you provided for your crew (e.g., retirement benefits; your portion of health, life, or disability insurance premiums):

$ |___|___|___| , |___|___|___| ☐ *not applicable – no benefits were provided for the crew this fiscal year*

SECTION C – Fishing Business Related Costs

Please record the total annual cost of these following items:

Mooring/ dockage fee $ |___|___|___| , |___|___|___| ☐ *not applicable – no mooring/dockage fees*	Vessel insurance (premium) $ |___|___|___| , |___|___|___| # of months insured: |___|___| ☐ *not applicable – vessel not insured*
Use of business vehicle $ |___|___|___| , |___|___|___| ☐ *not applicable – no vehicle expense*	Cell phone and VMS costs $ |___|___|___| , |___|___|___| ☐ *not applicable – no cell phone or VMS costs*
Business travel costs (not including vehicle costs) $ |___|___|___| , |___|___|___| ☐ *not applicable – no travel costs*	Business taxes $ |___|___|___| , |___|___|___| ☐ *not applicable – no business related taxes*
Professional fees (settlement fees, accounting, legal, etc) $ |___|___|___| , |___|___|___| ☐ *not applicable – no prof. fees*	Catch handling costs (auction fees, lumping, grading, transportation) $ |___|___|___| , |___|___|___| ☐ *not applicable – no handling fees*
Association fees (cooperative, fishing organization, etc) $ |___|___|___| , |___|___|___| ☐ *not applicable – no assoc. fees*	Non-crew labor services (Night watchman, etc. Do not include repair/maint costs) $ |___|___|___| , |___|___|___| ☐ *not applicable – no service costs*
Office expenses $ |___|___|___| , |___|___|___| ☐ *not applicable – no office expenses*	Permit and/or license fees $ |___|___|___| , |___|___|___| ☐ *not applicable – no perm./lic. fees*
Principal <u>paid</u> on business loans $ |___|___|___| , |___|___|___| ☐ *not applicable – no loans*	Interest paid on business loans $ |___|___|___| , |___|___|___| ☐ *not applicable – no loans*

SECTION D -
Other *Annual* Costs not Listed in Sections B or C

(please do not record trip costs such as fuel, oil, ice, supplies, etc. - observers collect this information at sea)

Cost	Description of other annual costs
$ |___|___|___| , |___|___|___|	|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
$ |___|___|___| , |___|___|___|	|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|
$ |___|___|___| , |___|___|___|	|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|

SECTION E - Typical Lay System

What was your primary fishery (based on revenue) this fiscal year? Please list only one (e.g., groundfish, scallops, etc.)

|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|

For the primary fishery you listed above, which best describes how the crew (including the captain) is paid:

☐ Clear lay (gross stock is split between boat and crew; then trip expenses are deducted from the crew's share)

☐ Broken lay (trip expenses are deducted from the gross stock; then split between boat and crew)

☐ Per-trip or hourly wage ☐ Other -- please describe in the comments section below

For clear or broken lay systems, what is the percentage share to the boat and crew? (should add to 100%)

|___|___| % Boat (owner) share |___|___| % Crew share (include hired captain's share)

For clear or broken lay systems, which trip expenses are normally deducted? (*check all that apply*)

☐ Fuel ☐ Water ☐ Oil/lubrication ☐ Lost/damaged gear ☐ Fishing quota or days-at-sea

☐ Food ☐ Bait ☐ Unloading fees ☐ Settlement fees ☐ Other: |___|___|___|___|___|

☐ Ice ☐ Electronics ☐ Cell phone ☐ General fishing supplies (hooks, bags, totes, gloves, etc.)

For the primary fishery you listed above, do you hire a captain? ☐ Owner operated ☐ Hired captain

How many years of experience does the captain have in the primary fishery you listed above? |___|___| years

What is the size of the crew in the primary fishery you listed above? |___|___| crew members
(include the captain)

Please use this space to provide additional information or comments

|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|

|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|

|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|

|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|

|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|

|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|___|

Publishing in *NOAA Technical Memorandum NMFS-NE*

Manuscript Qualification

This series represents a secondary level of scientific publishing in the National Marine Fisheries Service (NMFS). For all issues, the series employs thorough internal scientific review, but not necessarily external scientific review. For most issues, the series employs rigorous technical and copy editing. Manuscripts that may warrant a primary level of scientific publishing should be initially submitted to one of NMFS's primary series (*i.e., Fishery Bulletin, NOAA Professional Paper NMFS,* or *Marine Fisheries Review*).

Identical, or fundamentally identical, manuscripts should not be concurrently submitted to this and any other publication series. Manuscripts which have been rejected by any primary series strictly because of geographic or temporal limitations may be submitted to this series.

Manuscripts by Northeast Fisheries Science Center (NEFSC) authors will be published in this series upon approval by the NEFSC's Deputy Science & Research Director. Manuscripts by non-NEFSC authors may be published in this series if: 1) the manuscript serves the NEFSC's mission; 2) the manuscript meets the Deputy Science & Research Director's approval; and 3) the author arranges for the printing and binding funds to be transferred to the NEFSC's Research Communications Branch account from another federal account. For all manuscripts submitted by non-NEFSC authors and published in this series, the NEFSC will disavow all responsibility for the manuscripts' contents; authors must accept such responsibility.

