Survey Mode Effects on Valuation of Environmental Goods

Jason Bell  
Duke University

Joel Huber  
Duke University

W. Kip Viscusi  
Vanderbilt University

This paper can be downloaded without charge from the Social Science Research Network Electronic Paper Collection:  
http://ssrn.com/abstract=1739141
Survey Mode Effects on Valuation of Environmental Goods

Jason Bell¹, Joel Huber², and W. Kip Viscusi³*

1 Fuqua School of Business, Duke University, Durham, NC 27708.
2 Fuqua School of Business, Duke University, Durham, NC 27708.
3 Vanderbilt Law School, Vanderbilt University, 131 21st Avenue South, Nashville, TN 37203

E-Mail: jbb@duke.edu (Jason Bell), jch8@mail.duke.edu (Joel Huber), kip.viscusi@vanderbilt.edu (W. Kip Viscusi).

* Corresponding author: Tel: (615) 343-6835; Fax: (615) 322-5953.

Abstract: This article evaluates the effect of the choice of survey recruitment mode on the value of water quality in lakes, rivers, and streams. Four different modes are compared: bringing respondents to one central location after phone recruitment, mall intercepts in two states, national phone-mail survey, and an Internet survey with a national, probability-based sample. The modes differ in terms of the representativeness of the samples, non-response rates, sample selection effects, and consistency of responses. The article also shows that the estimated benefit value can differ substantially depending on the survey mode. The national Internet panel has the most desirable properties with respect to achieving the survey dimensions of interest.

Keywords: survey methods; environmental economics; environmental policy; stated preference; benefit-cost analysis; water quality.

1. Introduction

The choice of survey recruitment mode has a potentially important influence on the measurement of the value of environmental goods based on survey responses. There are two principal dimensions of influence that we analyze in this article. First, the mode may influence who chooses to respond to the survey, thus affecting the extent to which the responses reflect the valuations of the population of interest. Second, for the particular sample of respondents the survey recruitment mode affects whether the survey elicits their preferences accurately. Thus, the mode may affect how they respond to the
survey questions and the valuations that are elicited. This article explores the valuation of a single environmental commodity using different survey recruitment modes. We find that the manner in which potential respondents are recruited, the likelihood of respondents to self-select into or out of the survey, the costs in time and effort imposed by the survey, and the diligence with which respondents complete the survey task each can affect the estimated value of the good.

The substantive focus of the survey is the valuation of improvements in inland water quality. In particular, how much do people value increases in the quality of lakes, rivers, and streams? The quality dimensions of interest include recreational uses such as swimming and fishing, but also include ecological benefits to plants, fish, and wildlife that are associated with clean water. People may value many of these benefit components irrespective of whether the person visits lakes and rivers. As a consequence, attempts to elicit monetary values of water quality based on recreational visits to lakes and rivers cannot capture all the benefits associated with water quality. To obtain these values, some kind of survey approach that can elicit meaningful measures of water quality is essential. Because the survey structure we have designed involves an interactive computer-based valuation task, the survey mode must be able to both accommodate this approach and create a context in which a representative sample of respondents can give thoughtful responses to the valuation task.

To examine the influence of survey mode, this article investigates the differences in responses to an interactive computer survey using four survey modes: Internet panel, phone-mail, mall intercept, and central location. These modes differ in terms of the manner of recruitment, the costs they impose on the respondent, and the environment in which the survey is administered. The range of survey modes examined here is not exhaustive. For example, we do not consider either door-to-door surveys because of the increasingly high cost of obtaining a representative sample of respondents to an interactive computer survey based on door-to-door visits by a survey firm representative.

We find substantial differences across four modes in terms of monetary values of water quality. Differences arise both in terms of who responds to the survey and how they respond. Our examination of the differences in demographic characteristics of participants highlights the effect of the survey mode on the selection of respondents into the sample pool. By analyzing the predicted environmental benefit values controlling for demographic mix we also can demonstrate that there is an important selectivity effect that biases the empirical estimates in the econometric model. Our review of the performance of the survey modes also indicates substantial differences in the rates of inconsistency in answering the survey questions, which is one measure of how the survey mode affects respondents’ ability to give meaningful answers to the valuation task.

We also consider the important objectives of a successful survey administration, and examine how well each survey mode fostered those objectives. In particular, we conclude that a probability based Internet panel is best suited to the objective of tapping a representative sample of potential respondents. Internet panels mitigate the effects of respondents self selecting disproportionately into topics, such as the environment, in which they have strong interest; it limits the time and travel costs associated with completing the survey, and it enables the survey to be taken at home where the respondent is comfortable completing the survey. Phone-mail, central location, and mall intercept modes have favorable features with respect to some of these objectives, and may be preferable if cost is a major concern, but are problematic in one or more of them.
We begin by describing the survey instrument since the instrument will affect which survey modes are feasible and their relative merits in eliciting benefit values. We then examine the various modes used for fielding the survey. Following this discussion, we review the key dimensions on which surveys should be judged. Using these criteria, we analyze the valuation results obtained using each of the survey modes. Then follows an analysis of the extent of inconsistent responses in each survey mode, an important measure of the degree to which the survey mode helps or hinders the generation of accurate and useable individual valuations.

