



Research article

Cost and performance tradeoffs between mail and internet survey modes in a nonmarket valuation study

Robert M. Campbell ^{a, c, *}, Tyron J. Venn ^b, Nathaniel M. Anderson ^c^a University of Montana, Economics, Liberal Arts Room 407, 32 Campus Dr., Missoula, MT 59812, USA^b University of the Sunshine Coast, School of Business, Sippy Downs, Qld 4556, Australia^c United States Forest Service, Rocky Mountain Research Station, 800 East Beckwith Ave., Missoula, MT 59801, USA

ARTICLE INFO

Article history:

Received 13 December 2016

Received in revised form

28 December 2017

Accepted 10 January 2018

Keywords:

Choice modeling

Cost effectiveness

Internet surveys

Survey mode

Willingness to pay

ABSTRACT

Using the results of a choice modeling survey, internet, mail-only and mixed internet and mail survey modes were examined with regards to their cost-effectiveness, representativeness, and willingness to pay (WTP). The topical focus of the study was biomass energy generation preferences of the residents of Montana, Colorado and Arizona, USA. Compared to the mail and mixed mode samples, the internet-only mode produced a sample of respondents that was younger, more likely to have a college degree, and more likely to have a household income of at least \$100,000 per year. However, observed differences in the characteristics of the collected sample did not result in significant differences in estimates of WTP. The internet survey mode was the most cost-effective method of collecting the target sample size of 400 responses. Sensitivity analysis showed that as the target number of responses increased the cost advantage of internet over the mail-only and mixed mode surveys increased because of the low marginal cost associated with extending additional invitations.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

Stated preference nonmarket valuation studies rely on obtaining responses to surveys that present hypothetical markets for environmental goods and services that are not traded in actual markets. Contacting potential respondents and providing them with a survey has traditionally been performed using in-person interviews, telephone interviews, and mail contact. As internet use has increased rapidly in the United States, internet-based survey methods have emerged as a viable method for data collection (Pew Research Center, 2016). Internet-based surveys offer a number of advantages including reduced response time, the ability to provide large amounts of information to respondents, and low marginal cost per response relative to other survey modes (Berrens et al., 2003). However, as a relatively new method with generally lower response rates, questions still exist about the representativeness of samples collected by internet surveys and the effects of this mode on willingness to pay (WTP) estimates. Furthermore, there are high

fixed costs associated with setting up internet-based surveys that can offset the benefits of low marginal costs if a sufficient number of responses are not received.

The purpose of this paper is to evaluate whether an internet-based survey is an appropriate cost-effective alternative to mail-only and mixed mail and internet survey modes for nonmarket valuation, while also meeting the need to collect a representative sample and produce unbiased estimates of economic measures of interest, such as WTP. Data used in the evaluation was collected in an experiment conducted as part of a choice modeling exercise investigating public preferences for renewable woody biomass energy in three states in the western United States (Campbell, 2016; Campbell et al., 2016). The emphasis here is to provide a clear comparison of the cost and performance tradeoffs of mail and internet-based survey modes for a choice modeling survey, and also to provide new evidence regarding the representativeness of an internet sample and the quality of WTP estimates derived from it.

The paper proceeds by first reviewing the environmental valuation literature that has compared internet-based surveys to other methods. Then we provide a brief overview of the study that generated the data used in this analysis, which is described in detail elsewhere by Campbell (2016) and Campbell et al. (2016, 2018). Next, the methods and results of the comparison of the three

* Corresponding author. University of Montana, Economics, Liberal Arts Room 407, 32 Campus Dr., Missoula, MT 59812, USA

E-mail addresses: robert3.campbell@umontana.edu (R.M. Campbell), tvenn@usc.edu.au (T.J. Venn), nathanielmanderson@fs.fed.us (N.M. Anderson).

survey modes are presented. Finally, the findings and their implications for practitioners are discussed.

