Two Sources of Nonsampling Error in Fishing Surveys

J. Michael Brick, William R. Andrews, and John Foster

Abstract Two important nonsampling errors arise from nonresponse and noncoverage. Both nonresponse and noncoverage are forms of missing data and are potentially important contributors to the accuracy of survey estimates. This research examines the potential effects of both nonresponse and noncoverage in the Fishing Effort Survey (FES), a survey conducted by the National Oceanic and Atmospheric Administration (NOAA). The difficulty in evaluating nonresponse and noncoverage is that the values for the missing data are not available and proxy measures must be used. Using proxies we find that noncoverage results in very large biases, while the magnitude of nonresponse bias is negligible in comparison. Another important finding is that the rates of missing data due to nonresponse or noncoverage are not predictive of the magnitude of the biases.

1 Introduction

All sample surveys are subject to sampling and nonsampling errors which cause survey estimates to deviate from true population values. Survey sampling texts (e.g., Cochran, 1977 and Lohr, 2019) describe survey design techniques to lower sampling errors in a cost-effective manner, but these texts provide less guidance on reducing nonsampling errors. This research investigates two important sources of nonsampling errors that are forms of missing data: nonresponse and noncoverage. The effects of nonresponse and noncoverage errors in a survey used to estimate fishing effort are examined both separately and jointly. The joint analysis provides

J. M. Brick (✉)
Westat, Rockville, MD, USA
e-mail: mikebrick@westat.com

W. R. Andrews · J. Foster
NOAA Fisheries, Silver Spring, MD, USA
e-mail: rob.andrews@noaa.gov; john.foster@noaa.gov

insight into the relative magnitudes of the potential error and can help prioritize efforts to improve the overall quality of the survey.

Nonresponse affects virtually all surveys and has been the subject of many articles and texts (e.g., Groves & Couper, 1998; Särndal & Lundström, 2005, and Stoop, 2005). Falling response rates (Atrostic et al., 2001; Williams & Brick, 2017, and Luiten et al., 2020) both in the United States and internationally have greatly heightened interest in nonresponse. For example, Stedman et al. (2019) question whether low-response rates imply surveys are no longer valid research vehicles, and Groves (2006) discusses the value of probability samples with low-response rates.

The research on noncoverage error, another form of missing data, is more diffuse because the source of noncoverage differs across surveys. A more comprehensive review is done by Lessler and Kalsbeek (1992), who discuss nonresponse, noncoverage, and other nonsampling errors such as measurement error. Narrower research on specific instances of noncoverage is illustrated by the work on telephone surveys in the 1980s and 1990s (Thornberry & Massey, 1988), and then again when mobile devices were first introduced (Tucker et al., 2007). Similarly, web surveying has spawned research on noncoverage related to Internet access (Scherpenzeel & Bethlehem, 2010).

Nonresponse and noncoverage both result in missing data. As a result, the two sources of nonsampling error can be investigated, at least theoretically, using the same structure. For example, consider the bias in a sample estimate of the mean, $\bar{y}_{nm} = \sum_{k \in s_{nm}} d_k y_k / \sum_{k \in s_{nm}} d_k$, where d_k is the inverse of the probability of selection for unit k and s_{nm} is the set of non-missing data where the missingness is due to either nonresponse or noncoverage. The bias can be written as

$$Bias(\bar{y}_{nm}) \approx M(\bar{Y}_{nm} - \bar{Y}_m), \tag{1}$$

where M is the percent of missing data, $\bar{Y}_{nm}$ is the mean of the (possibly hypothetical) stratum of those who would provide data, and $\bar{Y}_m$ is the mean of the stratum of those who would not provide data.

Another popular way of characterizing nonresponse is as a function of the correlation between the probability of a sampled unit responding (its response propensity) and the outcome variable (Bethlehem, 1988). This stochastic representation is

$$Bias(\bar{y}_{nm}) \approx \bar{\phi}^{-1} \sigma_\phi \sigma_y \rho_{\phi y}, \tag{2}$$

where $\bar{\phi}$ is the population propensity of providing data, σ_ϕ and σ_y are the standard deviations of the propensity and y variable , and $\rho_{\phi,y}$ is the correlation between the propensity and y. This model is not typically used for noncoverage because coverage propensities—the probability a unit is on the sampling frame—is more difficult to postulate as being random.

For most surveys, either nonresponse or noncoverage errors are examined, but they are rarely examined jointly. One exception is Couper et al. (2007) who do deal with both sources. More theoretical work, like that of Little and Rubin (2019), treat

nonresponse and noncoverage as forms of missing data, but do not delve into the implications of the magnitudes of the biases from each source.