The representativeness of a survey sample can be assessed in two ways. First, for each survey mode it is possible to compare the sample characteristics with the national adult population. Second, since the Internet panel collects demographic information on all panelists, for that survey mode the presence of any sample selection effects can be estimated. We will examine demographic effects in terms of the choice to participate in the Internet panel survey and compare those effects with the demographics of the other survey modes to see the extent to which each mode experiences sample selection effects. Our concluding discussion reviews the reasons why we believe that a nationally representative Web-based panel drawn using a probability sample of the U.S. population is the most meaningful approach.

2. Survey Instrument

The survey used in this analysis focuses on the value of water quality for inland water—lakes, rivers, and streams. Specifically, the survey elicited the monetary value of lake and river quality in a respondent’s region. The dimensions of water quality that were highlighted in the survey are whether fish caught in them were safe to eat, whether swimming in the water could make one ill, and whether the lake or river supported a healthy environment of plants, fish, and other aquatic life. These dimensions reflect the water quality dimensions used by the U.S. Environmental Protection Agency (EPA) for its National Water Quality Inventory.¹

In order to avoid focusing on idiosyncratic aspects of the respondent’s region that could not be monitored and might affect responses in unpredictable ways, the survey asked respondents to think about a hypothetical move to one of two new regions that resembled their own region in terms of number of water bodies and general characteristics. The goal of this multiple-question set is to obtain a meaningful point valuation for each individual respondent for an unfamiliar, non-market good using a few relatively simple choices.² The survey approach uses a series of iterative paired comparisons patterned after the approach pioneered by Viscusi, Magat, and Huber (1991). Respondents make a series of pairwise regional choices that differ on two dimensions: water quality and cost of living.

Figure 1 shows the text of a representative question. Respondents first choose one of the two regions. Based on the individual response, the survey then alters the choice comparison to make the choices more equally valued, where the overall objective is to find the point of indifference between the two regions. As part of this iteration process, subsequent questions either reduce the difference in

---

¹ The National Water Quality Inventory is a measure of water quality conditions in the United States and can be found at http://water.epa.gov/lawsregs/guidance/cwa/305b/index.cfm
² For further information regarding the iterative choice method used in the survey instrument, see Magat, Huber, Viscusi (2000).
water quality between the two regions or reduce the difference in cost of living between regions to estimate the point at which the respondent is indifferent between the presented options. Even if the respondent does not indicate strict indifference between the options, the survey generates a bounded value for the dollar value of improved water quality that lies between the tradeoff rates for the answers to the two sequential questions bracketing the last switch. There is no bound where the respondent reaches a corner solution without ever switching. In that case valuation is estimated econometrically using censored regression models.

**Figure 1. Text of Water Quality Survey Question**

Imagine again that you must move to another region of the country. You have narrowed your choices down to two regions. They differ in only two ways, the quality of the water and the annual cost of living in the regions. They even have the same number of acres of lakes and miles of rivers within 2 hours or so of where you would live.

| Region 1 | Region 2 |
|----------|----------|
| Increase in Annual Cost of Living | $100 | $400 |
| More Expensive | More Expensive |

| Percent of Lake Acres and River Miles with Good Water Quality Water Quality |
| 50% | Good | Water | Quality |
| 65% | Good | Water | Quality |

| Which Region Would you prefer? |
| Region 1 | Region 2 | No Preference |
| * | * | * |

Once a respondent reached a corner solution we include other questions to test for the rationality. In the first question the last chosen option becomes dominated by the alternative. Respondents choosing the dominated option are informed that they have done so and given the opportunity to alter their response. Respondents who fail alter their response are labeled inconsistent. These inconsistent respondents either do not understand the process or are actively protesting it. The percent of respondents so protesting is an important measure differentiating the four survey modes.

3. Survey Modes

The survey was fielded fourteen times from August 1997 to October 2004. All of the surveys were restricted to those older than 18 years and were administered on computers. Table 1 provides...
information on the timing and implementation of the phone recruitment to a central location in North Carolina, the mall intercepts in North Carolina and Colorado, the national phone-mail, and the national Internet panel recruited by Knowledge Networks. These administrations produced a total of 5,122 completed surveys, each of which can be used to generate an estimate of a respondent’s value of a one percentage point change in water quality. This value and the demographic characteristics of each respondent are the principal basis of the analysis. We augment the examination of survey modes with additional measures such as the frequency of inconsistent responses.

