Fisheries Annual Fixed Cost Data Collection and Estimation Methodology: *An application in the Northeast, U.S.*

Chhandita Das[ab1]
NOAA Fisheries Service, Northeast Fisheries Sciences Center, Social Sciences Branch
15 Carlson Lane, Falmouth, MA 02540
dchhandita@gmail.com

a. NYISO, 10 Krey Blvd, Rensselaer, NY 12144.
b. Integrated Statistics, 16 Sumner Street, Woods Hole, MA 02543.

Corresponding Author: Chhandita Das, Email: dchhandita@gmail.com

[1] Present Address: 10 Krey Blvd, Rensselaer, NY 12144.

Introduction

In the US, National Oceanic and Atmospheric Administration Fisheries (NOAA Fisheries) is responsible for protecting the nation's ocean resources and habitat. NOAA Fisheries carries out its duties by following the standards in the Magnuson-Stevens Fishery Conservation and Management Act [26]. The act was created to manage and conserve the nation's fisheries, maintain the resiliency of the marine ecosystem and coastal communities via rebuilding overfished fisheries, protecting essential fish habitat and reducing bycatch. Further, the Marine Mammal Protection Act and the Endangered Species Act were enacted to preserve and recover protected marine species such as whales, sea turtles and corals while allowing for other non-consumptive uses [27, 28]. In the northeast region, the Social Sciences Branch (SSB) of the Northeast Fisheries Science Center (NEFSC) conducts socio-economic research and analyses on a range of issues to meet the legislative requirements of these conservation and management acts.

The SSB, on a regular basis, is involved in providing estimates of the economic and social impacts of proposed and final fishery management actions, estimating economic profitability of segments of the northeast fleet, and assessing the short-run and long-run financial viability of the industry. These analyses inform fisheries managers of the differential impacts some management actions may have on different sectors of the industry, ports and/or different groups of fishermen and fishing communities. Information on the cost of fishing, although necessary for properly carrying out these economic and financial analyses, is either missing or limited. Therefore, in the absence of complete information on cost, analysts produce results either based on only revenues ignoring the cost aspects or use only partial cost information. These estimates are able to give only an incomplete picture, and do not depict the true financial status of the fishery.

Typically, fisheries experience three major types of cost: *annual fixed costs*; which are incurred annually irrespective of whether any fishing activity takes place or not; *variable costs*, which are accrued when a fishing trip occurs; and *labor costs*, which consists of payments to crew and captain. These cost data along with the revenue data, and other fisheries complementary data enable a net revenue type analyses which help fisheries managers make more informed decisions. In addition, the cost data can be used in measuring other types of financial statements (eg. Cash-flow) which can give a comprehensive overview of the fishery as well. Liese *etal* [17] elaborate on different financial accounting methods for the Federal Shrimp Fishery in the Gulf of Mexico.

Comprehensive and consistent availability of fisheries cost data is essential for effective management decisions as well as to measure and track economic and financial performance of the industry. The NEFSC collects fisheries operating costs on a continuous basis as a part of fisheries observer program that focuses primarily on biological and fishing effort data. Fisheries catch and revenue data are also collected on a regular basis via mandatory vessel log book and dealer reporting. However, currently, there exists no established mechanism to collect fisheries annual fixed cost data on a regular basis. This study discusses a survey methodology that attempts to fill this data gap. The primary focus of this data collection effort is to gather cost information that are directly related to owning and maintaining a fishing vessel, ignoring wholesale, retail and processing. These costs, although vary across year, are typically fixed within a year hence referred to here as annual fixed costs. The methodology discussed here can easily be adopted by other agencies carrying out similar analysis requiring cost information.

The paper is organized as follows: Section 2 gives a brief description about the previous annual fixed cost data collection efforts undertaken by the SSB. Section 3 discusses the survey methodology and Section 4 describes the data. Section 5 explains the cost models, estimation results and Section 6 concludes the paper.

1. Background

In the past, the NEFSC collected fisheries annual fixed cost and labor cost data only periodically via various methods, such as focus group sessions, surveys, capital construction fund and National Marine Fisheries Services' special reporting [1, 8, 13, 22, 23]. In order to have these data available more frequently and comprehensively, the SSB implemented an organized data collection effort in 2007. During this period, data were collected via voluntary mail surveys. In each survey year, approximately 2,700 surveys were sent to the entire population of active fishing vessels in the northeast along with their federal permit renewal forms. The response rate for 2007 was around 21%, which fell to 17% in 2008 and eventually to 8% in 2009, prompting SSB to discontinue the survey. Following this effort, a considerable amount of time was spent in studying the pros and cons of this implementation. The survey methodology, administration, response rates and data were thoroughly analyzed to identify places for improvement. Details from this background study can be found in a NOAA technical report [2]. This study discusses an improved and modified survey effort that was launched in 2012 incorporating the lessons learned from the 2007-2009 survey implementations.

2. Survey Methodology

The ground work for the recent survey effort began in late 2010 and was finally implemented in 2012 and 2013 to collect data on costs incurred in 2011 and 2012, respectively. Reference to survey years in this document is based on the year during which costs were incurred. A significant amount of time was spent researching sampling methodology and improving the survey instrument. Several measures were taken to improve response rates and data validity. The major steps in this modified approach are explained below.

2.1 Population and Sample

During the 2007-2009 data collection effort, surveys were sent each year to all vessel owners who had a federal permit and were actively fishing. This process led to an active vessel owner receiving at least one survey every year; a vessel owner received multiple surveys if he owned multiple vessels. The data analysis showed that such a broad-based approach resulted in high non-response rates, a possible consequence of survey fatigue and annoyance among the respondents. Moreover, vessel owners often receive other fishing related surveys, adding to their burden which further contributed to non-response. To reduce respondent burden, the 2011-2012 surveys were sent to a selected sample of the population instead of the entire population of active fishing vessels. To ensure that all major groups of the fishing population were properly represented in the sample, a stratified sampling method was used. The strata were defined using the vessel characteristics from NEFSC vessel log book data and vessel permit database. After experimenting with several vessel characteristics on which to base strata, vessel length and fishing gears were chosen to define the strata.