Our objective is to examine the potential effects of both nonresponse and noncoverage in a particular survey to better understand the potential implications of each source. The findings help to better understand the optimal approach to managing resources in this survey to improve the accuracy of the estimates. More generally, the development of bias estimates due to each source of missing data for the same survey will help illuminate the nature of both nonresponse and noncoverage and suggest improved ways of thinking about these nonsampling errors for other surveys.

2 The Fishing Effort Survey

The survey that is the focus of our research is the Fishing Effort Survey (FES) conducted by the National Oceanic and Atmospheric Administration (NOAA). The FES is part of NOAA's Marine Recreational Information Program (MRIP), which produces estimates of recreational saltwater fishing catch—estimates that are used in managing fisheries. It is a cross-sectional, household survey that is conducted every 2 months. The key estimates are the total number of private boat and shore-based recreational, saltwater fishing trips taken by residents of coastal states.

The FES is an address-based sample (ABS) where the addresses are stratified into coastal and non-coastal sub-state regions defined by geographic proximity to the coast. Within each geographic strata, addresses are matched to the National Saltwater Angler Registry (NSAR), which is comprised of state lists of licensed saltwater anglers. This matching creates two additional strata: license-matched (households with one or more licensed anglers) and license-unmatched (households that cannot be matched to NSAR). The coastal and license strata were instituted to improve the efficiency of the sample. Within each stratum, addresses are selected in a single stage using simple random sampling. Weights include raking to household control totals derived from the American Community Survey (ACS) and a final poststratification adjustment to the number of households by coastal/non-coastal strata.

The state effort estimates of the number of shore fishing trips and boat fishing trips from the FES are then combined with independent estimates of average catch per trip from the Access Point Angler Intercept Survey (APAIS) to produce estimates of total recreational saltwater catch. Since the FES only samples people who reside in the state, an adjustment is made to the FES estimates to account for the noncoverage of nonresident anglers. Details on the survey protocol and results for 2020 are in the FES annual report (https://media.fisheries.noaa.gov/2021-08/MRIP-Fishing-Effort-Survey-2020-Annual-Report-V2.pdf). More details on estimation methods are given in Papacostas and Foster (2018).

For this research, we focus on the FES conducted in Waves 4 and 5 of 2020 (July–August and September–October time-periods) for four states where a nonresponse

Table 1 Sample sizes, number of completes, and response rates for 2020 standard Fishing Effort Survey and nonresponse follow-up surveys, by state

State	Standard survey			Nonresponse follow-up			Overall
	Sampled	Completed[a]	RR2[b]	Sampled	Completed[a]	RR2[b]	RR2[a]
Overall	28,650	7968	27.9%	15,993	3456	21.9%	42.4%
Florida	4222	1238	28.0%	2235	513	22.4%	42.6%
Massachusetts	6143	2010	31.5%	3160	774	24.3%	47.0%
New York	11,956	2714	26.1%	7253	1343	19.0%	39.8%
North Carolina	6329	2006	28.4%	3345	826	23.1%	43.7%

[a] Includes partial completes with some data that requires editing
[b] RR2 is the American Association of Public Opinion Research response rate 2

follow-up (NRFU) was conducted. The follow-up data are used in the analysis of potential nonresponse bias. The NRFU followed a subsample of the nonrespondents to the standard FES. Details of the NRFU data collection protocol are given in Andrews (2021).

The first columns of Table 1 show the sample sizes, number of completes, and response rates for the standard FES, where the data are aggregated over both waves in each state. Completed surveys include those where all the information requested is 100% reported plus partial completes which include missing or inconsistent information that can be resolved by editing or imputation. About 85–90% of completes require no editing. The overall response rate for the standard survey in these states during the two waves was 27.9%. The subsequent columns show the same information for the NRFU study. The NRFU response rate is based on the nonrespondents sampled for the NRFU. The overall response rates in the last columns combine the standard and NRFU data collection and are weighted to account for the subsampling in the NRFU. The overall response is 42.4%, computed using the AAPOR RR2 formula.

Table 2 shows the percent of households that took a fishing trip (either of any type of fishing or by boat or from the shore) and the mean number of trips of those that did fish by state, geographic area, and license status. We refer to the licensed-matched stratum as "Licensed" and the remainder are "Not licensed." These estimates are the standard estimates that do not include the NRFU effort. The table demonstrates the considerable variation in fishing by state, area, and license status.

3   Methods of Assessing Bias

The primary problem facing evaluations of nonresponse and noncoverage in surveys is that the values for the missing data are not available except in unusual circumstances. As a result, proxy measures of bias must be substituted for the missing values so that estimates of the effects of the missing data can be computed.