### Table 1: Characteristics and Timing of the Survey Modes

| Survey Mode                                      | Date           | Interviews |
|--------------------------------------------------|----------------|------------|
| Central Location, Research Triangle Park, NC     | August 1997    | 106        |
| Mall Intercept, Cary, Charlotte, Co. Springs, Denver | January 1998 | 303        |
| National Phone Mail 1                            | September 1999 | 33         |
| National Phone Mail 2                            | June 2000      | 53         |
| Internet Panel Pretest                           | December 2001  | 383        |
| Internet Panel Round 1                           | October 2002   | 184        |
| Internet Panel Round 2                           | February 2003  | 406        |
| Internet Panel Round 3                           | April 2003     | 580        |
| Internet Panel Round 4                           | April 2004     | 549        |
| Internet Panel Round 5                           | August 2004    | 516        |
| Internet Panel Round 6                           | October 2004   | 2009       |

The first survey mode that was used involved bringing respondents to a central location. This survey was administered by the marketing firm Johnston-Zabor and Associates in 1997 in Research Triangle Park, North Carolina. The survey firm recruited people by phone from a convenience sample of respondents that had completed surveys in the past. The survey firm asked people to visit a central location to complete the survey on a computer.

The second survey mode used involved a standard mall intercept survey administered in shopping malls in Cary and Charlotte, North Carolina, and Colorado Springs and Denver, Colorado in 1998 by the marketing firm, Consumer Pulse. Representatives of the firm recruited mall shoppers to participate in the survey using computers at the mall location.

The phone-mail mode was also conducted by Consumer Pulse in late 1999 and mid 2000. Nationwide random digit dialing was used to recruit the sample. After agreeing to participate in the study, respondents were sent a disk by mail, which they used to complete the survey on their own computers. After doing so, they returned the completed survey disk by mail. Those without computers were offered additional compensation to use a neighbor’s computer or a computer available at a public location. Though over 75% of this sample used their home computer, 8% used a computer at work, and 13% used a friend’s computer. Only about 3.5% of respondents used a public location.
Knowledge Networks (KN) conducted the Internet panel mode between 2001 and 2004. This sample consisted of people previously recruited by nationwide random digit dialing to join a panel to take surveys online. KN invited a group of panel members to participate in our survey.

While the survey questions were similar throughout the modes, there were some differences. First, the starting cost and quality differences between regions presented to respondents differed both within the Internet panel administration as well as between survey modes. For instance, respondents in the Internet panel were presented starting cost-to-quality ratios between $5 per 1% quality difference and $30 per 1% quality difference. Thus, depending on which starting ratio the respondent received, the initial questions asked whether the respondent was willing to pay $5 or $30 per 1% improvement in water quality. The central location mode had a starting tradeoff ratio of $4, and the phone-mail and mall intercept had a ratio of $10. As described in Huber, Viscusi, Bell (2008), higher starting points can result in higher final valuations. The influence of starting points on the respondent’s valuation is accounted for in the regression analysis.

Additionally, the survey modes differed in terms of the range over which the water quality differences spanned. The lowest, or baseline percentage of water rated of good quality, presented water quality rated good as ranging from 20% to 75% in the Internet panel and mall intercept, while the central location and phone-mail presented 50% as the baseline water quality. Previous research found that these starting points influence values, with higher baseline quality leading to lower cost-quality tradeoff values as levels of water quality have a diminishing marginal value to respondents. Even though the mean baseline quality was similar across modes, the potential influence of the starting level of water quality on valuations is also accounted for at the level of the individual in the regression analysis. Finally, the phone-mail and Internet panel surveys contained a slightly larger question set. If respondents in those surveys continued to choose the option with higher cost and higher water quality, they were asked one additional question relative to the other modes before being presented with the dominated choice. Using Figure 1 to illustrate, the earlier surveys would have presented high-valuation respondents a 15% quality difference (65%-50%), then 10% then 5%, then 0%. The later surveys would have added a question with a 3% quality difference. The practical effect of this additional question, all else equal, should be fewer respondents whose values are censored at high values for the phone-mail and Internet panels because of the additional opportunity to switch choices of region, and those surveys could generate higher tradeoff rates for values at the censored point ($300/5% or $60 for earlier surveys and $300/3% or $100 for later surveys).

4. Sample Selection Differences among Survey Modes

Survey participation is a function of

---

3 For additional information on the characteristics of the KN panel, see http://www.knowledgenetworks.com/knpanel/docs/KnowledgePanel%28R%29-Design-Summary-Description.pdf.
4 For a more complete discussion of starting point and information effects, see Huber, Viscusi, Bell (2008)
1. Ability of investigators to contact a potential respondent
2. Interest of a potential respondent in the topic
3. Total time and effort cost for a respondent to complete the survey
4. Ability of a respondent to be comfortable in the location where the survey is completed

The most desirable mode will increase the performance on dimensions (1) and (4) and attempt to engage the respondent to promote dimension (2) and limit the time and effort cost dimension (3).

We investigate how responses to survey questions can vary when the substantive difference between surveys was how the sample was recruited. All respondents answered the survey analyzed here electronically, either via the Internet or with a computer recording responses to a disk. Because of this similarity in administration, as well as the collection of personal characteristics of every respondent, the differences between survey responses can be measured largely on the basis of how and whether a respondent was recruited to participate.

