Effects of survey response mode, purported topic, and incentives on response rates in human dimensions of fisheries and wildlife research

Leif Anderson,[1] Matt Jans,[2] Adam Lee,[2] Christopher Doyle,[2] Heather Driscoll,[2] and James Hilger[3]

[1] *Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration*

[2] *ICF*

[3] *Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration*

Acknowledgements

We are very grateful to all anglers who took part in this study and to the Washington, Oregon, and California Departments of Fish and Wildlife for their assistance in drawing the samples. The authors acknowledge funding from NOAA Fisheries Science and Technology to conduct the pilot test experiment described here, as well as the overall economic survey. This article and its findings do not necessarily reflect the views of the National Oceanic and Atmospheric Administration (NOAA) or the National Marine Fisheries Service.

Abstract

Incorporating human values and behavior into natural resource management has been linked to more sustainable and efficient outcomes. Estimating these values often requires survey-based data collection. Declining response rates have amplified the need to more fully understand the influence of survey design features on response. In this study, we used an experiment to measure the effect of response mode (mail vs. web), purported topic (recreation vs. saltwater angling), and incentives ($0, $2, $5) on two outcomes of interest: screener response and extended questionnaire response. The context is a survey of saltwater anglers on the Pacific coast of the United States. We found increased probabilities of screener response with an incentive, with a stronger effect within the mail response mode. At the extended questionnaire stage, we found some effect of survey response mode that was moderated by topic and incentive. Our results help inform future survey efforts.

Keywords: survey; response rates; nonresponse; recreational fishing; financial incentive; survey framing; survey mode

Introduction

There is an increasing recognition that incorporating human values and behavior into the process of natural resource management can lead to better outcomes with respect to both sustainability and efficiency (Bunnefeld et al., 2011; Kellner et al., 2011; Levin & Anderson, 2016; Müller et al., 2011). Characterizing the bidirectional links between human behavior and changes in the quantity or quality of natural resources can be especially important in the context of extractive or consumptive uses of fish and wildlife, with direct implications for the mortality and long-term sizes of populations with significant ecological or economic importance (Abbott & Fenichel, 2013; Baruch-Mordo et al., 2009; Lee et al., 2017; Teel & Manfredo, 2010; Ward et al., 2008).

Conducting studies examining these issues can be challenging, as existing data sources are often insufficient to characterize linkages between different levels of human behavior and the resource, requiring researchers to often turn to primary data collected from a survey. Human dimensions surveys are becoming increasingly expensive, with decreasing response rates (Stedman et al., 2019) and changing demographics among target populations (Arlinghaus, 2006). The effectiveness of survey recruitment methods is also changing, suggesting an increasing need to understand the most cost-effective ways for increasing response rates and the quality of data in contemporary survey research (National Research Council of the National Academies, 2013).

Our study explores the effects of different survey design factors in one particular human dimensions context – an economic survey of saltwater anglers on the Pacific coast of the United States (U.S.). This survey involved characterizing the preferences of anglers using a discrete choice experiment (Anderson et al., 2013; Hunt et al., 2020; Lew & Larson, 2014). As is common in survey research, several

factors of the administration were not fixed, allowing design choices to be determined partially based on the anticipated effects that these factors have on important metrics such as response rates.[1]

Although it has been shown that many factors of survey administration can have significant effects on response rates (Guo et al., 2016; Kanuk & Berenson, 1975), we focused on three: response mode, purported topic, and incentives. A number of articles have examined effects of monetary incentives on varying survey contexts, providing substantial evidence that incentives are likely to increase response (Edwards, 2005; Gajic et al., 2012; Groves et al., 2000; Jobber et al., 2004; Mercer et al., 2015; Millar & Dillman, 2011; Singer & Ye, 2013; Willcox et al., 2010). The effect of response mode has also been frequently shown to impact the propensity to respond, with response rates for mail surveys typically higher than those for internet-based surveys (Converse et al., 2008; Shannon & Bradshaw, 2002). Purported topic has been less frequently studied and is perhaps less likely to affect response rates (Tourangeau et al., 2009). However, purported topic is closely related to the more general notions of topic salience and framing effects, and more salient topics are often shown to produce higher response rates (Goyder, 1987; Groves et al., 2004; Groves et al., 2006; Marcus et al., 2007; Stedman et al., 2019). Importantly, nonresponse bias is more likely to result from contexts in which the purported topic has a significant effect on response rates (Groves et al., 2006).

[1] Response rates affect the cost per completed response and are regularly used as a proxy for data quality. Although low response rates are often interpreted as direct evidence of nonresponse bias, a more accurate interpretation is that of potential bias; response rates have been shown to have mixed effects on bias when investigated empirically (Groves, 2006; Kreuter, 2013). Investigations of nonresponse bias are beyond the scope of this article.

The efficacy of survey design factors has been shown to depend on contextual elements that vary over time and across demographic differences between target populations (Dillman et al., 2014; Groves et al., 2013), and no existing studies that we are aware of have examined the joint effects of response mode, purported topic, and incentives in the context of recreational angler surveys. We therefore turned to primary research to fill this gap, using an experimental approach where distinct combinations of our three factors (treatments) are applied to subsets of the sampled population. Using a controlled experiment allowed an examination of treatment effects on survey response while holding any additional confounding factors constant. In addition, this assessment of joint effects provided the framework to quantify any potential interactions between factors, which is necessary to produce a more nuanced set of insights (e.g., whether the effect of incentive depends on response mode). The direct contribution of this work to the literature is to provide contemporary, context-specific estimates of the joint effects of response mode, purported topic, and incentives on response rates within a single study. This allows us to answer three research questions: (a) do response mode, purported topic, and incentives affect screener completion propensity; (b) do response mode, purported topic, and incentives affect extended questionnaire propensity among anglers who are eligible; and (c) do the effects of response mode, purported topic, and incentives moderate each other (i.e., are interactions between these three design factors present)? Our results also serve to inform future survey efforts among this angling population and in other similar contexts, such as studies of freshwater anglers or hunters. Natural resource researchers and practitioners can use these effects to help inform the design and reduce the cost of future survey efforts.

Methods

Experimental Design

We used a randomized experiment to examine the effect of survey administration factors

(response mode, purported topic, incentives) on the propensity to complete an eligibility screener and an extended questionnaire about fishing practices and preferences. This randomized experiment was carried out using a survey of licensed anglers in the U.S. states of Washington, Oregon, and California.