A multi-stage stratified random sampling method was used to select the sample. Vessels were first grouped into seven broad gear groups based on their principal gears. Vessels typically use multiple gears for their fishing activities, therefore principal gears were defined based on the gears that accounted for the maximum annual revenue earned by the vessels during that year. The broad gear groups were: dredge, gillnet, hand-gear, longline, pot/trap, trawl and purse/seine. There were few principal gears that did not fall under these seven broad categories; they were considered not relevant for core analyses and the associated vessels (about 20 each year) were excluded from the survey frame.

Within a gear category, costs are expected to vary by vessel size. Therefore, in the next step, vessels were further grouped into two length categories based on the mean length of the vessels within each gear group. Vessels were classified as large or small, depending on whether their length fell above or below the average vessel length for each gear group. These average vessel length and gear-groups which formed the basis of strata classifications, are shown in Table 1 for each survey year. The longline and purse/seine strata were very small, consisting of about 50 vessels or less; therefore they were not categorized further by length. This process resulted in 12 strata.

< Table 1 here >

The frequency distributions of the population and sample by strata and survey year are shown in Table 2. The potential respondent population consisted of all commercial fishing vessels holding a current northeast region permit in any fishery and actively fishing during the calendar year of study. A fishing vessel was considered active if it reported a positive amount of landings either through the northeast seafood dealer reporting system or vessel trip logs during the calendar year of study. The sample for 2011 consisted of 1,703 vessels, about 42% of the population of 4,008 vessels. For 2012, the sample size was 1,875, which was about 49% of the vessel population size of 3,821.[1] Vessels, selected for the 2011 sample were not considered for the 2012 sample selection. Moreover, additional steps were taken to prevent multiple vessel owners from being selected more than once.

< Table 2 here >

3.2. Survey Administration

The survey was offered via mail and online. It was expected that providing both options would improve overall response rates. The online version may be more convenient for some since it does not require the extra step of mailing the form. An advantage of online surveys is also that they are free of errors associated with illegibility issues. Dillman et al. [5] discusses the advantages and disadvantages of using mixed mode surveys.

A private consulting company was hired to administer the survey. The company was responsible for building and hosting the online survey via a secured web address and pre-testing both the online and the paper survey instruments with selected vessel owners. They also maintained a toll free phone line to address any questions or concerns that the survey recipients had. The survey was administered in several steps. First, selected vessel owners were sent a pre-notification letter. The letter notified vessel owners about the impending survey, stated the purpose of the survey and described how the cost data would be used. Research has shown that pre-notification letters improve response rates in mail surveys by 3 to 6 percentage points [3, 6, 14]. This step was followed by mailing the survey packet, which contained the instrument, a stamped self-addressed return envelope and information to access the online survey. The packet also included a one page flier explaining the objective and use of this survey, and requesting the vessel owner's co-operation. In the second year of the survey, the packet also contained a brief summary of the data obtained from the previous year.

After the mailing of the first survey packets, the non-respondents were followed up by first mailing a reminder post-card, then by mailing another survey packet, and finally by calling them by phone. Studies have shown that a "special contact" improves response rate in mail surveys [4, 5, 9]. The contractor was entrusted with carrying out all these steps including printing of all the survey materials. They were also responsible for contacting the respondents in an effort to obtain missing responses and clarifying incomplete responses as per guidelines provided by the SSB staff.

3.3. Survey Response Rates

[1] The initial plan was to select 50% of the population to be included in the sample for each year; however final sample sizes were revised due to budgetary restrictions.

The 2011 surveys were sent to 1,703 unique vessel owners, of which only 1,457 received the survey and therefore could have completed it. This number of final recipients is referred to as the valid sample size. The remaining surveys were either undeliverable due to bad addresses or were considered out-of-scope because the owners sold their boats or were inactive most of the year. The returned surveys included complete responses, incompletes, duplicates and refusals. The number of responses was 478 (80% mail; 20% online), of which 437 were considered to be complete or partially complete. Responses were considered partially complete if sufficient numbers of cost related questions in each section were answered. The sufficient number was determined by examining response for each question, and judging when and if the missing and zero responses could be deemed acceptable. If no cost related questions were answered, then that survey was not used for analysis. For 2012 survey, the valid sample size was 1,787, and 396 surveys were returned (85% mail; 15% online), of which 376 were complete or partially complete.

Table 3 shows the response rates along with various outcome rates that are often calculated in survey research. The *response rate* was calculated as the total number of complete and partially complete surveys as a percentage of the valid sample size. The *response rate* was 30% in 2011 and 21% in 2012. The *co-operation rate* was calculated by considering all the returned surveys that were complete, partially complete, or incomplete, which was 31% in 2011 and 22% in 2012. The *refusal rate* was 12% in 2011 and 18% in 2012. The *contact rate* was calculated by accounting for all the sample units, responders or non-responders, with whom a contact was made either via returned surveys, phone or email. Contact was made with 51% of the sample in 2011 and 59% in 2012. There were several non-responders with whom no contact was made. The *response*, *co-operation* and *refusal rates* were calculated as percentages of the valid sample size, and the *contact rate* was calculated as a percentage of the total sample size. Note, there is no standard way to calculate these rates as researchers often use different methods to calculate the same rates [30].

< Table 3 here >

Table 4 shows a further breakdown of the response rates by strata and survey year. Since the primary objective of this data collection effort is to evaluate commercial fisheries performance and management impacts, the responses from the recreational vessels were not included in data analysis. Because of this reason, the sample sizes displayed in the following table are lower than the numbers in Table 3. Moreover, in 2012, eight surveys were received that had no vessel identifier. Since these survey responses could not be put into appropriate strata, they were not included in the response rate breakdown calculations.

< Table 4 here >

3. Survey Layout and Data

The survey instrument had six sections. Section A focused on vessel related questions, such as vessel ownership type and vessel value; Section B contained questions about repair/maintenance and upgrade/improvement related costs. These cost questions were presented by seven separate categories, which were propulsion engine, deck equipment/other machinery, hull, fishing gear, wheelhouse and gear electronics, processing/refrigeration and safety equipment. Section C contained fishing business related costs itemized by seventeen separate expense categories (e.g. office cost, vehicle costs and permit fees).[2] Section D asked about trip related costs, itemized by ten separate items, and Section E had questions on crew size and crew payments. Section F had open questions that allowed respondents to enter other costs not reported elsewhere in the survey [24, 25].