An objective of any survey is to get a minimally biased sample of a target universe, which in our case is the adult U.S. population. There was, of course, no expectation that the results from a survey administered in a single region such as North Carolina or Colorado would reflect national preferences since the density of lakes and rivers varies substantially across the country. This regional survey served to explore how people would respond to the survey questions. The discussion below will highlight some of the regional differences that arise. However, our main interest here is with potential selection biases in the four recruiting modes that limit the ability of the researcher to project the results to any target universe. The limitations based on selective regional coverage are not inherent shortcomings of the survey mode as one could, following the previous example, use mall intercepts throughout the country.

There is a substantial literature regarding the effects of survey mode on responses, particularly due to the increasing difficulty over time of reaching potential survey respondents by phone, mail, and e-mail. This difficulty has led to the use of convenience samples, using both phone lists and opt-in Internet samples, samples drawn from mall intercept recruiting, and samples drawn from panels.

Of great concern is whether the mode by which a survey is implemented affects response rates, results, and the demographic characteristics of the surveyed sample. For instance, research has shown that questions presented visually can lead to different answers compared to questions presented aurally, particularly for questions using scales.5

Differences between respondents and non-respondents are also evident in previous survey research. For instance, Rhodes et al. (1990) found age, gender, urban/rural, and health related effects between early and late responders to a 1981 health-related survey in Spain using mail recruitment, with multiple telephone and in-person follow-ups for non-respondents. Such differences might indicate that non-responders are more like late responders than those who complete a survey early and without reminders or other follow-up.

5 Dillman et al. (2009)
Interest in the survey topic can also be a significant determinant on survey participation. This has been previously noted as a factor in survey participation for mail surveys. We investigate further, as to whether differences in level of interest translate into value differences across recruiting modes.

5. Demographic Differences among Modes

Because the survey modes differ in terms of their ability to reach the target population, one can expect differences in the demographic characteristics between survey modes. In each case we use the U.S. Census adult population as the reference point for determining the representativeness of the sample. Table 2 shows the portion of the sample that took the survey through KN probability based Internet panel. The sample characteristics closely match the demographics of the adult population in the United States. This matching is to some extent due to the fact that the demographics of the Internet panel are known before invitations are sent so that KN can draw a nationally representative sample of respondents for such studies. The close match is also due to the fact that potential respondents are already known to be willing to take surveys by their presence in the panel, making their participation more likely. However, this willingness may also lead to other differences related to panel membership, such as whether the taking of multiple surveys affects answers to the next survey.

![Table 2: Comparison of Sample to the National Adult U.S. Population](table)

---

6 See Martin (1994), Roberson and Sundstrom (1990).
7 See Taylor et al. (2009)
The other survey modes perform much less well in terms of matching national population characteristics. For instance, the phone-mail sample included dramatically fewer respondents under 35 years of age, almost no respondents with less than high school education, twice as many college graduates or higher, fewer minorities, many more married respondents, and very few respondents with incomes below $15,000. These differences are as expected, older, married, wealthier, and more educated respondents should be more likely to have the ability to go through the process of contact, being mailed survey materials, completing the survey, and mailing it back. More unfortunately, the phone-mail sample had a much higher incidence of respondents who have visited a lake or river in the last year than other modes. The bottom of Table 2 shows that over 90% of respondents in the phone-mail sample had visited lakes or rivers in the past year as compared to 68% with the other survey modes. This difference in visitation percentages may arise from self-selection, where potential
applicants decided whether to participate based on the subject matter of the survey. This bias can be limited but not eliminated by including the regression coefficients for whether the respondent visited lakes or rivers.

This self-selection effect is also problematic for the central location sample, also a mode where the sample was recruited by telephone. That sample had a similar percentage of respondents reporting a visit a lake or river in the last year—89%. As with the phone-mail mode, the central location sample also had fewer respondents who were young (18-24), though it had many more of the next age group (25-34), leading to a mild 31% vs. 35% oversampling of 18-34 year olds in the central location mode. The central location respondents were much less likely to be Hispanic, more likely to be married, wealthier, and more educated than the nation overall. Presumably, some of these differences with national population statistics are due to the same factors as phone-mail, but is also due to the location of the survey, Research Triangle Park in North Carolina. That area has a higher portion of college educated professionals, especially among those reachable by phone and who were comfortable enough with the area to make the trip to the survey location. Over 75% of the central location respondents had a college degree, even higher than the 47% with that level of education in that region overall.8

The mall intercept sample had expected differences from national demographics. That sample is skewed much younger than the other modes, has the highest rate of participation among black respondents, and is the only mode with a majority of unmarried respondents. Respondents in the mall intercept sample were also less educated and less wealthy than either mode first contacted by phone. These characteristics are not surprising, as respondents were recruited from the subsample of people visiting a shopping mall with enough time to interrupt their shopping trip to take 25 minutes to complete the survey. However, this sample seemed to suffer much less from the self-selection based interest in the topic of lakes and rivers, as 75% of respondents reported having visited a lake or river in the last year. This percentage is still higher than the 68% among the Internet panel sample, where panelists’ general likelihood to participate in surveys in general appears to minimize this self-selection. Presumably the largest self-selection factors for the mall intercept sample were the inclination to visit a mall and the ability to spare the time.