Table 2 Estimated percent of households that fished and mean number of trips, by state and stratum

States	Percent			Mean trips	
	Any fish	Boat fish	Shore fish	Boat	Shore
All four states	11.3	6.9	8.4	6.9	8.6
Coastal county	13.1	8.8	10.4	7.0	8.8
Non-coastal	2.3	1.5	2.4	5.0	5.6
Licensed	43.7	28.7	32.9	9.9	10.0
Not licensed	8.6	5.1	6.3	5.5	8.0
Florida	17.5	10.8	13.1	7.0	9.3
Coastal county	17.5	10.8	13.1	7.0	9.3
Non-coastal	–	–	–	–	–
Licensed	53.6	38.2	39.4	10.4	10.7
Not licensed	13.3	7.6	10.0	5.0	9.3
Massachusetts	8.6	5.3	6.0	6.5	6.2
Coastal county	10.1	6.4	6.8	6.7	6.5
Non-coastal	4.1	1.9	3.7	4.5	4.3
Licensed	58.6	32.8	45.1	10.4	8.6
Not licensed	6.3	4.0	4.2	5.0	4.9
New York	7.1	4.7	4.9	6.7	9.1
Coastal county	10.4	6.9	7.1	6.8	9.7
Non-coastal	1.7	1.0	1.2	5.1	3.0
Licensed	28.9	18.1	21.7	8.8	11.5
Not licensed	6.4	4.2	4.3	6.3	8.6
North Carolina	8.5	4.4	7.1	7.3	6.9
Coastal county	13.8	7.4	11.8	8.1	6.8
Non-coastal	4.4	2.0	3.5	5.0	7.0
Licensed	31.7	18.0	25.3	8.8	8.3
Not licensed	5.2	2.4	4.5	5.8	5.7

The proxy measures used in this analysis are discussed below for both nonresponse and noncoverage.

3.1 Nonresponse Bias

We estimate nonresponse bias in two ways. First, we compare the estimates from the standard survey to the estimates from the data including the standard and NRFU respondents. The difference is a proxy for potential nonresponse bias. This bias proxy assumes the combined data set is unbiased. Although this assumption is unlikely to hold with an overall response rate of 42.4%, Groves and Peytcheva (2008) found this method tends to produce larger estimates of nonresponse bias than other methods.

Table 3 Estimated nonresponse biases of percent who fished, by state and stratum

States	Percent any fish		Percent boat fish		Percent shore fish	
	NRFU	Early/late	NRFU	Early/late	NRFU	Early/late
All four states	0.1	0.1	0.0	0.2	0.1	0.0
Coastal county	0.2	0.2	0.1	0.2	0.2	0.0
Non-coastal	−0.6	0.2	−0.1	0.2	−0.6	0.0
Licensed	0.0	1.2	0.0	1.3	0.2	0.2
Not licensed	0.1	0.0	0.1	0.1	0.0	0.0
Florida	0.3	0.1	0.0	0.2	0.5	−0.3
Coastal county	0.3	0.1	0.0	0.2	0.5	−0.3
Non-coastal	–	–	–	–	–	–
Licensed	0.1	1.6	1.1	2.2	0.1	−0.1
Not licensed	0.3	0.0	−0.1	0.1	0.5	−0.2
Massachusetts	−0.4	0.7	0.1	0.4	−0.5	0.6
Coastal county	−0.4	0.8	0.1	0.5	−0.5	0.6
Non-coastal	−0.1	0.1	0.1	0.0	−0.2	0.3
Licensed	2.5	0.7	0.2	−0.1	2.4	0.5
Not licensed	−0.5	0.6	0.1	0.4	−0.6	0.5
New York	0.3	0.0	0.3	0.1	0.1	0.0
Coastal county	0.5	−0.1	0.3	0.2	0.1	0.0
Non-coastal	0.1	0.3	0.2	0.1	0.1	0.2
Licensed	−0.4	3.4	−1.2	2.1	−0.2	1.9
Not licensed	0.4	−0.2	0.3	0.1	0.1	−0.1
North Carolina	−0.7	−0.1	−0.4	−0.1	−0.6	−0.1
Coastal county	0.4	−0.3	−0.3	−0.4	0.9	−0.1
Non-coastal	−1.5	0.0	−0.6	0.2	−1.7	−0.2
Licensed	−0.5	−1.0	−1.1	−0.2	0.0	−0.9
Not licensed	−0.7	0.1	−0.4	0.1	−0.7	0.1

Another proxy is to compare estimates from early respondents (those who responded to the first mailing) to those of the combined early and late respondents in a level of effort analysis. The assumption that this difference gives an unbiased estimate of the bias is even less likely to hold than the NRFU assumption. Our primary goal for these estimates is to support the development of bounds on the potential bias in the next section.

Table 3 shows the two proxy estimates of nonresponse bias for three key estimates: the percent of households that did any fishing during the time period, the percent that fished from a boat, and the percent that fished from the shore. The NRFU nonresponse bias is computed using Eq. 1, which is equivalent to the difference between the standard estimate and the estimate that includes the NRFU respondents as well as the standard respondents. Both the standard and NRFU estimates went through the full set of adjustments except raking to the ACS.