The ethics of scientific research and scientific publishing are a serious matter. All manuscripts submitted to this series are expected to adhere -- at a minimum -- to the ethical guidelines contained in Chapter 2 ("Publication Policies and Practices") of the *Scientific Style and Format: the CSE Manual for Authors, Editors, and Publishers*, seventh edition (Reston VA: Council of Science Editors). Copies of the manual are available at virtually all scientific libraries.

Manuscript Preparation

Organization: Manuscripts must have an abstract, table of contents, and -- if applicable -- lists of tables, figures, and acronyms. As much as possible, use traditional scientific manuscript organization for sections: "Introduction," "Study Area," "Methods & Materials," "Results," "Discussion" and/or "Conclusions," "Acknowledgments," and "References Cited."

Style: All NEFSC publication and report series are obligated to conform to the style contained in the most recent edition of the *United States Government Printing Office Style Manual*. That style manual is silent on many aspects of scientific manuscripts. NEFSC publication and report series rely more on the *CSE Style Manual,* seventh edition.

For in-text citations, use the name-date system. A special effort should be made to ensure that the list of cited works contains all necessary bibliographic information. For abbreviating serial titles in such lists, use the guidance of the International Standards Organization; such guidance is easily accessed through the various Cambridge Scientific Abstracts' serials source lists (see *http://www.public.iastate.edu/~CYBERSTACKS/JAS.htm*). Personal communications must include date of contact and full name and mailing address of source.

For spelling of scientific and common names of fishes, mollusks, and decapod crustaceans from the United States and Canada, use *Special Publications* No. 29 (fishes), 26 (mollusks), and 17 (decapod crustaceans) of the American Fisheries Society (Bethesda MD). For spelling of scientific and common names of marine mammals, use *Special Publication* No. 4 of the Society for Marine Mammalogy (Lawrence KS). For spelling in general, use the most recent edition of *Webster's Third New International Dictionary of the English Language Unabridged* (Springfield MA: G. & C. Merriam).

Typing text, tables, and figure captions: Text, tables, and figure captions should be converted to Word. In general, keep text simple (*e.g.*, do not switch fonts and type sizes, do not use hard returns within paragraphs, do not indent except to begin paragraphs). Also, do not use an automatic footnoting function; all notes should be indicated in the text by simple numerical superscripts, and listed together in an "Endnotes" section prior to the "References Cited" section. Especially, do not use a graphics function for embedding tables and figures in text.

Tables should be prepared with a table formatting function. Each figure should be supplied in digital format (preferably GIF or JPG), unless there is no digital file of a given figure. Except under extraordinary circumstances, color will not be used in illustrations.

Manuscript Submission

Authors must submit separate digital files of the manuscript text, tables, and figures. The manuscript must have cleared NEFSC's online internal review system. Non-NEFSC authors who are not federal employees will be required to sign a "Release of Copyright" form.

Send all materials and address all correspondence to: Jarita A. Davis (Editor), Editorial Office, NMFS Northeast Fisheries Science Center, 166 Water Street, Woods Hole, MA 02543-1026.

Northeast Fisheries Science Center
Operations, Management & Information Division
Research Communications Branch
Editorial Office

National Marine Fisheries Service, NOAA
166 Water St.
Woods Hole, MA 02543-1026

MEDIA MAIL

Publications and Reports of the Northeast Fisheries Science Center

The mission of NOAA's National Marine Fisheries Service (NMFS) is "stewardship of living marine resources for the benefit of the nation through their science-based conservation and management and promotion of the health of their environment." As the research arm of the NMFS's Northeast Region, the Northeast Fisheries Science Center (NEFSC) supports the NMFS mission by "conducting ecosystem-based research and assessments of living marine resources, with a focus on the Northeast Shelf, to promote the recovery and long-term sustainability of these resources and to generate social and economic opportunities and benefits from their use." Results of NEFSC research are largely reported in primary scientific media (*e.g.*, anonymously-peer-reviewed scientific journals). However, to assist itself in providing data, information, and advice to its constituents, the NEFSC occasionally releases its results in its own media. Currently, there are three such media:

NOAA Technical Memorandum NMFS-NE -- This series is issued irregularly. The series typically includes: data reports of long-term field or lab studies of important species or habitats; synthesis reports for important species or habitats; annual reports of overall assessment or monitoring programs; manuals describing program-wide surveying or experimental techniques; literature surveys of important species or habitat topics; proceedings and collected papers of scientific meetings; and indexed and/or annotated bibliographies. All issues receive internal scientific review and most issues receive technical and copy editing.

Northeast Fisheries Science Center Reference Document -- This series is issued irregularly. The series typically includes: data reports on field and lab studies; progress reports on experiments, monitoring, and assessments; background papers for, collected abstracts of, and/or summary reports of scientific meetings; and simple bibliographies. Issues receive internal scientific review, but no technical or copy editing.

Resource Survey Report (formerly *Fishermen's Report*) -- This information report is a quick-turnaround report on the distribution and relative abundance of selected living marine resources as derived from each of the NEFSC's periodic research vessel surveys of the Northeast's continental shelf. There is no scientific review, nor any technical or copy editing, of this report.

OBTAINING A COPY: To obtain a copy of a *NOAA Technical Memorandum NMFS-NE* or a *Northeast Fisheries Science Center Reference Document*, or to subscribe to the *Resource Survey Report*, either contact the NEFSC Editorial Office (166 Water St., Woods Hole, MA 02543-1026; 508-495-2228) or consult the NEFSC webpage on "Reports and Publications" (*http://www.nefsc.noaa.gov/nefsc/publications/*).