6. Valuation Differences among Modes

Because each mode generated samples with different demographic characteristics, one would expect there to be differences between samples in the valuations of water quality. We used the two-tailed Tobit regression analysis shown in Table 3 to generate estimates of the water quality valuations that account for censoring effects. The two-tailed Tobit approach is appropriate because any given respondent, if a point estimate could not be determined, was assigned a low or high value determined by their choice on their final question. Their actual value could be much lower than the assigned low value or much higher than the assigned high value. The use of Tobit regressions accounts for that indeterminacy. As the results in Table 3 indicate, relative to the Internet panel, phone-mail and mall intercept modes generate greater valuations of improved water quality.

8 http://www.researchtriangle.org/uploads/pdfs/RTRP_Region.pdf
The estimates in Table 3 also demonstrate a general effect of demographic differences on values. As the table shows, values for regional water quality amenities differ based upon interest in the environment (as measured by membership in an environmental organization), direct use of the good through visits, household income, education, age, and minority status. However, even after accounting for those factors as well as differences in starting points and baseline water quality between surveys, the phone-mail and mall intercept samples still have substantially higher values than the Internet panel or central location administration.
Table 4 shows how these differences are reflected in value estimates. The top two rows show the differences in terms of raw estimates (the point in the survey where the respondent reached indifference between choices, was narrowly bound by answers to two questions, or reached the end of the set of questions without having a value bounded both high and low). However, those values do not account for differences in demographic characteristics, starting values, or baseline water quality among survey modes. In addition, it does not address the fact that the values for about 20% of respondents were censored either high or low when they reached the final question in the iterated question set.

Table 4: Estimated Water Quality Values, Censoring of Extreme Values, and Starting Points

| Variable                                      | Full Sample | Internet Panel | Phone Mail | Mall Intercept | Central Location |
|-----------------------------------------------|-------------|----------------|------------|----------------|------------------|
|                                               | 4851        | 4406           | 83         | 264            | 98               |
| Raw Regional Water Quality Mean Value         | $22.04      | $22.54         | $25.62     | $16.17         | $12.19           |
| Raw Median                                    | $15         | $15            | $15.38     | $13.64         | $13.33           |
| Using Full Sample Average Demographics        |             |                |            |                |                  |
| Estimated Regional Water Value (Log)          | $32.10      | $31.31         | $53.67     | $42.83         | $29.36           |
| Difference from Internet Panel (Log)          |             |                |            |                |                  |
| Censored High                                 | 11.7 %      | 10.5 %         | 36.1 %     | 23.9 %         | 15.3 %           |
| Censored Low                                  | 8.4 %       | 8.8 %          | 1.2 %      | 5.3 %          | 2.0 %            |
| Starting Ratio                                | $15.49      | $16.18         | $10        | $10            | $4               |
| Baseline Quality                              | 53.65 %     | 54.03 %        | 50 %       | 49.9 %         | 50 %             |

These differences are accounted for in two ways in the next two rows of Table 4. First, each mode was assigned the average demographic values of everyone in the complete sample of all surveys regardless of mode. Therefore, even though for example the central location sample was much older in practice and started with a lower initial cost-quality tradeoff, they are assigned the common year and starting point for this calculation. These adjustments to all measured factors except survey mode reveal the extent to which the survey mode affects estimated values of regional water quality. Compared to the Internet panel, the phone-mail sample had estimated values 71% higher, mall intercept 37% higher, and the central location sample 6% lower.

Thus while the regression accounts for systematic differences between the panels, there appear to be differences in participation not reflected in the demographic, visitation, or survey text differences that increase valuations for phone-mail and mall intercept surveys. Since all the surveys were implemented on computers, the differences then arise largely from self selection and the environment where computers were used. This latter factor is discussed below in an examination of the inconsistency of responses.
7. Level of Inconsistent Responses among Modes

Aside from dramatic differences in valuation of water quality, the four different survey modes differed in terms of how often respondents passed the consistency test. A consistency test is important because it is a measure of the extent to which respondents understand and are attentive to the survey task, and accordingly minimize irrational or protest responses.

For this survey, consistency was assessed for respondents who consistently choose the lower cost or the higher quality region. For these respondents, a dominated choice was provided, where the region they had consistently chose was clearly worse than the other option. This choice either had the same water quality in both regions while one had a higher cost of living (for respondents who had been choosing higher quality) or had the same cost of living while one had a higher water quality (for respondents who had been choosing lower cost of living). If the respondent chose the dominated region, the computer informed the respondent about that inconsistency. If the respondent still persisted in that choice, that respondent was deemed inconsistent and not included in the regular analysis.