[2] The 2012 instrument splits section C into two sections to better cater to the multiple vessels owners by distinguishing between costs related to the specific vessel identified in the survey and the costs specific to the entire business.

Costs were grouped into five categories for summarization: repair/maintenance (RM), upgrade/improvement (UI), business, haul-out, and operating or variable costs. For upgrade and improvement costs, depreciation factors were applied to each subcategory using a seven year expected lifetime and were obtained from the IRS online resources [11]. All costs that are fixed within a year were summed together and reported as annual fixed costs, which were also used for modelling. Variable costs and labor costs are not modeled in this paper. Variable costs were collected in this survey only to validate these costs collected from other sources.

The data were corrected for any identifiable errors and anomalies before their use in analysis. To do this, each survey response was extensively reviewed to identify data reporting errors. Respondent comments and remarks from surveys were carefully noted and cost values were updated as appropriate. Next, data were plotted at the individual level, aggregate level and by vessel characteristics to identify outliers. After removal of invalid survey responses and outliers, there were 658 usable responses (366 from 2011 and 292 from 2012). There were about 50 surveys in 2012 which could not be used for analysis, as these respondents reported their firm level business costs, but did not provide their vessel numbers and hence their vessel level business costs could not be calculated.

Non-response bias was checked by comparing respondent and non-respondent across frame variables and by comparing response rates across strata [16]. Information on several frame variables is available about the population from which the sample was drawn. This information was used to compare mean estimates of the vessels' physical characteristics (gross tons, horse power, length and vessel age) and total revenues for respondents and non-respondents. The statistically significance of these mean differences were tested by using t-tests, and the results from these t-tests are shown in Table 5. The significance tests for each stratum were also conducted and reported in the appendix. Most of these t-tests showed no significant differences except in vessel horse power, gross tons and length for hand-gear strata and vessel age for large trawls.

< Table 5 here >

Next, strata frequencies in respondents were compared with the same in population (Table 6). The table shows that in both survey years, some strata were over-represented and some were under-represented among respondents. A chi-square test shows that the differences in strata frequencies of the population and the respondents were statistically significant in 2011 and highly statistically significant in 2012. This indicates that the data should be weighted to account for the under and above representation of these strata while reporting aggregate summaries based on respondent values. The strata weights were calculated as the reciprocal of the probability of a respondent being in a stratum: $w_i = 1/P_i$; where $P_i = n_i/N_i$; n_i is the respondents frequency in stratum i and N_i is the population frequency of stratum i [18].

< Table 6 here >

Weighted summary statistics of the costs from both survey years are shown in Table 7. For this part of the analysis, all 2011 dollar values were converted to 2012 dollars.[3] As the table shows, on average, most of these costs are consistent across the two years except the repair/maintenance costs and operating costs. The average repair/maintenance cost was $30,441 in 2011, about 1.5 times more than the average repair/maintenance cost in 2012 ($19,953). This result is not surprising since it is expected that vessels might incur higher repair costs in some years and lower costs in another. The average upgrade/improvement cost was $3,358 in 2011 and $2,442 in 2012. The average business cost was around $40,000 in both years. The table also shows haul-out costs, which were around $2,000 on average. Mean operating cost was $80,854 in 2011 and $58,175 in 2012. A detailed breakdown of these costs is shown in Table 8 for both years combined. For this table, vessel lengths were categorized into three size groups: over 80 ft, between 40 and 80 ft and less

[3] Producer price index for unprocessed fin fish was used for this conversion.

than 40 ft. As expected, large vessels are associated with higher costs in each cost category. The modeling framework for estimating and predicting costs is discussed in the following section.

< Table 7 here >

< Table 8 here >

4. Estimation and Prediction

5.1. The Modeling Framework

As discussed above, although cost information is critical in several analyses relevant to commercial fisheries management, it is extremely challenging to obtain this information. Surveys are a useful tool to meet this gap. 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 expect a 100% response rate. As this analysis showed, even after taking several measures to improve response rates, less than 60% of the vessels selected to be in the sample could be reached either via mail or phone, and fewer of them actually returned the survey (Table 3). Therefore, it is particularly important to build robust statistical models with reasonable accuracy that can estimate costs for all the vessels that were not surveyed or did not respond to the survey.

In this application, an aggregate annual fixed cost model was estimated which is the sum of repair/maintenance, upgrade/improvement, business and haul-out costs. Typically, researchers use an average cost criterion or an ordinary least square method (OLS) to estimate fisheries costs [7, 13]. Common practice is to estimate OLS with the log transformation of the cost values and then obtain predictions in raw scale by exponentiating the predicted cost values in log scales. This re-transformation, though frequently used, causes bias [12, 19, 20]. The bias is greater if there is heteroscedasticity (non-constant variance) in the log-transformed model. To correctly predict cost when using the log-transformed linear model estimation, proper adjustment should be applied with an anti-log-transformation. However, this adjustment process for un-biasedness can be time and labor intensive, and thus makes the simple OLS method more complicated. An alternative method to avoid these limitations is to use a generalized linear model (GLM).[4]

A GLM can be viewed as a differentially weighted non-linear least square estimation method. The advantages in using a GLM approach are that: (a) there is no re-transformation bias; (b) no adjustment is needed for an anti-log transformation; and (iii) GLM does not assume constant variance. GLM is also a preferred method for analyzing skewed data, which are often encountered with cost data. GLMs are applied widely in analyzing human health expense data where skewness of the distribution is common [15, 21, 31].