The early/late bias estimate is the difference between the estimate based only on the early respondents to those based on all respondents to the standard protocol

(NRFU data are not included in the early/late estimates). The early estimate is computed as a domain of the fully weighted standard estimate rather than repeating the weighting steps for this domain.

The nonresponse bias estimates in the table are all relatively small. For example, the 30 bias estimates for the variable any fishing (the eight estimates for each of the four states except Florida which does not have any non-coastal areas) have a mean of 0.2% points. It is also interesting that 10 of the 30 bias estimates are negative because the primary bias concern for the FES is based on the hypothesis that the survey might be subject to topic interest bias (Groves et al., 2004). This hypothesis states that anglers would be more likely to respond to the survey than those who do not fish, so that the extra effort (e.g., the follow-up and more mailings) would increase the proportion of those who did not fish. Since 10 of the 30 estimates are negative (fewer anglers in responding sample) and the sizes of the estimates are relatively small, the NRFU data provide no evidence of substantial nonresponse bias due to topic interest.

3.2   Noncoverage Bias

For noncoverage in the FES, we simulate the effects of not including portions of the population because the ABS frame contains nearly 100% of all residential addresses (Battaglia et al., 2016). As a result, the estimates from the FES are subject to minimal noncoverage, except from nonresident anglers. Noncoverage scenarios are simulated by excluding (1) non-coastal addresses and (2) addresses that are not matched to the license register. These two types of restrictions of the sample have been researched as methods to improve the efficiency of the sample.

The noncoverage bias estimates are computed as the difference between the estimate restricted to either the coastal county or licensed-matched stratum (licensed) and the full sample estimate. The bias estimates were computed by using the full set of weighting procedures except raking, but using only the respondents from the "covered" stratum.

Table 4 shows the percent of fishing households that are excluded under the two scenarios. The magnitude of missing data is relatively small when coverage is restricted to the coastal stratum. In contrast, when the data are limited to the licensed households, the missing data rates are larger and roughly similar to those resulting from nonresponse.

Figure 1 shows both the estimated nonresponse and noncoverage biases for the three estimates of the percent of households that fished in each state and across the four states. The noncoverage bias estimates are labeled "Licensed" and "Coastal" to denote the covered part of the frame, while the nonresponse bias estimates are labeled "Early/late" and "NRFU" as described above. The noncoverage biases are all positive and generally large. Positive biases are expected because addresses in licensed addresses and coastal counties fish more often than those in the non-coastal and not licensed addresses. The magnitude of the noncoverage biases also vary substantially among states.

Table 4 Percent missing data, by source and state

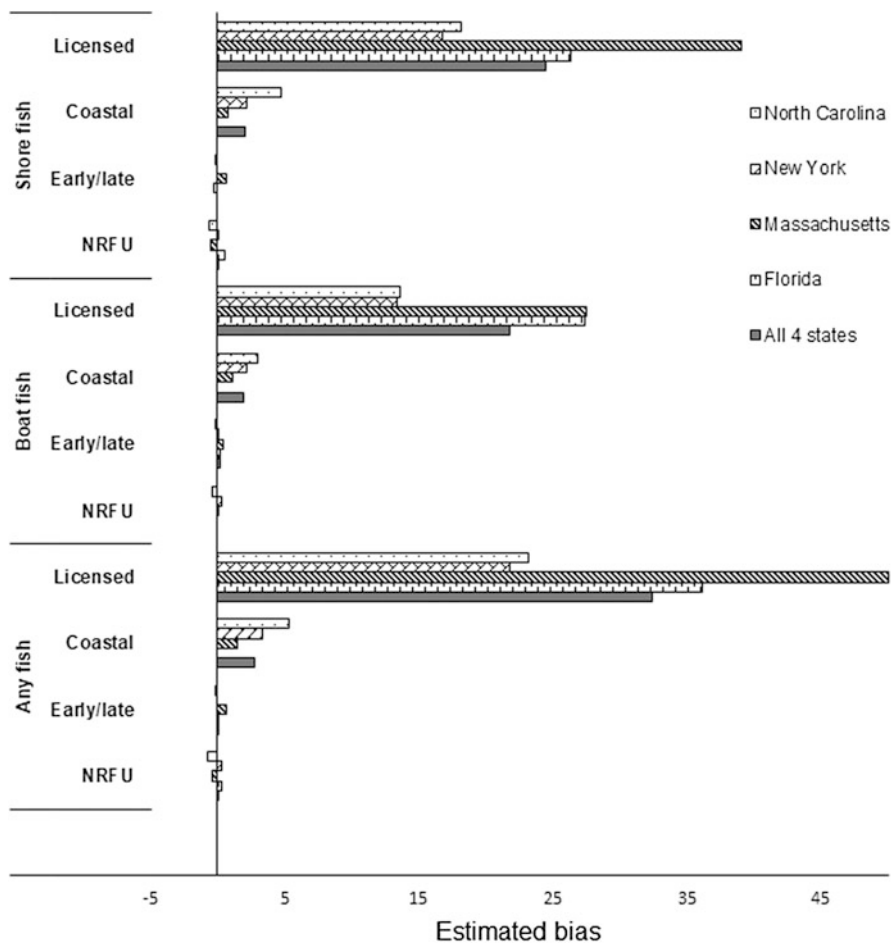| State	Nonresponse	Noncoverage	
		Licensed	Coastal
Overall	57.6	70.0	7.0
Florida	57.4	63.0	0.0
Massachusetts	53.0	70.0	11.0
New York	60.2	86.0	9.0
North Carolina	56.3	53.0	29.0