2. Review of previous studies of survey modes

Dillman et al. (2014) described a high-quality survey as one that: a) provides a known opportunity for all members of the study population to be included in the sample, b) collects a sufficiently large sample of the population in a random fashion, c) encourages respondents to provide accurate information through well designed questions and information, and d) minimizes the probability that respondents to the survey differ systematically from people who choose not to respond. The degree to which these goals are met will determine the amount of error that is present in the data in the form of coverage error, sampling error, measurement error, and nonresponse error, respectively (Dillman et al., 2014). Coverage error occurs when the list from which sample is drawn does not accurately represent the population in ways that are important to the survey. Sampling error occurs as a result of surveying only some members of the sample frame. Nonresponse error occurs when nonrespondents differ from survey respondents in some way that influences estimates. Measurement error arises from unwillingness or inability of respondents to provide accurate answers. These potential sources of error can be difficult to disentangle from one another post hoc, and impossible to quantify individually without a study specifically design to do so. However, as they relate to this study, all sources of potential error have either been controlled for across survey modes or explored in the paper as outlined in Section 4.3.

The mode of a survey is defined as the method of administration used in data collection, and commonly includes in-person interviews, telephone interviews, or self-administration via mail or the internet. Survey modes may differ in their ability to minimize these sources of error as a result of differences in response rates, ability to collect a representative sample, effects on valuation estimates, and per-unit cost of obtaining usable responses. Previous research shows that internet-based surveys generally have been found to generate lower response rates than other contact methods, suggesting an internet sample may be more prone to nonresponse bias than other modes. Unsurprisingly, Marta-Pedroso et al. (2007) found higher response rates to in-person interviews (84%) than random internet contact (5.1%). Sinclair et al. (2012) found higher response rates for a random mail-survey contact (30.2% for personalized invitations and 10.5% for generic invitations) compared to a random contact internet survey (4.7% for personalized invitations and 2.2% for generic invitations).

Internet panels provide one potential solution to the problem to low internet response rates. Internet panels are groups of people that stand ready to participate in surveys, and consist of participants that are most often self-selected in response to some form of solicitation, or pre-recruited, sometimes based on a probability sampling design (e.g. Knowledge Networks, now known as GfK Knowledge Panel), and sometimes based on convenience samples¹. Both Lindhjem and Navrud (2011a) and MacDonald et al. (2010) found higher response rates for mail-contact than pre-recruited internet panels. The surprising result that an internet mode using panels of people who had already agreed to participate in surveys

failed to achieve higher response rates than a mail mode relying on random contact, speaks to the challenges of achieving high response rates with internet surveys. As outlined in Hays et al. (2015) it can be recruitment into the panels, rather than the response to an individual survey, that results in a lower effective response rate for panel-based internet surveys. However, panel methods can be improved. Olsen (2009) achieved a 63.6% response rate from a pre-recruited internet panel and 60.3% using a mail survey mode. Berrens et al. (2003), Schonlau et al. (2002), and Lindhjem and Navrud (2011b) provide detailed discussions of different types of internet panels and their relative attributes.

Although a large and growing proportion of households in the United States have access to the internet, the level of access differs between socioeconomic groups, with lower access amongst seniors, people with low educational attainment, and low household income (Perrin and Duggan, 2015). Also of concern is the ability to obtain responses from people who live in rural areas (Perrin and Duggan, 2015; Pew Research Center, 2015). This raises the concern that internet-based surveys may exacerbate the issues of coverage error that already exists with other survey modes in terms of collecting samples that are wealthier and better educated than the population as a whole. If the target population is the population as a whole and a representative sample cannot be collected, the preferences of the population may not be accurately estimated, and biased estimates of the economic values of interest may result. Published survey mode studies suggest that, on average, internet respondents tend to be younger, wealthier and better educated than mail and in-person interview respondents (Olsen, 2009; MacDonald et al., 2010; Windle and Rolfe, 2011). Mixed-mode sampling approaches (e.g. internet and mail sampling used together) have been suggested as a way to reach segments of the population that tend to have lower access to the internet (Champ, 2003). Mixed-mode surveys provide respondents with the option to respond via either internet, or mail, thus allowing respondents without internet-access or sufficient computer skills a means of participation. This is obviously only true if contact is made through the mail, not if contact is made only via an internet-based communication, such as email.