The response mode factor of the experimental design consisted of two levels: mail and web (i.e., internet-based). The complexity of some sections of the questionnaire precluded a telephone mode from consideration. We defined purported topic as the subject that was used to frame the questionnaire and its purpose within the introductory survey materials. Two levels of the purported topic factor provided the framework to test whether framing the topic as (general) recreation versus (more specific) saltwater fishing in the introductory questionnaire materials that respondents received as part of the full protocol affected the propensity to respond. These two levels represented the actual survey topic of saltwater fishing and a more broadly-defined topic that includes fishing as well as other recreation activities.[2] Three levels were used for the incentive factor: $0, $2, and $5. Multiple non-zero levels provided the basis to help answer the related questions of whether an incentive should be used, as well as how large such an incentive should be, if used. We used a full 2x2x3 factorial design to allow for the testing of both main effects, as well as interactions among factors. Different combinations of these factor levels (treatments) were administered to randomly drawn subsamples to test whether there were statistically significant differences in the response rates among these treatments.

[2] The hypothesized effect of the recreation purported topic was unknown at the time of the experiment. Although it seems likely that a more general recreation topic might encourage response from the anglers who consider other recreation activities more important or did not fish in saltwater (defined as ineligible for our survey), it could have the opposite effect on more avid saltwater anglers.

Our sample was allocated as equally as possible to each treatment. Specifically, an overall sample of 4,000 was equally split between treatments with response modes of mail and web, between the topic framed as recreation and saltwater fishing, and almost equally among the three levels for the incentive factor ($0 = 1,344, $2 = 1,328, $5 = 1,328). Overall, the test of each treatment was conducted using either 332 or 336 license holders, depending on the level of the incentive factor (Table 1). The overall sample in each state was drawn randomly from licensing databases and sampled anglers were randomly assigned to treatments.[3]

Table 1 about here

Survey Administration

The target population for our study was all adult licensed anglers who had fished in saltwater off the coasts of Washington, Oregon, or California in the previous 12 months. The saltwater angling population could not be uniquely identified using only the information contained in the licensing databases due to the generality of many license categories across the three states. Instead, we used screening questions to determine eligibility.

[3] Out of 4,000 addresses, two were removed from the analysis because they requested a mode other than the one they were assigned (web respondents who requested a mail questionnaire). These two records were dropped from analysis as they were considered contamination with respect to the mode manipulation. Of the remaining 3,998, 858 (21%) responded to the screening questionnaire. In total, 224 addresses responded to the extended questionnaire. Fifty-five of those in the mail condition did not respond to the mailed screener questionnaire, and only provided their "screening information" with their extended questionnaire response. These cases are not counted as "screener complete" because they did not respond to that particular contact attempt and method.

Contextual differences in species targeting and catch composition across the three states and between Northern and Southern California were used to create four regional variants of the questionnaire. Sample sizes were split equally among these four regions for the experiment. As the underlying goal of the experiment was to select a single treatment to administer the final questionnaire across the four regions, the region variable was not interacted with the factors of the experimental design in the subsequent estimation.

Our overall study protocol consisted of an initial recruitment mailing and up to four follow-up contacts.[4] All anglers included in the sample received an invitation letter by mail, describing the study and requesting participation. The wording of our invitation letters varied by purported topic. The saltwater fishing version of our letter stated that the purpose of our study was to help understand and manage fish stocks, whereas the recreation version of our letter described the research more broadly about recreation activities. For approximately two-thirds of the sample, this initial contact included an incentive, split equally between $2 and $5. Anglers assigned to the mail return mode also received a two-page screener questionnaire with the invitation that was used to determine eligibility (Table 2).

Table 2 about here

Anglers assigned to the mail return mode who were determined to be eligible received a hardcopy of a 16-page questionnaire as the next contact. The web return mode sample received a personalized link within the invitation letter to a web questionnaire that included screener questions, followed by the full web questionnaire for anglers determined to be eligible.

[4] The study received human subjects review approval under the Paperwork Reduction Act: OMB control number: 0648-0750.

Differences in purported topic were carried through the cover page of the mail questionnaire. The graphics on the cover of the saltwater fishing topic questionnaire depicted only target saltwater species illustrations. The cover of the general recreation topic questionnaire included illustrations of fishing along with other common recreation activities.

In the follow-up contacts, the differences related to purported topic and response mode were maintained, as the wording of the reminder letters and the cover page of the mail questionnaire were framed as either saltwater fishing or general recreation activities, and the repeated invitations pushed potential respondents to either a mail or web based version on their assignment of mode in the first stage of the experiment.

Outcomes of Interest

We addressed two key outcomes: screener completion and extended questionnaire completion. Addressing these response steps sequentially allowed us to assess various points in the survey process where nonresponse can enter, and whether design factors impacted that risk. First, we explored the propensity to respond to the screening effort. Across both modes, the only eligibility screening question asked whether the sampled angler had taken any saltwater fishing trips within the previous 12 months. However, the screening processes were slightly different in each mode. For the mail response mode, screening questions could be completed either via the mailed screener form or through answers to the follow-up contacts on the extended mail questionnaire for anglers who did not respond initially to the screener form. In this analysis, addresses in the mail mode were only considered a “complete screener” if they returned the mail screener form. Addresses assigned to the web mode provided their screener information in the web questionnaire once they logged in and before the extended web questionnaire questions. This decision was made so that we could evaluate response to the first contact attempt and

method in each sample (i.e., screener form return in the mail sample and screener questions online in the web sample).

The second outcome of interest was response to the extended questionnaire. In the web sample, a screening question immediately preceded the extended questionnaire and therefore all web extended completes also completed the screener. For the mail mode, most of the sample completed the mailed screener questionnaire prior to receiving the extended mail questionnaire. However, as mentioned above, 55 extended mail completes only provided their eligibility information by answering an equivalent eligibility question on the extended questionnaire, so were not considered to be a completed screener for this analysis.[5] For both response modes, we defined an extended questionnaire as complete if a respondent included answers to at least one question that followed the screening questions.

Estimation of Effects

All outcomes of interest were discrete indicator variables using an individual angler as the unit of analysis. We estimated a set of logistic regression models to calculate the effects of response mode, purported topic, and incentive amount on screener completion and extended questionnaire completion. To characterize the potential robustness of the results, we provided separate estimates for models that included only main effects and models that included interactions among the design factors. In addition to the factors in the experimental design, we included the main effect of region in all models.