Fig. 1 Estimated nonresponse and noncoverage bias estimates for percent of households reporting any, boat, and shore fishing

The figure clearly shows the noncoverage biases are much larger than the nonresponse biases. The biases from excluding addresses without a license are especially large. For example, the overall noncoverage bias resulting from exclusion of non-coastal addresses is 2.8% points. When addresses without a license match are excluded, the estimated bias is 32.4% points. The corresponding nonresponse biases are less than 0.2% points.

Another key result is that the rate of missing data is a poor indicator of the potential bias from the different sources. For example, for the estimate of fishing prevalence (any fish), excluding non-coastal counties in North Carolina from the sample results in a bias that is more than 13 times higher than the nonresponse bias despite having a much lower rate of missing data. When we examine the joint effect of nonresponse and noncoverage, the dominant contribution of noncoverage is apparent. Table 3 shows that nonresponse bias is very small for all the strata used for simulating noncoverage bias (coastal and licensed). Because nonresponse bias was so small compared to noncoverage bias, we did not directly try to simulate the correlation between the two types of bias. If we assume the effects of the two sources of missing data are independent, an assumption of additive biases that probably overestimates bias, the nonresponse bias adds only slightly to the large positive biases due to noncoverage. At least in the FES, the effect of reducing coverage even to the coastal areas would swamp any nonresponse bias.

4 Bounds on Bias Estimates

The relatively low nonresponse bias estimates are not surprising. An earlier NRFU study done in 2012–2013 in the same four states also found no significant nonresponse bias. Other research also found that excluding non-coastal counties led to higher than desired noncoverage bias and that only including licensed addresses had substantial biases. As a result of these earlier findings, the decision was made to cover all addresses in the FES.

Despite these findings, the response rates in the FES are still low enough that nonresponse bias remains a concern (Stokes et al., 2021). When the fishing effort survey transitioned from a telephone survey to a mail survey, the estimates of the percent of adults fishing increased two- to threefold. The topic interest bias hypothesis was viewed as a realistic cause because the sampled households could see the whole survey immediately and understand the questions being asked.

It is important to understand the FES is subject to other nonsampling errors such as recall error (Andrews et al., 2018). However, recall and many of the other nonsampling errors would tend to reduce the proportion of estimated people who fished.

Since we are interested in looking at bounds, we construct bounds on the nonresponse bias with the topic interest hypothesis in mind. Early on Cochran (1977) discusses bounding the estimated nonresponse bias of a proportion and concluded the bounds could be "distressingly wide" if nonresponse was not negligible.

Following this approach but only allowing the bias to be positive, consistent with the topic bias hypothesis, we obtain the maximum possible bias in this direction. This value is derived by assuming that all fishing households in the sample responded (i.e., 100% response rate for fishing households) and that all nonrespondents were non-fishing households. For example, the FES NRFU had a 45.3% response rate across the four states, so we assume the remaining 54.7% did not fish. With this assumption, the maximum bias is 6.1% points for any fishing, 3.8% points for boat fishing, and 4.5% points for shore fishing. These are large biases relative to the size of the estimates given in Table 2, but assuming a 100% response rate for those who fish is extreme. Furthermore, even these extreme nonresponse assumptions result in nonresponse biases that are far smaller than the noncoverage biases in Fig. 1.

The bounding approach of Montaquila et al. (2008) can be modified to give more insight into the possible bias. Define the response propensities for those who fish (any fish, shore fish, or boat fish) to be ϕ_1 and for those who do not fish to be ϕ_2. Let P be the proportion who fish. The expected response rate is $\bar{\phi} = P\phi_1 + (1 - P)\phi_2$. Taking expectations of the estimated proportion ($\hat{p}$) over both the sample design and response mechanism, the bias of an estimated proportion is

$$Bias(\hat{p}) = P(\phi_1\bar{\phi}^{-1} - 1). \tag{3}$$

See Hedlin (2020) who derives the same expression.