Table 5 shows the differences in the percentage of inconsistent respondents for each survey mode. While overall, 5.3% of respondents were inconsistent, that rate ranged from a low of 3.5% for the phone-mail sample to a high of 12.9% for the mall intercept sample. In terms of simple t-tests, the Internet panel has significantly fewer inconsistent respondents than the complete sample of surveys, as well as significantly fewer insignificant responses at the high value portion of the decision tree compared to the other modes (t statistics of 5.04 and 4.92 respectively). The mall intercept sample was the worst on both the high and low value questions (t statistics of -6.10, -5.90, and -2.0 respectively). While phone-mail had the lowest nominal level of inconsistency, that difference was not significantly lower than the overall sample or than the Internet panel.

Table 5: Percent Inconsistent Responses by Survey Mode

| Percent Inconsistent | N   | Total   | Inconsistent at High Value | Inconsistent at Low Value |
|----------------------|-----|---------|---------------------------|--------------------------|
| Survey Mode, Internet Panel | 4627 | 4.78 %  | 3.28 %                    | 1.49 %                   |
| Survey Mode, Phone - Mail     | 86  | 3.49 %  | 1.16 %                    | 2.33 %                   |
| Survey Mode, Mall Intercept  | 303 | 12.87 % | 9.90 %                    | 2.97 %                   |
| Survey Mode, Central Location| 106 | 7.54 %  | 6.60 %                    | 0.94 %                   |

The demographic differences predicting inconsistent respondents are reflected in the probit regression reporting marginal effects in Table 6. The income and visitation differences suggest that not all inconsistency may be due to inattentiveness or confusion, some of the respondents classified this way may be registering a sort of protest or merely an insistence on the direction of previous answers. After accounting for demographic and starting point differences, the mall intercept sample is more inconsistent overall and at the high level relative to the Internet panel sample, and the phone-mail sample is less inconsistent at the high level.

Table 6: Probit Regressions Predicting Inconsistency using Demographic Characteristics

---

*a*
The poor performance of the mall intercept sample in terms of inconsistent respondents causes some concern. These respondents were invited to interrupt shopping to go into a room off a noisy but otherwise stimulating mall. If these respondents are more likely to be uncomfortable, rushed, or in some way affected by participating in the survey due to their attention being focused elsewhere, then this survey mode may be inappropriate, particularly for more complicated surveys.
The better performance of the Internet panel and the phone-mail sample illuminates this possibility. For these samples, respondents generally completed the survey in their own homes at a convenient time of their choosing. This comfort may have resulted in better attention, less hurriedness, and therefore more thoughtful and fewer inconsistent responses.

Since Internet panel and phone-mail respondents were generally able to complete the survey in their own homes, the lower rate of inconsistency is understandable. The central location and mall intercept respondents may have felt hurried or simply not completely comfortable in a foreign environment. Central location respondents may have outperformed mall intercept since the former committed to a specific trip to participate. Mall intercept respondents were merely convinced to participate during an already planned outing, and thereby may not have given the survey as much attention as the central location group.

8. Non-response Characteristics

Unfortunately, it is difficult to precisely identify effects of survey modes when non-response is involved. Except for the Internet panel, little information is available about those who declined to participate.

The Internet panel sample provides some insight about the characteristics of those who declined to participate in that survey. While offering little information about people who refuse to take surveys at all, it is instructive to identify those members of the Internet panel who declined the invitation to participate. Knowledge Networks’ Internet panel has a broad set of basic demographic characteristics that is collected independently of our survey. Table 7 compares demographic information for those who completed the survey with those who declined to participate, and Table 8 presents a probit regression reporting significant marginal effects of those characteristics on the probability of participation in the survey. In general, older and more educated invitees agreed to complete the survey, while those at the top income category and minority groups were less likely to agree.

Table 7: Non-Response Characteristics in Internet Panel

| Variable       | Completed Survey | Declined Invitation |
|----------------|------------------|---------------------|
| Income         | $51,671          | $50,862             |
| Top Income Category | 1.9 %     | 3.4 %               |
| Years of Education    | 13.17      | 12.82               |
| Age (Years)       | 44.70           | 37.37               |
| Black            | 13.4 %          | 20.5 %              |
| Other            | 6.4 %           | 7.5 %               |
| Hispanic         | 10.6 %          | 14.8 %              |
| Female           | 49.0 %          | 50.5 %              |
| Married          | 56.4 %          | 49.0 %              |
| Northeast        | 18.5 %          | 18.8 %              |
| South            | 36.2 %          | 38.1 %              |
| West             | 21.6 %          | 22.0 %              |
Table 8: Probit Regressions Predicting Non-Response in Internet Panel