For our main effects models, both screener completion and extended questionnaire completion were estimated as a function of indicator variables describing differences from the omitted reference

[5] Thus, extended questionnaire completion should be interpreted as mostly, but not fully, contingent on the distinct screener effort.

levels of each factor: (a) mail (mode); (b) recreation (purported topic); (c) $2 and $5 (incentive); and (d) Northern California, Southern California, and Oregon (region). As an illustrative example, in the model for extended questionnaire completion, the coefficient on mail measures the difference in the propensity to complete the questionnaire for anglers assigned to the mail response mode relative to anglers assigned to the web response mode. The full models for screener completion and extended questionnaire completion added all two- and three-way interactions with the exception of region. We omitted interactions between region and design factors, as the goal of the experiment was to select a single survey design approach to implement across all four regions. We estimated all models in Stata 16.

To help quantify and communicate the effect of changes in the levels of each design factor on the predicted probability of response to both the screener and extended questionnaire, we calculated and then graphed the incremental effects using Stata's `margins` and `margins plot` functions, respectively. In logistic regression models, the estimated coefficients on the interaction terms alone cannot be used for determining the nature of any interaction effects on the predicted probabilities (Ai & Norton, 2003). Therefore, we estimated and plotted post-hoc marginal effects following a widely-accepted approach (Mitchell, 2020; Mize, 2019). All post-hoc contrasts were tested using the default balanced option in Stata because the experimental design led to equal or very close to equal sample sizes in each manipulated design feature. The full set of margins for screener and extended questionnaire response, broken out by the full factorial of the design factors, is available from the authors upon request for readers interested in further detail than is provided in this article.

Results

A total of 858 anglers responded to the screening questionnaire and 224 anglers responded to the extended questionnaire. Although it is also important to compare respondent demographics to the population to help gauge the representativeness of the data, this was not possible in our context due to

data limitations of the license frame. First, there was no way to know which anglers were eligible within the license frame (hence the need for a screener in our study). Second, there are almost no demographic data within any of the three state licensing databases.

Prior to presenting regression results, the bivariate associations reported in Figure 1 demonstrate that screener response (Figure 1a) was significantly associated with mode (Pearson χ^2 = 30.59, $p < .001$), and incentive (Pearson χ^2 = 143.57, $p < .001$), but not topic (Pearson χ^2 = 0.27, p = .601).

Figure 1 about here

Extended response, conditional on both screener response and eligibility (Figure 1b), was not associated with any of the three manipulated design features. However, we retained all design treatments in the screener and extended questionnaire regressions that follow to test for a Simpson's paradox (Kievit et al., 2013), and the potential that main effects and interactions emerged after controlling for region.

In addition to highlighting general differences in response propensities, we used these bivariate plots to help inform general model specification for the subsequent logistic regression analysis. In particular, note that the $2 incentive had a large impact on screener response propensity relative to no incentive, and that the incremental effect of a $5 incentive was relatively minor relative to a $2 incentive.[6] This suggests a potential simplification to the regression plans: collapsing the $2 and $5

[6] Although post-hoc tests from a model with only incentive predicting screener complete showed that each of the incentive levels were significantly different from each other (the difference between $2 and $5 was statistically significant (z = 2.11, p = .035), detailed results not shown), the largest difference was clearly between no incentive and any incentive.

levels of incentive. As this simplification provided a considerable reduction in complexity, especially in models that include interactions between multiple factors, we adopted a combined $2 and $5 incentive level in what follows.

Predicting Screener Completion

Regression estimates from the four logistic models provided a more comprehensive depiction of the effects of the design factors and region on response rates (Table 3). For screener completion, the main effects only model estimates (Table 3, 1a) show that mode, incentive, and region had significant effects. Using a mail response mode produced a statistically significant higher screener response propensity than using a web mode (Odds Ratio (OR) = 1.57; $p < .001$), and a $2 or $5 incentive produced a more substantial increase (OR = 3.12; $p < .001$). All regions exhibited lower completion rates than did Washington.

Table 3 about here

The screener completion interactions model estimates (Table 3, 1b) were qualitatively similar to the main effects only model estimates. Here again, all regions had lower completion rates relative to Washington. The characterization of effects for the three factors in our experimental design requires a more careful interpretation given the included interactions among factors in this model. For example, the effect of an incentive is a function of the positive main effect (OR = 1.65; $p = .008$) and interactions that characterize how the effect of an incentive varies as a function of the other factors. The contrasts of mode, incentive, and the cross-effect of mode by incentive were all significant (Table 4). This shows that increased response propensity was generated using a mail response (Wald $\chi 2 = 3.62$, $p = .057$)[7] or

[7] Both $\alpha = .05$ and $\alpha = .10$ levels were used to asses statistical significance due to the exploratory nature of the research.

including any incentive with the screener (Wald $\chi 2 = 124.74$, $p < .001$), and that these effects were increased further when combined (mail response, any incentive; Wald $\chi 2 = 19.34$, $p < .001$).

Table 4 about here

To further illustrate these higher order effects, we examined the interaction between mode and incentive on predicting screener completion (Figure 2). When no incentive was given, there was no effect of mode ($\Delta = -0.02$, Wald $\chi 2 = 2.05$, $p = .152$). Adding an incentive increased the probability of screener response in both the mail ($\Delta = 0.23$, Wald $\chi 2 = 187.79$, $p < .001$) and web modes ($\Delta = 0.09$, Wald $\chi 2 = 28.75$, $p < .001$). However, the effect was much larger for the mail mode. The differential between the mail and web mode response probabilities was greatest when any incentive was given ($\Delta = 0.12$, Wald $\chi 2 = 50.49$, $p < .001$). The highest level of response was achieved from respondents invited through the mail and asked to return a mail screener form. Taken together, any amount of incentive increased response, but the incentive had a stronger effect in mail than web mode. See Table 5 for first and second differences as tests of specific effects based on Figure 2.

Figure 2 and Table 5 about here

Predicting Extended Questionnaire Completion Given Screener Response and Eligibility

In the models predicting extended questionnaire completion (conditional on screener completion and eligibility), fewer design factors had significant effects. In the extended questionnaire main effects only model (Table 3, 2a), none of the three design factors had statistically significant effects. Among the regions, only Northern California showed an increased probability of extended questionnaire response in comparison to Washington (OR = 2.34; $p = .022$).

The extended questionnaire interactions model (Table 3, 2b) depicted a more nuanced set of effects with a combination of significant main effect and interaction coefficients. The contrasts of topic

by mode and topic by mode by incentive were both statistically significant (Table 4). This shows that the probability of response to the extended questionnaire was increased by framing the topic as recreation while also using the mail response mode (Wald $\chi 2 = 6.00$, $p = .014$) and that adding an incentive to this combination further increased response (Wald $\chi 2 = 4.93$, $p = .026$).