These equalities are used to show how the response propensity or response rate in the any fish group (ϕ_1) and the bias is related. We take the NRFU as the basis for our values, with the NRFU overall response rate of $\bar{\phi} = 45.3\%$, and its estimate of the proportion who did any fishing of $\hat{p} = 0.112$. Plugging in these values, we obtain the ϕ_1 required to produce a bias of a specified amount shown in Fig. 2. The figure also shows shore fishing ($\hat{p} = 0.083$) and boat fishing ($\hat{p} = .0069$) curves. The estimate is unbiased if the responses rates are equal ($\phi_1 = \phi_2$).

The maximum positive nonresponse bias for each type of fishing is achieved when the response rate for those who fish (any fish, shore fish, or boat fish) is 100%. For example, for any fish, this bias is 6.1% points and is achieved when any fish $\phi_1 = 100\%$, as mentioned above. More reasonable values for ϕ_1 are some multiple of the response rate for those who do not fish. For example, bounds might be set by allowing ϕ_1 to be $1.2\phi_2$ (a response rate of 54.4%), $1.4\phi_2$ (63.4%), and $1.6\phi_2$ (72.5%). In practice, even a multiple of 1.2 is unusual and would imply a large topic interest effect.

Inspecting Fig. 2 at the points where the curves intersect these response rates shows the biases are relatively small. For example, for shore fishing, the biases are: 1.4, 2.4, and 3.1% points, respectively. For boat fishing, the biases are smaller.

The approach used by Hedlin (2020) bounds the potential nonresponse bias using Eq. 2 by considering different values of $\rho_{\phi,y}$. Hedlin shows that unless the correlation is high, the relative bias is small when the mean response propensity (response rate) is greater than 30%.

Applying this method, we assume the people who fish have a response propensity of ϕ_1, and the overall response propensity is $\bar{\phi}$. The correlation is
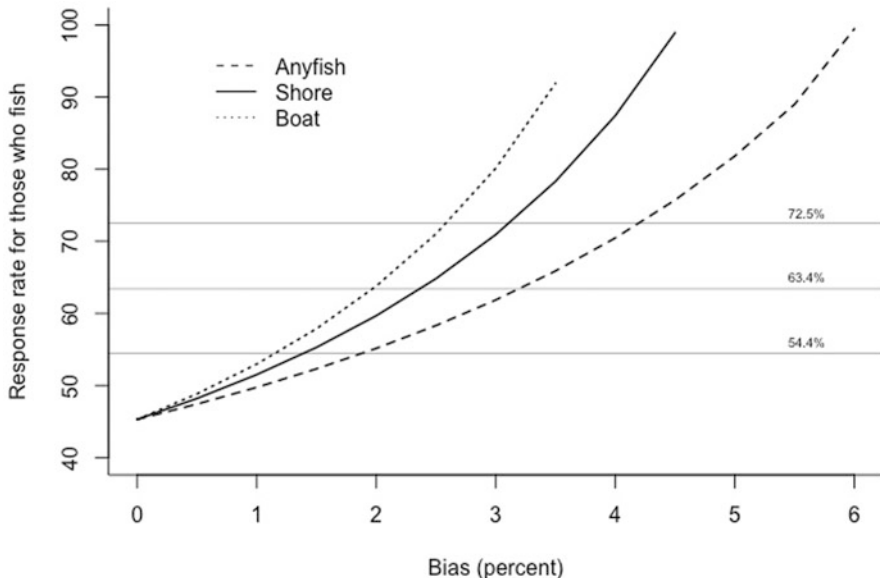
Fig. 2 Relationship between bias of estimate of any fishing, shore fishing, and boat fishing and response rate of those who fish, when overall response rate is 45.3% and estimated fishing proportions are 0.112, 0.083 and 0.069, respectively

$$\rho_{\phi,y} = \frac{P(\phi_1 - \bar{\phi})}{\sqrt{\bar{\phi}(1 - \bar{\phi})P(1 - P)}} \tag{4}$$

If we assume the extreme bound of $\phi_1 = 100\%$ and substitute for the other quantities in Eq. 4 using the NRFU values from above, the correlation for any fishing is 0.39, for shore fishing is 0.33, and for boat fishing is 0.30. With $\phi_1 = 1.6\phi_2$ (72.5%), the correlations are less than 0.20 for all three statistics. In other words, the correlations required to produce very large nonresponse biases are consistent with very extreme response rate assumptions for the fishing households.

Throughout this evaluation, the sample and weighting methods used to reduce potential nonresponse have not been taken into consideration. For example, the stratification by coastal geography and by license-match status have proven to be very effective in terms of identifying addresses with higher proportions of fishing as shown in Table 2. If the above analysis were repeated within stratum rather than overall, the within-stratum homogeneity of the fishing proportions and response rates would result in even less potential nonresponse bias.