Since fisheries cost distribution is right skewed, similar to health expense data, GLM is an appropriate approach. The skewness of the cost distribution is evident from Figure 1. 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

---

[4] *As an alternative estimation criterion, Heckman two-step estimation method was also explored in order to assess and correct for non-response bias, which is common in survey research [10, 29]. Except few gear dummies, almost all variable had no significant impact in the selection model. In the second step of the cost estimation model, the Inverse Mills Ratio, derived from the selection model, was also insignificant indicating absence of significant non-response bias. Based on these results and the mean comparison of key characteristics between respondents and non-respondents (Table 5 and A1), the GLM was chosen as the final model.*

GENMOD procedure in SAS.[5] Manning and Mullahy [20] suggests several tests to choose between the log-transformed OLS and GLM. They suggest using the GLM estimator if the log-scale residuals have coefficients of kurtosis of less than three. Moreover, they also suggest using GLM over log-transformed OLS if the OLS residuals on the log scale are heteroscedastic in some or all of the covariates. In this application, both tests suggest consideration of GLM.[6] In addition, the strata weights computed above were used in the WEIGHT statement of the SAS GENMOD syntax.

< Figure 1 here >

To identify the best predictors for the GLM, several continuous and categorical variables were constructed using vessel information from various data reporting sources within NEFSC (Table 9). For model estimations, both years of data were combined. Principal gear categories were further grouped into three major gear types: *static, mobile and purse/seine.* The *Mobile* gear group included dredge and trawl, whereas the *static* gear groups included gillnet, longline, pot/trap and hand-gear. Although there were 658 responses in the estimation set, information about vessel characteristics were not available for all vessels.

Table 9 shows that the vessels included in the estimation sample, on average, were 44 ft long, built about 23 years ago, weighed 36 gross tons, and were equipped with 428 horse power engines. The majority of these vessels used static gear (80%), were constructed of fiberglass (74%) and landed their catch in the New England region (78%).

< Table 9 here >

The final model specification was chosen based on the log-likelihood values, MAE (Mean Absolute Error) and RMSE (Root Mean Squared Error). RMSE was calculated by (a) taking the square of the differences between the predicted and reported costs; (b) taking the mean of these squared differences; and (c) then taking the square root of the mean. A low RMSE indicates a better fit [21].

5.2. Results and Cost Predictions

The GLM estimates for the cost model are shown in Table 10. The interpretation of the coefficients in a log-linear model is different from a linear model. In a log-linear model the impact on the dependent variable is evaluated in terms of the exponent of the estimated coefficients.

< Table 10 here >

All the variables are significant in the model (Table 10). The estimates show that vessels which are large and are made of fiberglass materials have higher annual costs. Vessels primarily using mobile gears have significantly higher annual costs than vessels primarily using static gears. This might be because vessels with mobile gear undergo more repair maintenance costs and hence incur higher costs annually. The vessel age variable has a negative coefficient implying vessels that are older, that's vessels that are being used for a several years are associated with a lower annual costs than vessels which are relatively newer. A dummy variable capturing the primary port region is significant in this model. This result is expected since several business costs such as mooring costs, travel costs and insurance costs may vary by geographical area. Estimates show that the vessels that chose Mid Atlantic as their primary port region have lower annual costs than the vessels choosing other parts of the northeast as their primary port.

---

[5] *Gamma distribution is a preferred distribution choice for the error term in cost analysis with GLM [15, 21].*
[6] *Kurtosis values varied from 2 to 3 based on different model specifications.*

The model estimates were used to predict costs for all the vessels that were actively fishing during 2012, and for which the independent variables could be defined. Predicted costs at the raw scales were obtained by taking the exponent of the predicted costs at log scale. Table 11 shows the summary statistics of the predicted costs by strata and overall. Cost summaries for the strata with less than 3 vessels are not reported for confidential reasons. The predicted costs were calculated only for year 2012, and are reported in the same table with the reported costs for the same year for ease of comparison. Overall, average predicted *annual fixed costs* cost is $70,369 which is higher than the costs reported by the survey respondents. The breakdown by strata however shows mixed results. For some strata, the reported and predicted costs are consistent such as large dredge and pot/trap. On the other hand, for some strata there is large divergence between the predicted and reported costs such as large handgear, purse/seine and small trawl vessels. For large handgear, the predicted costs were about 2.6 times higher than the reported costs and for purse/seine and small trawls it was about 1.2 times higher than the reported costs. Figure 2 displays boxplots of the reported and predicted cost distributions.

< Table 11 here >

<Figure 2 here >

5. Discussion

Economic data on the costs of operating commercial fishing businesses are needed to accurately inform fisheries managers about the impacts of proposed regulations, assess the state of the fishery over time and calculate performance metrics under continuously changing management regimes. A survey methodology for collecting cost data is described in this paper. The data cleaning process and data summaries by vessel characteristics are also discussed. An econometric model was developed to enable cost predictions for vessels that were not surveyed or did not respond, and for time periods beyond the survey years. In future, these cost estimates will be used to conduct net revenue type analysis which will be tremendously helpful in understanding and tracking fishery performances. In addition, a cash-flow type analysis can also be calculated with these data.

Annual fixed cost data, along with trip cost and labor cost information, provides a complete picture of the state of commercial fishing businesses, including profitability and return on investments, and the long term financial viability of the fishery. Therefore, it is necessary to put forward effort towards obtaining reliable, timely, and updated information on commercial fishing fixed costs data. However, gathering this information on fishing cost is very challenging for several reasons. Fishermen are often not willing to share their economic information either out of mistrust regarding the use of the data or from fear of their proprietary business information getting into the wrong hands. Moreover, collecting cost information at the vessel level becomes quite challenging in cases where several fishermen own multiple vessels and costs are shared among these multiple vessels. Poor record keeping may also contribute to the availability of these data. The biggest hurdle in getting responses is the general reluctance among the vessel owners towards data reporting which is due to their already existing different types of data reporting requirement.

Going forward, the finding from this analysis will be used to improve survey administration, response rates and data quality in the faces of challenges discussed above. As seen during the 2007-2009 and the latest implementation, repeating the survey in consecutive years have a negative impact in response rates. With that in mind, the survey will be conducted only in regular intervals instead of repeating it every year. Feedback from the contracting company will be used to make the survey administration more efficient. In addition, data summaries and analysis based on the collected cost data will be made available to the fishermen with the expectation of better cooperation from the vessel owner in future data collection efforts. Another strategy to encourage response rate might be to shorten the length of the survey, that is, targeting to gather cost data at the aggregate level rather than at the item level as attempted here. Following these steps can help agencies engaged in similar data collection activities carry out a successful project.