| Accepted Invitation to Participate in Survey | dF / dx   | Std. Err. |
|---------------------------------------------|-----------|-----------|
| Log (Income)                                | 0.0025    | 0.0067    |
| Top Income Category                         | -0.1697***| 0.0465    |
| Years of Education                          | 0.0100*** | 0.0023    |
| Age                                         | 0.0050*** | 0.0004    |
| Black                                       | -0.1107***| 0.0181    |
| Other                                       | -0.0490** | 0.0250    |
| Hispanic                                    | -0.0805***| 0.0196    |
| Female                                      | -0.0084   | 0.0115    |
| Married                                     | 0.0047    | 0.0123    |
| Northeast                                   | -0.0225   | 0.0186    |
| South                                       | -0.0209   | 0.0157    |
| West                                        | -0.0208   | 0.0181    |

| observed P                                  | 0.7531    |
| predicted P                                 | 0.7650    |
| Observations                                | 5642      |
| LR chi2(11)                                 | 310.13    |
| Prob > chi2                                 | 0.0000    |
| Pseudo R2                                   | 0.0492    |
| Log likelihood                              | -2998.2614|

*Notes: Coefficients have been transformed to equal marginal effects; * significant at 10%; ** significant at 5%; *** significant at 1%

Comparing Table 7 with Table 2, phone-mail mode most closely tracks these results. Phone-mail had more educated, older, fewer Black, and fewer Hispanic respondents, all significant predictors of participation in the probit regression reporting marginal probability effects in Table 8. In addition, the phone-mail respondents were more likely to be married and had higher incomes than the Internet panel respondents, both of which were seen nominally in Table 6 though not statistically significant in the probit.

The central location mode only corresponded with non-response expectations in terms of years of education and Hispanic ethnicity. Otherwise, these respondents were somewhat younger, possibly due to the travel requirement to participate at a central location.

The mall intercept mode also accorded with expectations in terms of education and Hispanic ethnicity. However, these respondents were younger, more likely to be Black, had lower incomes, and were less likely to be married than respondents in the Internet panel. This could be due to the
demographic characteristics of shoppers at the mall where the respondents were recruited, as well as mall shoppers who had enough time available to complete a survey.

So, generally, each mode’s demographic makeup corresponds, to some extent, to the factors seen in the Internet panel that affected participation, but each also has participation affected by features of its own sampling characteristics. The consistent effects of such factors on survey participation are cause for concern in terms of achieving a nationally representative sample, and provide a reason to use panels that are less affected by self selection.

9. Conclusions

We examined two ways in which four survey recruiting modes might influence the valuation outcomes. First, we showed that survey mode influences those who choose to respond to the survey. Second, we demonstrated that the different survey modes affects whether the survey elicits their preferences accurately.

Significant and large differences were found in the valuation of an environmental good across recruitment modes of survey administration. These differences persist even when demographic and survey question differences are accounted for, and are most likely associated with oversampling respondents who are interested in the topic when recruited by phone. This effect seems to be smallest for the probability sampled Internet panel, where respondents agree in advance to take a number of surveys on a range of topics.

Further research might investigate other modes, such as door-to-door surveying to compare the trade-off in inconsistency associated with location comfort against time constraint, as well as non-response characteristics. In addition, interest in environmental goods and issues could be assessed in surveys on unrelated topics to determine the extent of self-selection by topic even among the Internet panel. Such an assessment would serve to determine a baseline to measure how the level of interest affects participation.

What features are best suited for assembling a representative and unbiased sample? A survey mode for a study should have characteristics that lead to as representative a sample as feasible, including factors that maximize the ability to project the sample to a well-defined universe, minimizing self selection in participation by those with particular interest in or disinclination towards the survey topic, minimizing and compensating for the costs incurred by respondents to complete the survey task, and maximizing the ability of the respondent to be comfortable in the manner and location in which the survey is administered.

As Table 9 summarizes, the Internet panel best achieves these goals among the survey modes investigated. The availability of demographic information among panelists gives the Internet panel the ability to produce a sample that is nationally representative in terms of the most common demographic characteristics. While the fact that panelists agree in advance to take surveys on a variety of topics does not eliminate the possibility of self-selection, it minimizes this effect better than any of the other modes where potential respondents are informed in advance of the survey topic. The Internet panel also minimizes time and effort costs associated with completing the survey task by allowing respondents to complete the survey in their own home and at a time of their choosing.
Table 9: Participation Factors and the Performance among Survey Modes

| Ability of investigators to project to a well-defined universe | Internet Panel | Phone-Mail | Mall Intercept | Central Location |
|---------------------------------------------------------------|----------------|------------|----------------|------------------|
| Good. Though panelists must be recruited to the panel by phone, once members are generally willing to complete surveys, and the characteristics of invitees are available. | Fair. Households are difficult to reach by phone, and those who can more easily be reached may have different demographic characteristics than the US adult population. | Poor. Invitees are already present, but the demographics of mall visitors may be different than US population. | Fair. Households are difficult to reach by phone, and those who can more easily be reached have different demographic characteristics than the US adult population. |