Table 4 about here

We further illustrated these higher-order effects by examining the highest significant interaction in the regression model (Figure 3), presenting the three-way interaction between topic, mode, and incentive as two two-way interactions between incentive and mode, paneled by topic. Partial interactions within each purported topic were tested for significance at the $\alpha = .10$ level because neither were significant at the $\alpha = .05$ level. This difference in testing was reflected using 90% confidence intervals (Figure 3). The interaction between incentive and mode – the difference in the effect of incentive ($0 - $2 or $5) evaluated at different levels of mode (mail - web)[8] – was significant among sampled anglers assigned to the recreation topic ($\Delta = 0.32$, Wald $\chi 2 = 3.48$, $p = .062$), but not among anglers assigned to the saltwater fishing topic ($\Delta = -0.24$, Wald $\chi 2 = 2.13$, $p = .145$). Specifically, when the topic was recreation and there was no incentive, mail mode lead to higher response than web ($z = 2.44$, $p = .015$). Table 6 shows tests of first and second differences for the predicted marginal propensities. Similar to screener nonresponse, we calculated first differences for both mode and incentive, and then second differences among those effects. We repeated that process separately for each topic. As discussed above, mail mode obtained a higher response propensity than web when the purported topic was recreation and there was no incentive. The second difference (i.e., difference in mode effects

[8] Given the model structure (Table 6), this is also equal to the difference in the effect of mode (mail - web) evaluated at different levels of incentive ($0 - $2 or $5).

between incentive levels) was also significant only for the recreation topic. Neither incentive nor mode influenced response propensity when the topic was saltwater fishing.

Figure 3 and Table 6 about here

Discussion

In this experiment, we evaluated the effects of response mode, purported topic, and incentives on response rates to a survey of saltwater anglers at the screener stage and at the extended questionnaire stage. Estimates from logistic regression models showed that factors of survey design have both statistical and practical significance, highlighting the importance of a thoughtful design process. At the screener stage, increased probabilities of response were generated by including any amount of incentive and the incentive had a stronger effect in mail response mode (relative to web). This increased effect of incentive for the mail screener was possible due to how closely the incentive and the response task were tied; mail mode respondents saw the incentive and could fill out the screener form immediately without the added step of accessing a web-enabled device. Without an incentive, there was no effect of mode.

There were fewer significant effects at the extended questionnaire stage for a few potential reasons. First, the design factors had already acted on respondent behavior, so further incremental effects on response are perhaps less likely.[9] Second, there is increased separation between some of the factors and the opportunity to respond to the extended questionnaire. In particular, the incentive accompanied earlier contacts. A significant positive effect was found for extended questionnaire

[9] This type of relationship has been noted in other studies. For example, Schmid et al. (2019) found that participation in an online questionnaire was increased with an incentive, but once participants were recruited, the net effect on completing the questionnaire was zero.

completes for the mail mode (relative to web), but only when the topic was recreation and there was no incentive.

It is important to note that the combined effect of any experimental factor, over both the screener and extended questionnaire stages, cannot be directly seen by examining only the regression results for the extended questionnaire response. To see this, note that the later stages of a multi-stage data collection process are often conditional on responses to earlier stages. In our context, respondents must first respond to the screener before transitioning to the extended questionnaire. If an experimental factor, such as incentive, has a strong positive impact on response at the screener stage and a null impact at the extended questionnaire stage, the combined effect remains positive. In this article, we provide the effects broken out by the stage of the data collection, as this provides more information about the particular channel by which nonresponse can enter the process.

In the specific context of identifying saltwater anglers from within a population of general fishing license holders, using a distinct screener was a necessary precursor to an extended questionnaire. In this two-stage process, survey design factors are likely to have relatively larger effects on the first stage (screener completion) relative to their effects on the second stage (extended questionnaire completion).[10] In contrast, many human dimensions surveys would not need a distinct screener to be administered, provided the frame used for sampling either allowed direct identification of the eligible population or the proportion of eligible individuals was sufficiently large. In survey research applications where no screener survey is needed, the results we provide at the screener completion stage are likely to be more relevant than the conditional extended questionnaire results.

[10] The results from this experiment offer support to this assumption, as the survey design factors had considerably more impact on the first-stage (screener) of the two-stage data collection.

Observed differences between the effects of survey design factors on mail and web mode responses are likely to be a product of the relative temporal proximity of the study invitation, screener survey, and extended questionnaire. Compared to web mode respondents, mail mode respondents received less time between the invitation and screener form, and a longer break between the screener and extended questionnaire. For example, we found that no design factors had an incremental impact on extended questionnaire completion for web mode respondents, given the completion of a screener by an eligible respondent. This is not surprising, as respondents to the web screener are immediately transitioned to the web-based extended questionnaire. Although the complexity of the overarching questionnaire – a discrete choice experiment – precluded the consideration of telephone as a response mode, it would be interesting to expand this experiment to include telephone response mode within the context of a simpler survey design.

Comparison to the Literature

Although there are many studies that have examined effects of particular survey design factors on response rates, there are few studies that explored the individual and combined effects of survey mode, purported topic, and incentives among a similar population. Empirical studies exploring these factors together in a single study within any population are also limited. To provide some context related to the generalizability of our results to other populations, we compared the results we found in this study of anglers with the general results found in the literature. We provided these comparisons broken out by survey design factor as follows.

Incentives

Comparing our results to the existing literature, we found agreement that pre-paid incentives boost response rates (Edwards, 2005; Gajic et al., 2012; Groves et al., 2000; Jobber et al., 2004; Millar & Dillman, 2011; Singer & Ye, 2013; Willcox et al., 2010), but this effect does not necessarily increase

linearly with the incentive amount (Edwards, 2005; Mercer et al., 2015; Singer & Ye, 2013). Findings from our experiment help refine these results for researchers planning multi-stage surveys by highlighting that the reduction in nonresponse from an incentive primarily impacts the first (in our context, screener) stage of the data collection.

Response Mode

Response rates from questionnaires administered by mail have been shown to be higher than those from questionnaires administered by the internet (Barrett et al., 2017; Converse et al., 2008; Manfreda et al., 2008; Shannon & Bradshaw, 2002). In the context of angler surveys, web surveys that were administered through e-mail only have been shown to have lower response rates than those administered by mail or telephone (Barrett et al., 2017). Even though we did not rely on e-mail for recruitment, as we used a mailed invitation letter in both the mail and web response mode treatments, higher response rates for the mail mode seem to confirm these results for the screener. In the extended questionnaire, the only significant effect of mail mode was for the recreation topic with no incentive. Overall, our experiment adds additional context to existing evidence that mail surveys are likely to achieve higher response rates than web surveys.