Acknowledgments

The study was funded NOAA Fisheries. The author thanks Andrew Kitts and Dr. Tammy Murphy for their contribution to this project and John Walden for his feedback on the manuscript. The author also thanks the contracting company, *Eastern Research Group*, for administering the survey. Thanks are also due to hundreds of fishing vessel owners who responded to the survey, participated in the focus-groups and provided numerous feedbacks along the way.

References

[1] Crutchfield SR, Gates JM. The Impact of extended fishery Jurisdiction on the New England otter trawl fleet. Marine Resource Economics 1985; 2(2):153-173.

[2] Das Chhandita. An overview of the annual cost survey protocol and results in the northeast (2007-2009). NOAA Technical memorandum NMFS NE-226 2013.

[3] Dillman DA. The design and administration of mail surveys. Annual Review of Sociology 1991; 17:225-249.

[4] Dillman DA, Christenson JA, Carpenter EH, Brooks R. Increasing mail questionnaire response: a four state comparison. American Sociological Review 1974; 39 (5):744-756.

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

[6] Fox RJ, Crask MR, Kim J. Mail survey response rate a meta analysis of selected techniques for inducing response. Public Opinion Quarterly 1988; 52:467-491.

[7] Funk RD, Griffin WL, Mjelde JW, Ozuna T, Ward JM. A method of imputing and simulating costs and returns in fisheries. Marine Resource Economics 1998; 13(3):171-183.

[8] Gautam A, Kitts AW. Data description and statistical summary of the 1983-1992 cost-earnings data base for northeast U.S. commercial fishing vessels: a guide to understanding and use of the data base. NOAA Technical Memorandum NMFS-NE-112. 1996.

[9] Heberlein TA, Baumgartner, R. Factors affecting response rates to mailed questionnaires: a quantitative analysis of the published literature. American Sociological Review 1978; 43(4): 447-462.

[10] Heckman JJ. Sample selection bias as a specification error. Econometrica 1979 (47): 153-161.

[11] IRS depreciation and amortization for the fishing industry. http://www.irs.gov/Businesses/Small-Businesses-&-Self-Employed/Depreciation-and-Amortization-for-the-Fishing-Industry. Accessed on March, 2015.

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

[13] Jin D, Thunberg E, Kite-Powell H, Blake K. Total factor productivity change in the New England groundfish fishery: 1964-1993. Journal of Environmental Economics and Management 2002; 44(3):540-556.

[14] Kanuk L, Berenson C. Mail surveys and response rates: a literature review. Journal of Marketing Research 1975; 12(4):440–453.

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

[16] Lian C. West coast limited entry groundfish cost earnings survey: protocol and results for 2008. NOAA Technical Memorandum NMFS-NWFSC-121 2012.

[17] Liese, C, Travis M, Pina D, Water, R. The annual economic survey of Federal Gulf Shrimp permit holders: report on the design, implementation and descriptive results for 2006. NOAA Technical Memorandum NMFS-SEFSC-584 2009.

[18] Lohr SL. Sampling: design and analysis, second edition. Cengage Learning 2010.

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

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

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

[22] National Marine Fisheries Service. Basic economic indicators: Atlantic and Pacific groundfish 1932-1972, Current fisheries statistics No 6271, COM-74-11638, Economic Research Laboratory, Washington D.C. 1974.

[23] National Marine Fisheries Service. Revenue, costs and returns from vessel operation in major U.S. fisheries. Washington D.C. 1977.

[24] Northeast Fisheries Science Center Annual Cost Survey 2012. http://www.nefsc.noaa.gov/read/socialsci/pdf/Annual_Cost_Survey_Final_Version_April_24_2013.pdf. 2013.

[25] Northeast Fisheries Science Center Annual Cost Survey 2011. http://www.nefsc.noaa.gov/read/socialsci/pdf/2012-08-06_NE_Annual_Cost_Survey.pdf. 2012.

[26] NOAA Fisheries, 2007. Magnuson-Stevens Reauthorization Act Details. http://www.nmfs.noaa.gov/sfa/laws_policies/msa/msa_2007.html. Accessed on April, 2015.

[27] NOAA Fisheries, Protected Resources, Marine Mammal Protection Act, 1972. http://www.nmfs.noaa.gov/pr/laws/mmpa/. Accessed on April, 2015.

[28] NOAA Fisheries, Protected Resources, Endangered Species Act, 1973. http://www.nmfs.noaa.gov/pr/laws/esa/ Accessed on April, 2015.

[29] Sales AE, Plomondon ME, Magid DJ, Spertus JA, Rumsfeld JS. Assessing response bias from missing quality of life data: the Heckman method. Health and quality of life outcomes 2004; 2(49):1-10. *DOI: 10.1186/1477-7525-2-49.*

[30] The American Association for Public Opinion Research. Standard definitions: final dispositions of case codes and outcome rates for surveys, seventh edition 2011.

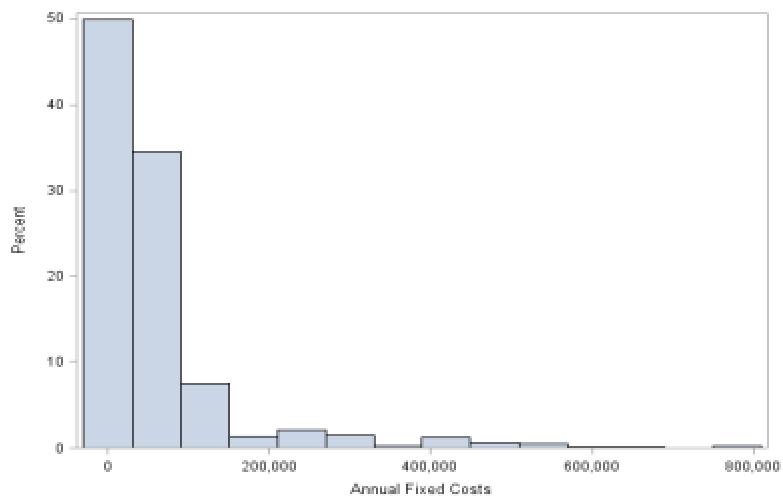
[31] Wu EQ, Birnbaum HG, Zhang HF, Ivanova JI, Yang E, Mallet D. Health care costs of adults treated for attention-deficit/hyperactivity disorder who received alternative drug therapies. Journal of Managed Care Pharmacy 2007; 13(7): 561-569.