For noncoverage, the bias estimates given earlier are those that we would obtain if no special weighting adjustments were used to reduce the bias. As a result, we discuss the potential to reduce noncoverage biases by weighting. This serves two purposes. First, it provides an alternative way to judge the size of the noncoverage bias if the frame only included coastal addresses or license addresses. Second, it

provides another angle on our research goal of exploring the relative magnitude of the nonresponse and noncoverage biases under conditions more favorable to noncoverage.

For the exclusion of addresses in non-coastal areas, an adjustment like that used for nonresident anglers could be employed using data from the APAIS. However, adjusting to a relatively small sample, such as the APAIS, is less efficient than using totals from a census or a very large, high-response rate survey like the ACS. Furthermore, although calibration generally reduces biases for estimates of totals, it is much less effective for estimates of proportions such as the proportion who fish. To be effective for adjusting estimates of proportions such as those studied here, the calibration data would need to be broken into classes or cells with differential coverage rates. Small surveys do not have adequate sample size to provide accurate estimates by classes. The exclusion of those in addresses that do not match to a fishing license is even more problematic. Asking license status in an in-person intercept survey such as the APAIS is fraught with problems since fishing without a license is illegal in many cases. As shown by Tourangeau and Yan (2007), this is precisely the situation in which large biases are common.

5   Discussion

Our analysis explores both nonresponse and noncoverage biases for the FES. The nonresponse bias analysis was feasible largely due to a nonresponse follow-up study. Noncoverage biases were estimated by artificially excluding data that were collected, where including only coastal areas or including only addresses with fishing licenses are designs that have been examined in practice because they are efficient in terms of finding anglers to complete the survey.

The nonresponse bias for the FES is relatively small except under very unexpected assumptions. The 2020 NRFU study and the analysis of the early and late respondents find only small biases. This finding is consistent with an earlier NRFU study. Both of those studies found no evidence to support the topic interest bias hypothesis that would result in overestimates of fishing prevalence.

When bounds based on different assumptions about the response rates for the people who fished and those who did not fish were constructed, nonresponse bias remains small under reasonable assumptions. Substantial nonresponse bias occurs only under the most extreme and unrealistic assumptions (such as assuming everyone who fished responded to the survey). We also translate these response rate assumptions into correlations between fishing and responding, showing again the nonresponse biases are small under realistic assumptions. For the four states in this study, the nonresponse bias for both shore fishing and boat fishing is likely to be no greater than 1–2% points even under relatively unusual assumptions (ratios of response rates of 1.2). This finding contrasts with the very large biases associated

with noncoverage in the FES. Large biases occur when the sample is restricted to either just coastal addresses or to addresses with licenses. Weighting methods to reduce these biases are feasible, but available external data sources are unlikely to reduce the noncoverage biases to be close to the magnitude of the nonresponse biases.

An important conclusion is that all missing data are not equivalent. In the FES at least, data that are missing because of noncoverage result in much larger biases than data that are missing due to nonresponse. The noncoverage biases are large even though the missing data rate for the coastal estimates average 7% and go up to 30% in North Carolina. The license noncoverage biases are much larger than even those in the coastal stratum. Despite missing data due to nonresponse being over 50%, the nonresponse bias is small. Clearly, missing data rates are not predictive of biases.

The important determinant of bias is the difference in the characteristics of the missing and non-missing data. This concept is simple to understand for noncoverage. If the percent of people who fish is very different for the covered and non-covered, then the bias will be large and weighting adjustments are unlikely to reduce the biases significantly. The bias estimates for the noncoverage due to sampling only coastal areas and licensed-matched addresses were very large for the percent who fished.

While the post-survey distinction of a respondent stratum and nonrespondent stratum has value in formulating a bias expression like Eq. 1 for nonresponse, it is a model that has limited conceptual appeal. If all sampled units have some chance of responding, then a nonrespondent stratum does not exist. Instead, the response propensity model given by Eq. 2 is more consistent with data collection experiences. For example, multiple attempts are made to interview the same units because the decision to participate is not fixed and may depend on a host of factors.

Another difference is that weighting adjustments for noncoverage rely exclusively on external data, but nonresponse adjustments can use data collected in the survey itself as well as external data. Typical nonresponse weighting class adjustments transfer weights from the nonrespondents to the respondents in the same class before calibration to external data. The survey data allow examination of the homogeneity of the response propensities within the classes. Noncoverage weighting adjustments also typically use classes to reduce bias, but the model for the adjustment cannot be evaluated from the survey data itself. Thus, weighting adjustments may be more effective for nonresponse.

These findings suggest that noncoverage may result in larger biases than nonresponse, even when the missing data rates due to noncoverage are much lower than those due to nonresponse. Surveys need to consider more than just missing data rates when deciding on survey designs and levels of effort. With the FES, saving resources by reducing coverage and using those resources to increase the response rate to the survey would likely increase the biases of the estimates. Each survey needs to evaluate its potential for biases, but the FES results reveal that relying on missing data rates to do this may be very misleading.