Purported Topic

Comparing our results to the literature requires broadening the scope beyond the framing of saltwater fishing and general recreation to other topics. Existing research has demonstrated that people with more involvement or interest in a topic have higher response rates (Goyder, 1987; Groves et al., 2004; Groves et al., 2006; Marcus et al., 2007; Stedman et al., 2019) and that differential framing of the survey topic within an invitation to participate in a survey, or framing of individual questions, can influence response (Schuldt et al., 2015; Tourangeau et al., 2009). In our context, the framing of the survey topic as saltwater fishing was likely to be relatively more salient to the eligible population, and

this is especially true of avid saltwater anglers. For other less avid saltwater anglers who also engage in other recreation activities, it is possible that the recreation topic was relatively more salient.[11] Whether the net of these opposing effects was an increase or decrease in response rate was an empirical question. However, there are some reasons to believe that, in our context, the potential effects of purported topic were likely to have been moderated relative to general population surveys with more sensitive topics (Tourangeau et al., 2010). First, there was overlap between the two levels of purported topic in our experiment, as the topic of saltwater fishing was wholly contained within the more general topic of recreation. Second, the sample frame consisted of license-holding anglers – albeit some of whom only fished in freshwater – and as a result, even the ineligible sample might have considered the saltwater fishing topic to be somewhat salient. Given this contextual limitation to our experiment, it is perhaps not surprising that we found no significant main effects of topic in either the screener or extended questionnaire. Existing literature suggests that the differential effect of interest is moderated with the use of incentives (Groves et al., 2004; Groves et al., 2000). In the extended questionnaire model, the two significant interactions exhibited a similar pattern, as the interaction effect of purported topic by mode was completely offset when an incentive was included.

Recommendations for Similar Surveys

For researchers conducting human dimensions surveys in similar contexts, the results we provided yield some recommendations. There was strong support that using incentives has a significant positive effect on response rates. In a multi-stage survey administration, where a screener survey is required to identify eligible individuals from within a larger population, the effect of an incentive is likely

[11] Note that a difference in response related to avidity would likely produce nonresponse bias in statistics of interest, such as measures of effort and annual trip expenditures.

to be concentrated to when the incentive was received.

The effects of purported topic and response mode were more context-specific. Mail mode returns were higher for the screener, and any potential incremental effect of mode on extended questionnaire response seemed to be mitigated through the use of an incentive. It is well known that using web surveys can increase problems related to coverage, selection, and nonresponse bias (Duda & Nobile, 2010; Stedman et al., 2019; Sterrett et al., 2017). Although these potential issues were minimized in our context by sampling from license databases and not relying only on the coverage provided by records with email addresses, these issues should be considered when deciding which mode to use for administering a survey. In our context, the ineligible portion of the population was similar to the eligible portion, as all were licensed anglers. Therefore, the finding that purported topic had little effect throughout this experiment might not hold in other contexts in which the general topic is less salient to individuals in the sample frame (e.g., drawing sample from the general population for a highly targeted survey).

The differences in predicted response rates among survey design factors is, of course, only one component of a larger benefit-cost analysis. The decision between a mail and web-based survey, and whether to include a monetary incentive, contribute to the total cost of data collection. Achieving a larger number of completed questionnaires is an obvious benefit, but financial costs need to be considered as well. Whether the added costs are outweighed by the increased response rates depends on the context-specific costs of administering the questionnaire using different modes and the direct cost of any incentives. Consideration of these costs alongside the relative benefits of different rates of response, as informed by the results of this study, should help increase the efficiency of conducting future human dimensions surveys.

References

Abbott, J. K., & Fenichel, E. P. (2013). Anticipating adaptation: A mechanistic approach for linking policy and stock status to recreational angler behavior. *Canadian Journal of Fisheries and Aquatic Sciences*, *70*(8), 1190–1208. https://doi.org/10.1139/cjfas-2012-0517

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, *80*(1), 123–129. https://doi.org/10.1016/S0165-1765(03)00032-6

Anderson, L. E., Lee, S. T., & Levin, P. S. (2013). Costs of delaying conservation: Regulations and the recreational values of exploited and co-occurring species. *Land Economics*, *89*(2), 371–385. https://doi.org/10.3368/le.89.2.371

Arlinghaus, R. (2006). Understanding recreational angling participation in Germany: Preparing for demographic change. *Human Dimensions of Wildlife*, *11*(4), 229–240. https://doi.org/10.1080/10871200600802889

Barrett, B. N., van Poorten, B., Cooper, A. B., & Haider, W. (2017). Concurrently assessing survey mode and sample size in off-site angler surveys. *North American Journal of Fisheries Management*, *37*(4), 756–767. https://doi.org/10.1080/02755947.2017.1324543

Baruch-Mordo, S., Breck, S. W., Wilson, K. R., & Broderick, J. (2009). A tool box half full: How social science can help solve human–wildlife conflict. *Human Dimensions of Wildlife*, *14*(3), 219–223. https://doi.org/10.1080/10871200902839324

Bunnefeld, N., Hoshino, E., & Milner-Gulland, E. J. (2011). Management strategy evaluation: A powerful tool for conservation? *Trends in Ecology & Evolution*, *26*(9), 441–447. https://doi.org/10.1016/j.tree.2011.05.003

Converse, P. D., Wolfe, E. W., Xiaoting Huang, & Oswald, F. L. (2008). Response rates for mixed-mode surveys using mail and e-mail/web. *American Journal of Evaluation*, *29*(1), 99–107. https://doi.org/10.1177/1098214007313228

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*.

Duda, M. D., & Nobile, J. L. (2010). The fallacy of online surveys: No data are better than bad data. *Human Dimensions of Wildlife*, *15*(1), 55–64. https://doi.org/10.1080/10871200903244250

Edwards, P. (2005). Meta-analysis of randomised trials of monetary incentives and response to mailed questionnaires. *Journal of Epidemiology & Community Health*, *59*(11), 987–999. https://doi.org/10.1136/jech.2005.034397

Gajic, A., Cameron, D., & Hurley, J. (2012). The cost-effectiveness of cash versus lottery incentives for a web-based, stated-preference community survey. *The European Journal of Health Economics*, *13*(6), 789–799. https://doi.org/10.1007/s10198-011-0332-0

Goyder, J. (1987). *The silent minority: Nonrespondents on sample surveys*. Polity Pr.