Appendix

< Table A1 here. >

Figure 1: Histogram of annual fixed cost distribution

Figure 2: Boxplots of reported and predicted annual fixed costs


Percent
50
40
30
20
10
0
0
200,000
400,000
600,000
800,000
Annual Fixed Costs

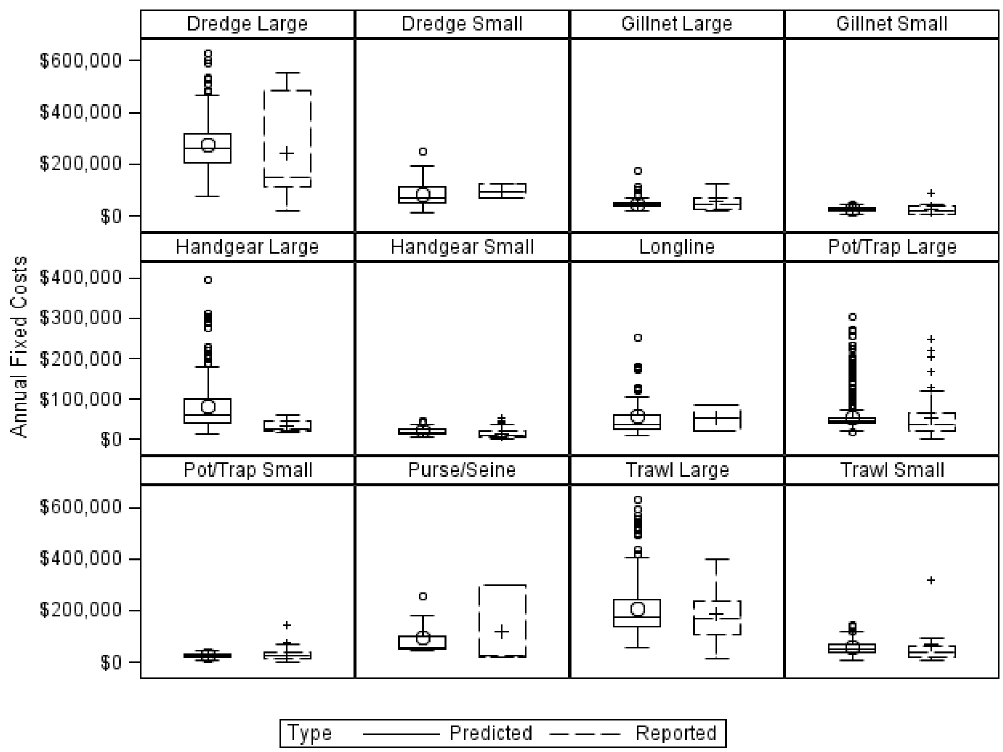

Dredge Large
Dredge Small
Gillnet Large
Gillnet Small
Handgear Large
Handgear Small
Longline
Pot/Trap Large
Pot/Trap Small
Purse/Seine
Trawl Large
Trawl Small
Annual Fixed Costs
$600,000
$400,000
$200,000
$0
$400,000
$300,000
$200,000
$100,000
$0
Type
Predicted
Reported

Table 1: Average vessel length by principal gear groups and survey year

Gear Group	Average Length (feet) 2011	Average Length (feet) 2012
Dredge	72.47	72.07
Gillnet	40.15	40.09
Handgear	38.51	38.63
Longline	44.91	46.89
Pot/Trap	37.98	38.09
Trawl	60.99	60.85
Purse/Seine	61.03	59.54

Table 2: Frequency distribution of the sample and population by strata and survey year

Strata	Population 2011	Sample 2011	Population 2012	Sample 2012
Large Dredge	326	163	316	86
Small Dredge	199	100	193	88
Large Gillnet	140	70	140	63
Small Gillnet	142	71	143	66
Large Handgear	271	137	227	28
Small Handgear	577	216	491	202
Longline	42	21	54	34
Large Pot/Trap	898	336	683	396
Small Pot/Trap	941	353	1,112	694
Purse/Seine	13	7	13	6
Large Trawl	226	111	218	90
Small Trawl	233	118	231	123
All	***4,008***	***1,703***	***3,821***	***1,875***

Table 3: Types of communication with the survey recipients and their rates by year

Response Types	Definition	Frequency (%) 2011	Frequency (%) 2012
Response Rate	Complete and partially complete	437 (30.00)	375 (20.98)
Co-operation Rate	Returned surveys: complete, partially complete, incomplete and duplicate	454 (31.16)	394 (22.05)
Refusal Rate	Number of refusals (via phone, email or returned surveys with written refusal)	178 (12.22)	317 (17.74)
Contact Rate	All forms of communications received from the survey recipients: survey responses, refusals and out of scope status information.	876 (51.44)	1,106 (59.05)
Sample Size	*Number of vessels selected to be in the sample*	*1,703*	*1,875*

Valid Sample size	*Considering only those vessels who received a survey*	*1,457(85.55)*	*1,787 (95.36)*

Table 4: Response frequency by strata and by year

Strata	2011			2012		
	Sample Frequency	**Response Frequency**	**Response Rates (%)**	**Sample Frequency**	**Response Frequency**	**Response Rates (%)**
Large Dredge	143	29	20.28	83	16	19.28
Small Dredge	83	11	13.25	86	4	4.65
Large Gillnet	60	24	40.00	61	14	22.95
Small Gillnet	58	16	27.59	62	12	19.35
Large Handgear	28	4	20.69	27	7	25.93
Small Handgear	114	43	37.72	186	45	24.19
Longline	19	5	26.32	33	3	9.09
Large Pot/Trap	278	80	28.78	380	92	24.21
Small Pot/Trap	297	96	32.32	656	128	19.63
Purse/Seine	6	3	50.00	5	3	60.00
Large Trawl	101	33	32.67	86	22	25.58
Small Trawl	100	28	28.00	112	12	10.71
All	***1,287***	***372***	***28.90***	***1,777***	***358***	***20.19***

Table 5: Non-response bias test (Mean comparison of key characteristics)