References

Andrews, W. R. (2021). *Evaluating nonresponse bias in the MRIP fishing effort survey.* https://apps-st.fisheries.noaa.gov/pims/main/public?method=DOWNLOAD_FR_DATA& record_id=2018

Andrews, W. R., Papacostas, K. J., & Foster, J. (2018). A comparison of recall error in recreational fisheries surveys with one- and two-month reference periods. *North American Journal of Fisheries Management, 38*, 1284–1298.

Atrostic, B., Bates, N., & Silberstein, A. (2001). Nonresponse in US government household surveys: Consistent measures, recent trends, and new insights. *Journal of Official Statistics, 17*, 209–226.

Battaglia, M. P., Dillman, D. A., Frankel, M. R., Harter, R., Buskirk, T. D., McPhee, C. B., DeMatteis, J. M., & Yancey, T. (2016) Sampling, data collection, and weighting procedures for address-based sample surveys. *Journal of Survey Statistics and Methodology, 4*, 476–500.

Bethlehem, J. (1988). Reduction of nonresponse bias through regression estimation. *Journal of Official Statistics, 4*, 251–260.

Cochran, W. (1977). *Sampling techniques*. Wiley.

Couper, M., Kapteyn, A., Schonlau, M., & Winter, J. (2007). Noncoverage and nonresponse in an internet survey. *Social Science Research, 36*, 131–148.

Groves, R. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly, 70*, 646–675.

Groves, R., & Couper, M. (1998). *Nonresponse in household interview surveys*. Wiley.

Groves, R., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly, 72*, 167–189.

Groves, R., Presser, S., & Dipko, S. (2004). The role of topic interest in survey participation decisions. *Public Opinion Quarterly, 68*, 2–31.

Hedlin, D. (2020). Is there a 'safe area' where the nonresponse rate has only a modest effect on bias despite non-ignorable nonresponse. *International Statistical Review, 88*, 642–657.

Lessler, J., & Kalsbeek, W. (1992). *Nonsampling error in surveys*. Wiley.

Little, R., & Rubin, D. (2019) *Statistical analysis with missing data*. Wiley.

Lohr, S. (2019). *Sampling: Design and analysis*. Chapman and Hall/CRC.

Luiten, A., Hox, J., & de Leeuw, E. (2020). Survey nonresponse trends and fieldwork effort in the 21st century: Results of an international study across countries and surveys. *Journal of Official Statistics, 36*, 469–487.

Montaquila, J., Brick, J. M., Hagedorn, M., Kennedy, C., & Keeter, S. (2008). Aspects of nonresponse bias in RDD telephone surveys. In: L. Tucker, B. de Leeuw, J. Lavrakas, L. Sangser (Eds.), *Advances in telephone survey methodology* (pp. 561–586). Wiley.

Papacostas, K., & Foster, J. (2018). National Marine Fisheries Service's Marine Recreational Information Program survey design and statistical methods for estimation of recreational fisheries catch and effort. https://www.fisheries.noaa.gov/resource/document/survey-design-and-statistical-methods-estimation-recreational-fisheries-catch-and

Särndal, C. E., & Lundström, S. (2005). *Estimation in surveys with nonresponse*. Wiley.

Scherpenzeel, A., & Bethlehem, J. (2010). How representative are online panels? Problems of coverage and selection and possible solutions. In *Social and behavioral research and the internet* (pp. 105–132).

Stedman, R. C., Connelly, N. A., Heberlein, T. A., Decker, D. J., & Allred, S. B. (2019). The end of the (research) world as we know it? Understanding and coping with declining response rates to mail surveys. *Society & Natural Resources, 32*, 1139–1154.

Stokes, S. L., Williams, B. M., McShane, R. P., & Zalsha, S. (2021). The impact of nonsampling errors on estimators of catch from electronic reporting systems. *Journal of Survey Statistics and Methodology, 9*, 159–184.

Stoop, I. (2005). *The hunt for the last respondent*. Sociaal en Cultureel Planbureau.

Thornberry, O., & Massey, J. (1988). Trends in United States telephone coverage across time and subgroups. In: R. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls, J. Waksberg (Eds.), *Telephone survey methodology* (pp. 25–49). Wiley.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133*, 859–883.

Tucker, C, Brick, J. M., & Meekins, B. (2007). Household telephone service and usage patterns in the United States in 2004: Implications for telephone samples. *Public Opinion Quarterly, 71*, 3–22.

Williams, D., & Brick, J. M. (2017). Trends in US face-to-face household survey nonresponse and level of effort. *Journal of Survey Statistics and Methodology, 6*, 186–211