Groves, R. M., Presser, S., & Dipko, S. (2004). The role of topic interest in survey participation decisions. *Public Opinion Quarterly*, *68*(1), 2–31. https://doi.org/10.1093/poq/nfh002

Groves, Robert M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, *70*(5), 646–675. https://doi.org/10.1093/poq/nfl033

Groves, Robert M., Couper, M. P., Presser, S., Singer, E., Tourangeau, R., Acosta, G. P., & Nelson, L. (2006). Experiments in producing nonresponse bias. *Public Opinion Quarterly*, *70*(5), 720–736. https://doi.org/10.1093/poq/nfl036

Groves, Robert M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2013). *Survey methodology*. Wiley.

Groves, Robert M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation. *Public Opinion Quarterly*, *64*(3), 299–308. https://doi.org/10.1086/317990

Guo, Y., Kopec, J. A., Cibere, J., Li, L. C., & Goldsmith, C. H. (2016). Population survey features and response rates: A randomized experiment. *American Journal of Public Health*, *106*(8), 1422–1426. https://doi.org/10.2105/AJPH.2016.303198

Hunt, L. M., Phaneuf, D. J., Abbott, J. K., & Fenichel, E. P. (2020). Per trip changes to the economic value of Ontario, Canada anglers fishing the Laurentian Great Lakes under target species transitions. *Human Dimensions of Wildlife*, 1–16. https://doi.org/10.1080/10871209.2020.1800144

Jobber, D., Saunders, J., & Mitchell, V.-W. (2004). Prepaid monetary incentive effects on mail survey response. *Journal of Business Research*, *57*(4), 347–350. https://doi.org/10.1016/S0148-2963(02)00385-5

Kanuk, L., & Berenson, C. (1975). Mail surveys and response rates: A literature review. *Journal of Marketing Research*, *12*(4), 440. https://doi.org/10.2307/3151093

Kellner, J. B., Sanchirico, J. N., Hastings, A., & Mumby, P. J. (2011). Optimizing for multiple species and multiple values: Tradeoffs inherent in ecosystem-based fisheries management. *Conservation Letters*, *4*(1), 21–30. https://doi.org/10.1111/j.1755-263X.2010.00132.x

Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00513

Kreuter, F. (2013). Facing the nonresponse challenge. *The ANNALS of the American Academy of Political and Social Science*, *645*(1), 23–35. https://doi.org/10.1177/0002716212456815

Lee, M.-Y., Steinback, S., & Wallmo, K. (2017). Applying a bioeconomic model to recreational fisheries management: Groundfish in the northeast United States. *Marine Resource Economics*, *32*(2), 191–216. https://doi.org/10.1086/690676

Levin, P. S., & Anderson, L. E. (2016). When good fences make bad neighbors: Overcoming disciplinary barriers to improve natural resource management. *Coastal Management*, *44*(5), 370–379. https://doi.org/10.1080/08920753.2016.1208034

Lew, D. K., & Larson, D. M. (2014). Is a fish in hand worth two in the sea? Evidence from a stated preference study. *Fisheries Research*, *157*, 124–135. https://doi.org/10.1016/j.fishres.2014.04.005

Manfreda, K. L., Bosnjak, M., Berzelak, J., Haas, I., & Vehovar, V. (2008). Web surveys versus other survey modes: A meta-analysis comparing response rates. *International Journal of Market Research*, *50*(1), 79–104. https://doi.org/10.1177/147078530805000107

Marcus, B., Bosnjak, M., Lindner, S., Pilischenko, S., & Schütz, A. (2007). Compensating for low topic interest and long surveys: A field experiment on nonresponse in web surveys. *Social Science Computer Review*, *25*(3), 372–383. https://doi.org/10.1177/0894439307297606

Mercer, A., Caporaso, A., Cantor, D., & Townsend, R. (2015). How much gets you how much? Monetary incentives and response rates in household surveys. *Public Opinion Quarterly*, *79*(1), 105–129. https://doi.org/10.1093/poq/nfu059

Millar, M. M., & Dillman, D. A. (2011). Improving response to web and mixed-mode surveys. *Public Opinion Quarterly*, *75*(2), 249–269. https://doi.org/10.1093/poq/nfr003

Mitchell, M. N. (2020). *Interpreting and visualizing regression models using Stata* (Second). Stata Press.

Mize, T. (2019). Best practices for estimating, interpreting, and presenting nonlinear interaction effects. *Sociological Science*, *6*, 81–117. https://doi.org/10.15195/v6.a4

Müller, B., Quaas, M. F., Frank, K., & Baumgärtner, S. (2011). Pitfalls and potential of institutional change: Rain-index insurance and the sustainability of rangeland management. *Ecological Economics*, *70*(11), 2137–2144. https://doi.org/10.1016/j.ecolecon.2011.06.011

National Research Council of the National Academies. (2013). *Nonresponse in social science surveys: A research agenda*. National Academies Press. https://doi.org/10.17226/18293

Schmid, B., Balac, M., & Axhausen, K. W. (2019). Post-car world: Data collection methods and response behavior in a multi-stage travel survey. *Transportation*, *46*(2), 425–492. https://doi.org/10.1007/s11116-018-9968-2

Schuldt, J. P., Roh, S., & Schwarz, N. (2015). Questionnaire design effects in climate change surveys: Implications for the partisan divide. *The ANNALS of the American Academy of Political and Social Science*, *658*(1), 67–85. https://doi.org/10.1177/0002716214555066

Shannon, D. M., & Bradshaw, C. C. (2002). A comparison of response rate, response time, and costs of mail and electronic surveys. *The Journal of Experimental Education*, *70*(2), 179–192. https://doi.org/10.1080/00220970209599505

Singer, E., & Ye, C. (2013). The use and effects of incentives in surveys. *The ANNALS of the American Academy of Political and Social Science*, *645*(1), 112–141. https://doi.org/10.1177/0002716212458082

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*(10), 1139–1154. https://doi.org/10.1080/08941920.2019.1587127

Sterrett, D., Malato, D., Benz, J., Tompson, T., & English, N. (2017). Assessing changes in coverage bias of web surveys in the United States. *Public Opinion Quarterly*, *81*(S1), 338–356. https://doi.org/10.1093/poq/nfx002

Teel, T. L., & Manfredo, M. J. (2010). Understanding the diversity of public interests in wildlife conservation. *Conservation Biology*, *24*(1), 128–139. https://doi.org/10.1111/j.1523-1739.2009.01374.x

Tourangeau, R., Groves, R. M., & Redline, C. D. (2010). Sensitive topics and reluctant respondents: Demonstrating a link between nonresponse bias and measurement error. *Public Opinion Quarterly*, *74*(3), 413–432. https://doi.org/10.1093/poq/nfq004

Tourangeau, Roger, Groves, R. M., Kennedy, C., & Yan, T. (2009). The presentation of a web survey, nonresponse and measurement error among members of web panel. *Journal of Official Statistics*, *25*(3), 299–321.