Sample Units	Variable	Response Status	Mean (SD)	Pr > |t|
All Vessels	Vessel length (feet)	Non-Respondents Respondents	45 (18) 44 (42)	0.08
	Vessel age (years)	Non-Respondents Respondents	23 (12) 23 (12)	0.71
	Gross tons	Non-Respondents Respondents	38 (36) 36 (32)	0.23
	Vessel horsepower	Non-Respondents Respondents	439 (275.6) 428 (282.2)	0.36
	Vessel revenue ($)	Non-Respondents Respondents	273,193 (503,493) 243,196 (421,180)	0.11

Table 6: Non-response bias test (Response rates comparison)

Strata	2011		2012	
	Population frequency (%)	**Respondent frequency (%)**	**Population frequency (%)**	**Respondent frequency (%)**
Dredge-Large	9.03	7.63	8.28	3.70
Dredge-Small	5.65	2.72	5.06	1.14
Gillnet-Large	3.91	6.54	3.67	3.98
Gillnet-Small	3.97	4.36	3.75	3.42
Handgear-Large	1.84	1.09	5.95	1.99
Handgear-Small	9.81	11.41	12.71	12.53
Longline	1.17	1.36	1.41	0.85
Pot/Trap-Large	25.18	21.79	17.92	26.21
Pot/Trap-Small	26.33	26.15	29.16	36.18
Purse/Seine	0.36	0.54	0.34	0.85
Trawl-Large	6.20	8.72	5.71	5.69
Trawl-Small	6.54	7.63	6.05	3.42
Significance test	***Chi-square 21.66*** ***DF 11*** ***Pr>chisq <.027***		***Chi-square 55.90*** ***DF 11*** ***Pr>chisq <.0001***	

Table 7: Weighted summary statistics of the major cost categories and haul-out cost

Cost Categories	**2011**			**2012**			**Combined**		
	N	**Mean ($)**	**SD ($)**	**N**	**Mean ($)**	**SD ($)**	**N**	**Mean ($)**	**SD ($)**
Repair/Maintenance	352	30,441	189,282	343	19,953	88,134	695	24,972	149,099
Upgrade/Improvement (After depreciation)	217	3,358	14,445	218	2,442	10,470	435	2,880	12,681
Business	331	40,055	220,614	292	42,467	242,831	623	41,231	231,137
Haul-out cost	266	2,079	14,772	253	2,547	17,694	519	2,318	16,263
Operating cost	340	80,854	455,898	343	58,175	292,963	683	68,813	384,288
Annual fixed cost	366	69,040	352,470	292	65,750	312,128	658	67,515	334,960

Table 8: Weighted summary statistics of the major cost categories and haul-out cost by vessel length groups

Cost Category	**Length Groups**	**N**	**Mean ($)**	**SD ($)**

Repair/Maintenance	Over 80ft	42	105,916	357,757
	40ft-80ft	280	29,583	139,277
	<40ft	373	9,209	27,771
Upgrade/Improvement (After depreciation)	Over 80ft	27	5,778	18,862
	40ft-80ft	172	3,798	15,251
	<40ft	236	1,669	7,385
Business	Over 80ft	41	209,239	443,654
	40ft-80ft	234	42,567	180,402
	<40ft	348	14,690	47,491
Haul-out cost	Over 80ft	20	10,619	49,806
	40ft-80ft	198	2,770	15,672
	<40ft	301	1,139	7,335
Operating cost	Over 80ft	42	293,290	723,621
	40ft-80ft	273	83,351	391,233
	<40ft	368	22,178	59,226
Annual fixed cost	Over 80ft	42	316,754	636,453
	40ft-80ft	262	73,657	253,570
	<40ft	354	25,051	63,644

Table 9: Description of variables for estimation

Variable	**Definition**	**Mean (SD)**
Length	Vessel length in feet	44 (17)
Length_sq	Length*length	2194 (1958)
Gtons	Vessel gross tons	36 (46)
Vhp	Vessel horse power	428 (304)
Vhplen	Vessel horse power per feet	9.66 (4.29)
Age	Vessel age (in years)	23 (12)
Categorical variables		*Frequency Percentage*
Fglass	Fiberglass construction	74 %
Mobile	Mobile gear (dredge, trawl)	20 %
Reg_MA	Primary port in Mid-Atlantic region (New England is the excluded region).	22 %

Table 10: GLM estimates for annual fixed cost model

Parameter	Estimate	Standard Error	Pr > ChiSq
Intercept	7.4039	0.2964	<.0001
age	-0.0134	0.0025	<.0001
length	0.0934	0.0092	<.0001
Length_square	-0.0004	0.0001	<.0001
Mid-Atlantic Port	-0.2077	0.0671	0.002
Fiber Glass	0.3008	0.097	0.0019
Mobile gear	0.2721	0.0894	0.0023
Log-likelihood: -7,398.660; No. of Observations used: 643			

Table 11: Summaries of reported and predicted costs by strata for 2012 survey respondents and population

Strata	Reported			Predicted		
	N	Mean ($)	SD ($)	N	Mean ($)	SD ($)
Dredge Large	11	240,073	185,884	316	272,869	88,403
Dredge Small	x	x	x	193	82,942	45,026
Gillnet Large	12	53,745	32,163	139	46,250	16,838
Gillnet Small	11	27,395	22,629	139	24,092	8,653
Handgear Large	4	30,759	20,132	227	79,902	61,629
Handgear Small	37	14,099	12,596	490	18,708	7,515
Longline	x	x	x	54	55,949	49,602
Pot/Trap Large	76	53,277	50,843	684	53,550	35,895
Pot/Trap Small	106	28,465	20,370	1114	26,970	8,015
Purse Seine	3	115,537	159,833	13	93,800	64,508
Trawl Large	18	185,735	106,935	218	278,474	396,071
Trawl Small	10	68,508	92,143	231	61,036	22,813
Overall	***292***	***54,688***	***78,218***	***3,818***	***70,369***	***87,174***

Table A1. Non-response bias test (Mean comparisons of key characteristics by strata)

Strata	Variable	Response Status	Mean (SD)	Pr > |t|
Dredge-Large	Vessel length	Non-Respondents	84 (9)	0.30