| Interest by potential respondent in the topic | Internet Panel | Phone-Mail | Mall Intercept | Central Location |
|-----------------------------------------------|----------------|------------|----------------|------------------|
| Good. Since panelists already agree to take a variety of surveys, self-selection by topic is lessened. | Poor. Phoned invitees can opt in if particularly interested or opt out if they do not feel they are knowledgeable about the topic. | Poor. Invited shoppers can opt in if particularly interested or opt out if they do not feel they are knowledgeable about the topic. | Poor. Phoned invitees can opt in if particularly interested or opt out if they do not feel they are knowledgeable about the topic. |

| Total time and effort costs for respondents to complete the survey | Internet Panel | Phone-Mail | Mall Intercept | Central Location |
|-------------------------------------------------------------------|----------------|------------|----------------|------------------|
| Excellent. Invitation to participate sent by e-mail, survey completed in the home. | Fair. Invitation by phone, survey disk by mail, survey completed by most respondents at home, and completed survey returned by mail. | Good. Invited shoppers are already at the survey location. | Poor. Phoned invitees must travel to survey location. |

| Ability of respondent to be comfortable in the location where the survey is completed | Internet Panel | Phone-Mail | Mall Intercept | Central Location |
|--------------------------------------------------------------------------------------|----------------|------------|----------------|------------------|
| Excellent. Respondents complete the survey at a time of their convenience in their own home. | Excellent or Good. Most respondents complete the survey at a time of their convenience in their own home. Some might travel to a location with an available computer to complete the survey. All respondents must return the materials by mail. | Poor. Respondents complete the survey in an unfamiliar location at the time of the shopping trip. | Poor. Respondents complete the survey in an unfamiliar location at a scheduled time when the central location is open. |
Other survey modes perform well in regards one or more of these factors, but none perform overall as well as the Internet panel. For instance, while phone-mail allows most respondents to complete the survey in their home at a convenient time, for this survey it required additional software or require non-computer users to take the survey elsewhere. Mall intercept minimizes travel time since respondents are already at the survey location, but their lack of consistency may derive from feeling rushed by the interruption of their trip or lack of comfort in an unfamiliar survey environment.

Inconsistency in responses to survey questions can also indicate problems with a particular survey mode, as it indicates lack of attention and thoughtfulness towards the survey task. The mall intercept sample was most often associated with inconsistent responses in the survey, perhaps due to rushed or uncomfortable respondents. Phone-mail and the Internet panel modes, where respondents were most able to complete the survey at a time and place of their choosing, had the least such inconsistency.

Non-response, as measured using invited Internet panelists who declined to participate, is associated with a variety of demographic characteristics, and those characteristics are reflected to some extent in the make-up of each of the other survey modes. Those differences were most similar to the phone-mail mode, where respondents reached by random digit dialing as were the original recruitment of the Internet panel. Other modes had some such similarities, which were somewhat confounded by characteristics particular to each survey mode.

In general, it is clear that the choice of survey mode can affect the estimation of the value of an environmental good. The Internet panel seems to minimize such effects among survey modes investigated.

Acknowledgements

This article draws on research supported by a grant from EPA, Cooperative agreement # PI-83359201-0.

References and Notes

1. Dillman, D.; Phelps, G.; Tortora, R.; Swift, K.; Kohrell, J; Berck, J; Messer, B. Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet. *Soc. Sci. Research*. 2009, 38, 1-18.
2. Huber, J; Viscusi, W.K.; Bell, J. Reference Dependence in Iterative Choices. *Organizational Behavior and Human Decision Processes*. 2008, 106, 143-152.
3. Magat, W.; Huber, J.; Viscusi, W.K. with the assistance of Jason Bell. An Iterative Choice Approach to Valuing Clean Lakes, Rivers, and Streams. *J. Risk and Uncert.* 2000, 21, 7-43.
4. Martin, C. L. The impact of topic interest on mail survey response behavior. *Journal of the Market Research Society*. 1994, 36, 327–337.
5. Roberson, M. T.; Sundstrom, E. Survey design, return rates, and response favorableness in an employee attitude survey. *Journal of Applied Psychology*. 1990, 75, 354–357.
6. Rodes, A.; Sans, S.; Balana, L.L.; Paluzie, G. Recruitment methods and differences in early, late and non-respondents in the first MONICA-Catalonia population survey. *Rev Epidemiol Sante*. 1990, 38, 447-453.

7. Taylor, P.; Nelson, N.; Grandjean, B.D.; Anatchkova, B.; Aadland, D. Mode Effects and Other Potential Biases in Panel-based Internet Surveys: Final Report. 2009.

8. Viscusi, W.K.; Magat, W.; Huber, J., Pricing Environmental Health Risks: Survey Assessments of Risk-Risk and Risk-Dollar Tradeoffs for Chronic Bronchitis. *J. Environ. Econ. Manag*. 1991, 21, 32-51. Reprinted in Richard Carson, ed., The Stated Preference Approach to Environmental Valuation, Vol. 1: Foundations, Initial Development, Statistical Approaches (Burlington: Ashgate Publishing, 2007).