Ward, K. J., Stedman, R. C., Luloff, A. E., Shortle, J. S., & Finley, J. C. (2008). Categorizing deer hunters by typologies useful to game managers: A latent-class model. *Society & Natural Resources*, *21*(3), 215–229. https://doi.org/10.1080/08941920701831913

Willcox, A. S., Giuliano, W. M., & Israel, G. D. (2010). Effects of token financial incentives on response rates and item nonresponse for mail surveys. *Human Dimensions of Wildlife*, *15*(4), 288–295. https://doi.org/10.1080/10871201003736047

Table 1: Allocated sample sizes and completed responses by randomized design factor

Response Mode	Purported Topic	Incentive	Sample Allocated (n)	Screener Complete (n)	Extended Complete (n)
Mail (n=2000)	Recreation (n=1000)	$0	336	31	20
		$2	332	110	18
		$5	332	114	18
	Saltwater Fishing (n=1000)	$0	336	33	18
		$2	332	97	22
		$5	332	116	27
Web (n=2000)	Recreation (n=1000)	$0	336	36	5
		$2	332	59	17
		$5	332	86	27
	Saltwater Fishing (n=1000)	$0	336	44	10
		$2	332	67	27
		$5	332	65	15
		Total	4,000	858	224

Table 2: Contents of contacts by design factor

	Purported Topic		Response Mode		Incentive		
	Recreation	Saltwater Fishing	Mail	Web	$0	$2	$5
Initial Contact							
Invitation letter:	Framing: recreational survey	Framing: saltwater fishing survey		Included URL			
Screener survey:	Questions: recreation and fishing	Questions: saltwater fishing only	In mail packet	Preceded web survey			
Cash incentive:					No	$2	$5
Survey Instrument	Cover: recreational activities	Cover: fishing graphics	Mailed to eligible anglers	Immediately followed web screener	No	No	No
Reminder	Framing: recreational survey	Framing: saltwater fishing survey		Included URL	No	No	No

Table 3. Logistic regression estimates of survey design factors and region on screener and extended questionnaire responses

	Prob(Complete Screener)						Prob(Complete Extended | Eligible)					
	n = 3998						n = 366					
	(1a) Main Effects Pseudo R^2 = .06			(1b) Interactions Pseudo R^2 = .06			(2a) Main Effects Pseudo R^2 = .03			(2b) Interactions Pseudo R^2 = .05		
	Coeff.	Odds ratio	*p*	Coeff.	Odds ratio	*p*	Coeff.	Odds ratio	*p*	Coeff.	Odds ratio	*p*
Intercept	-2.05	0.13	<.001	-1.55	0.21	<.001	0.05	1.05	.877	0.78	2.17	.213
Purported Topic: Recreation (v. Saltwater)	0.04	1.04	.589	-0.23	0.79	.330	-0.04	0.96	.841	-1.49	0.23	.073
Mode: Mail (v. Web)	0.45	1.57	<.001	-0.33	0.72	.177	0.16	1.17	.483	-0.99	0.37	.149
Incentive: $2 or $5 (v. $0)	1.14	3.12	<.001	0.50	1.65	.008	0.31	1.36	.222	-0.29	0.75	.660
Region (each v. Washington)												
Oregon	-0.23	0.80	.031	-0.23	0.79	.030	-0.29	0.75	.278	-0.34	0.71	.215
Northern California	-0.57	0.57	<.001	-0.57	0.56	<.001	0.85	2.34	.022	0.84	2.31	.026
Southern California	-0.63	0.53	<.001	-0.64	0.53	<.001	0.48	1.62	.121	0.46	1.59	.143
Topic x Mode:												
Recreation + Mail	-	-	-	0.16	1.18	.645	-	-	-	2.60	13.52	.008
Topic x Incentive:												
Recreation + $2 or $5	-	-	-	0.35	1.42	.202	-	-	-	1.25	3.48	.167
Mode x Incentive:												
Mail + $2 or $5	-	-	-	0.98	2.67	<.001	-	-	-	1.06	2.89	.172
Topic x Mode x Incentive												
Recreation + Mail + $2 or $5	-	-	-	-0.21	0.81	.606	-	-	-	-2.48	0.08	.026

Table 4. Contrasts of marginal linear predictions in interaction models for screener (1b) and extended questionnaire (2b) responses

	(1b) Screener, Interactions		(2b) Extended, Interactions	
Contrast	Chi-sq.	*p*	Chi-sq.	*p*
Purported Topic	0.07	.784	0.42	.516
Mode	3.62	.057	0.63	.429
Incentive	124.74	<.001	0.77	.379
Topic x Mode	0.09	.759	6.00	.014
Topic x Incentive	1.55	.213	0.00	.988
Mode x Incentive	19.34	<.001	0.10	.750
Topic x Mode x Incentive	0.27	.606	4.93	.026

Table 5. First and second differences of incentive and mode for screener response

Response Mode	Incentive	Probability (SE)	Effect of Mode at Incentive Levels	Effect of Incentive at Mode Levels
Mail	$0	0.10 (0.01)		
	$2 or $5	0.33 (0.01)		Prob (Mail, $0) – Prob (Mail, $2 or $5) = -0.23***
Web	$0	0.12 (0.01)	Prob(Mail, $0) – Prob (Web, $0) = -0.02	
	$2 or $5	0.21 (0.01)	Prob (Mail, $2 or $5) – Prob (Web, $2 or $5) = 0.12***	Prob (Web, $0) – Prob (Web, $2 or $5) = -0.09***
			Difference by Incentive = -0.14 ***	*Difference by Mode* = -0.14 ***

*** $p < .001$

Table 6. First and second differences of incentive and mode, by topic, for extended questionnaire response