	(feet)	Respondents	85 (8)	
	Vessel age (years)	Non-Respondents Respondents	26 (12) 27 (13)	0.36
	Gross tons	Non-Respondents Respondents	152 (33) 153 (34)	0.80
	Vessel horsepower	Non-Respondents Respondents	780 (281.7) 990 (744.7)	0.08
	Vessel revenue ($)	Non-Respondents Respondents	1,464,853 (670,539) 1,501,037 (619,323)	0.74
Dredge-Small	Vessel length (feet)	Non-Respondents Respondents	53 (12) 47 (11)	0.10
	Vessel age (years)	Non-Respondents Respondents	27 (12) 30 (14)	0.47
	Gross tons	Non-Respondents Respondents	51 (38) 36 (31)	0.12
	Vessel horsepower	Non-Respondents Respondents	412 (162.7) 378 (150.2)	0.42
	Vessel revenue ($)	Non-Respondents Respondents	540,170 (442,485) 357,459 (101,293)	0.17
Gillnet-Large	Vessel length (feet)	Non-Respondents Respondents	44 (3) 45 (6)	0.45
	Vessel age (years)	Non-Respondents Respondents	21 (11) 23 (9)	0.22
	Gross tons	Non-Respondents Respondents	23 (12) 26 (16)	0.34
	Vessel horsepower	Non-Respondents Respondents	417 (184.2) 406 (130.9)	0.70
	Vessel revenue ($)	Non-Respondents Respondents	173,498 (139,056) 225,920 (210,023)	0.17
Gillnet-Small	Vessel length (feet)	Non-Respondents Respondents	35 (5) 35 (4)	0.94
	Vessel age (years)	Non-Respondents Respondents	27 (15) 27 911)	0.86
	Gross tons	Non-Respondents Respondents	13 (7) 13 (6)	0.67
	Vessel horsepower	Non-Respondents Respondents	288 (266.2) 318 (83.7)	0.13
	Vessel revenue ($)	Non-Respondents Respondents	103,077 (80,251) 74,502 (88,858)	0.17
Handgear-Large	Vessel length (feet)	Non-Respondents Respondents	54 (15) 46 (6)	***0.00***
	Vessel age (years)	Non-Respondents Respondents	27 (14) 31 (17)	0.41
	Gross tons	Non-Respondents Respondents	42 (25) 29 (13)	***0.01***
	Vessel horsepower	Non-Respondents Respondents	815 (471.7) 692 (320.9)	0.26
	Vessel revenue ($)	Non-Respondents Respondents	18,197 (49,571) 119,80 (11,592)	0.26

Handgear-Small	Vessel length (feet)	Non-Respondents Respondents	30 (5) 28 (5)	***0.00***
	Vessel age (years)	Non-Respondents Respondents	22 (12) 22 (11)	0.63
	Gross tons	Non-Respondents Respondents	10 (7) 8 (6)	0.08
	Vessel horsepower	Non-Respondents Respondents	349 (218.7) 302 (155.8)	***0.02***
	Vessel revenue ($)	Non-Respondents Respondents	9,728 (23,417) 11,779 (21,027)	0.44
Longline	Vessel length (feet)	Non-Respondents Respondents	44 (16) 49 (19)	0.56
	Vessel age (years)	Non-Respondents Respondents	26 (11) 20 (12)	0.24
	Gross tons	Non-Respondents Respondents	37 (37) 39 (43)	0.93
	Vessel horsepower	Non-Respondents Respondents	376 (195.8) 425 (114.7)	0.34
	Vessel revenue ($)	Non-Respondents Respondents	273,787 (372,942) 181,383 (199,302)	0.32
Pot/Trap-Large	Vessel length (feet)	Non-Respondents Respondents	44 (9) 43 (8)	0.39
	Vessel age (years)	Non-Respondents Respondents	18 (11) 17 (10)	0.09
	Gross tons	Non-Respondents Respondents	28 (26) 28 (22)	0.91
	Vessel horsepower	Non-Respondents Respondents	474 (181.9) 477 (168.7)	0.87
	Vessel revenue ($)	Non-Respondents Respondents	188,474 (213,324) 175,082 (143,921)	0.36
Pot/Trap-Small	Vessel length (feet)	Non-Respondents Respondents	34 (4) 34 (4)	0.07
	Vessel age (years)	Non-Respondents Respondents	21 (11) 21 (11)	0.79
	Gross tons	Non-Respondents Respondents	13 (6) 14 (6)	0.48
	Vessel horsepower	Non-Respondents Respondents	308 (123.6) 300 (120.9)	0.39
	Vessel revenue ($)	Non-Respondents Respondents	98,287 (100,736) 105,248 (102,781)	0.37
Purse/Seine	Vessel length (feet)	Non-Respondents Respondents	63 (16) 58 (15)	0.65
	Vessel age (years)	Non-Respondents Respondents	26 (11) 19 (13)	0.39
	Gross tons	Non-Respondents Respondents	110 (74) 60 (61)	0.25
	Vessel horsepower	Non-Respondents Respondents	456 (145.1) 603 (240)	0.27
	Vessel revenue ($)	Non-Respondents Respondents	1,025,573 (671,380) 788,776 (605,612)	0.55

Trawl-Large	Vessel length (feet)	Non-Respondents Respondents	77 (14) 76 (11)	0.54
	Vessel age (years)	Non-Respondents Respondents	29 (9) 33 (10)	***0.02***
	Gross tons	Non-Respondents Respondents	128 (48) 128 (41)	0.97
	Vessel horsepower	Non-Respondents Respondents	661 (434.1) 669.6 (395.7)	0.90
	Vessel revenue ($)	Non-Respondents Respondents	719,081 (584,272) 725,798 (649,955)	0.95
Trawl-Small	Vessel length (feet)	Non-Respondents Respondents	45 (8) 47 (7)	0.13
	Vessel age (years)	Non-Respondents Respondents	29 (13) 31 (17)	0.39
	Gross tons	Non-Respondents Respondents	30 (18) 32 (14)	0.39
	Vessel horsepower	Non-Respondents Respondents	335(115.6) 319 (109.2)	0.42
	Vessel revenue ($)	Non-Respondents Respondents	148,066 (194,263) 174,376 (170,208)	0.39