Topic	Response Mode	Incentive	Probability (SE)	Effect of Mode at Incentive Levels	Effect of Incentive at Mode Levels
Recreation	Mail	$0	0.73 (0.08)		
		$2 or $5	0.63 (0.06)		Prob (Mail, $0, recreation) – Prob (Mail, $2 or $5, recreation) = 0.10
	Web	$0	0.36 (0.13)	Prob (Mail, $0, recreation) – Prob (Web, $0, recreation) = 0.37**	
		$2 or $5	0.59 (0.06)	Prob (Mail, $2 or $5, recreation) – Prob (Web, $2 or $5, recreation) = 0.05	Prob (Web, $0, recreation) – Prob (Web, $2 or $5, recreation) = -0.23
				Difference by Incentive = 0.32*	*Difference by Mode* = 0.32*
Saltwater Fishing	Mail	$0	0.48 (0.08)		
		$2 or $5	0.66 (0.06)		Prob (Mail, $0, saltwater) – Prob (Mail, $2 or $5, saltwater) = -0.18*
	Web	$0	0.70 (0.12)	Prob (Mail, $0, saltwater) – Prob (Web, $0, saltwater) = -0.23	
		$2 or $5	0.64 (0.06)	Prob (Mail, $2 or $5, saltwater) – Prob (Web, $2 or $5, saltwater) = 0.02	Prob (Web, $0, saltwater) – Prob (Web, $2 or $5, saltwater) = 0.06
				Difference by Incentive = -0.24	Difference by Mode = -0.24

*$p < .10$, **$p < .05$, ***$p < .001$

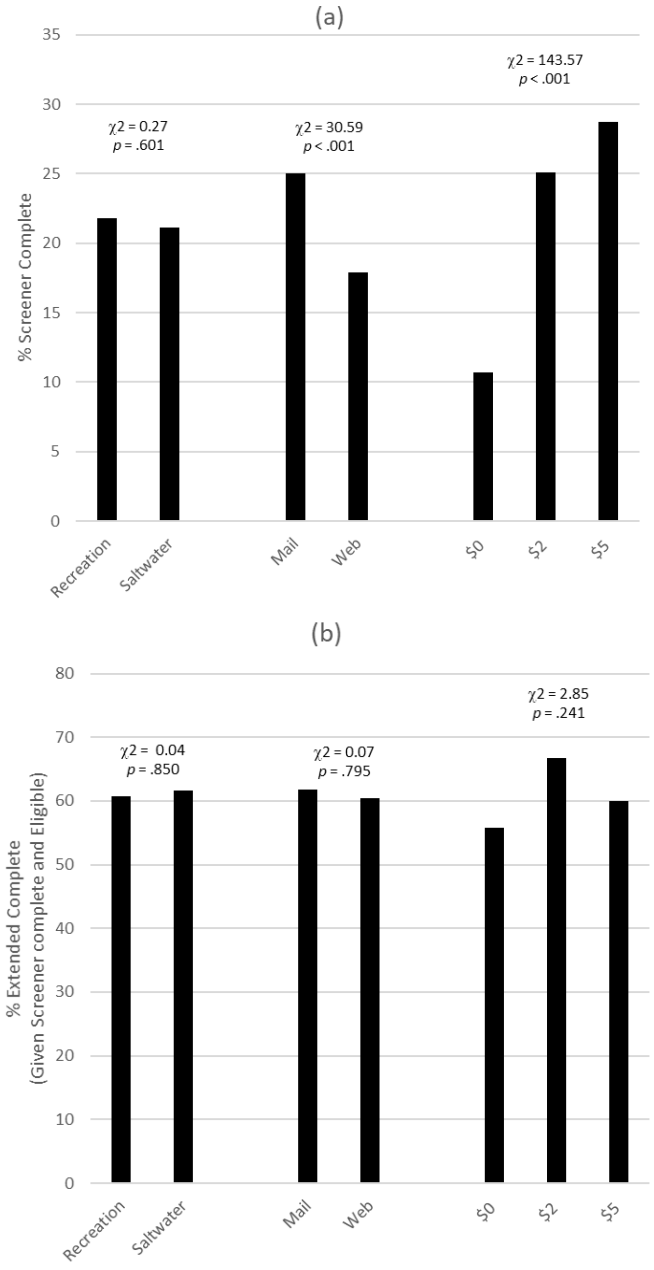

Figure 1. Bivariate effects of mode, purported topic, and incentive on (a) screener and (b) extended completes

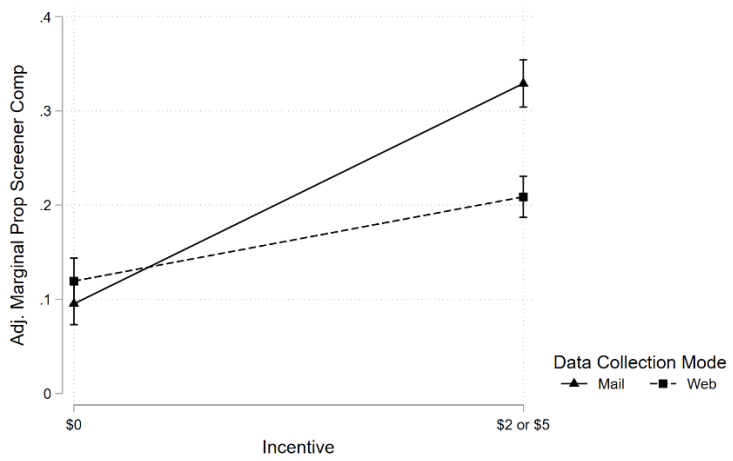

Figure 2. Effect of incentive and mode on screener response

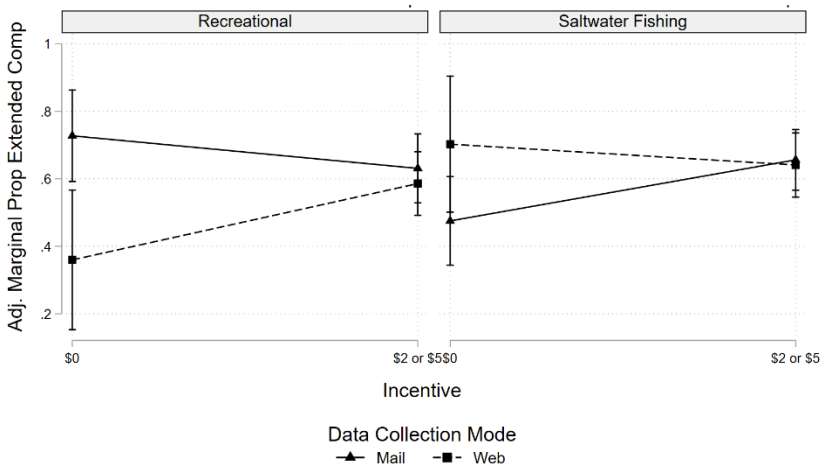

Figure 3. Effect of mode, incentive, and topic on extended questionnaire response