The purpose of nonmarket valuation surveys is to produce estimates of economic value for nonmarket goods and services. Like estimates of other parameters of interest, the magnitude and quality of results can be negatively impacted by the presence of nonresponse error from low response rates, coverage error from a non-representative sample, and measurement error associated with systematic variation in responses among survey modes. Research findings regarding the effect of survey mode on the magnitude and quality of valuation estimates are mixed. Some studies found no significant differences between internet and other survey modes (Covey et al., 2010; Fleming and Bowden, 2009; Lindhjem and Navrud, 2011a; Olsen, 2009). Bell et al. (2011) and Mjelde et al. (2016) on the other hand, both found that internet samples produce statistically significantly lower estimates of economic value than other survey modes. Olsen (2009) found lower estimation precision and lower certainty in choice (as measured through the variance of unobserved effects, and responses to debriefing questions for certainty) from an internet sample than one collected by mail. However, they also found a lower rate of protest responses (from zero bidders who were identified as protest responders in debriefing questions) than in the internet sample. Lindhjem and Navrud (2011a) found no evidence of difference in “don’t know” responses and protest responses between internet and face-to-face interviews. Based on their review of multiple studies that compared WTP estimates from internet surveys with other modes, Lindhjem and Navrud (2011b) concluded that there is little evidence to suggest that responses obtained from internet

¹ The internet survey mode in this study did not rely on internet panels. The stratified random sample was drawn from a sample frame of physical mail addresses and potential respondents were contacted via mail about participating in a single survey.

² According to the U.S. Environmental Protection Agency sensitive groups are defined as older adults and children, and persons with heart, lung or respiratory diseases (United States Environmental Protection Agency, 2017).

surveys are of lower quality (with regards to precision, respondent certainty, and prevalence of protest responses) than other modes.

Survey research is often conducted under a budget constraint that limits sampling intensity, which makes cost-effectiveness an important quality of any sampling design. The low marginal cost associated with sending additional invitations after fixed costs of designing and hosting the internet-based survey have been incurred, has been cited as a reason for the favorable cost-effectiveness of internet surveys (Berrens et al., 2003). The cost of an additional email invitation is close to zero, and the cost of sending an additional mail invitation to respond to an internet-based survey is also low compared to an additional contact for a mail-based survey that requires the printing and mailing of a survey booklet. As sample size increases, low marginal costs can overcome the disadvantage of high fixed costs and relatively low response rates associated with the internet mode. However, there are some important aspects of the internet mode that must be considered. Traditional mail and telephone surveys frequently rely on a sample frame generated from a database of known addresses or phone numbers, which are sometimes a matter of public record, or by using random digit dialing in the case of the telephone mode. Analogous databases of email addresses can be difficult to procure and may become quickly out of date, therefore researchers must use alternative methods to reach internet users. One option is to solicit internet responses by mail, which links known household addresses to internet responses (Campbell et al., 2016; Schonlau et al., 2002). As previously discussed, internet panels are also an option, but these have been critiqued for having low-response rates (Lindhjem and Navrud, 2011b; MacDonald et al., 2010).

Despite its practical importance in conducting surveys, cost is rarely discussed explicitly in studies that evaluate alternative modes. Fleming and Bowden (2009) conducted a travel cost survey using both an internet-based and a mail-based survey and found the internet-based survey to be more cost-effective in their collection of 640 responses. Survey cost-effectiveness research in other social science disciplines and in medical research have generally found internet-based surveys to be more cost-effective than mail-based surveys (Cobanoglu et al., 2001; Weible and Wallace, 1998). Schleyer and Forrest (2000) found that cost-effectiveness is dependent on sample size, with internet being more cost-effective than mail-only for target sample sizes greater than 275. Nevertheless, Sinclair et al. (2012) found that even with a large sample size, mail contact was the most cost-effective survey mode. The low cost of mail contact for Sinclair et al. (2012) is likely due in part to their choice of a single contact with no follow up, which is significantly less expensive than multiple contacts in the commonly employed Dillman et al. (2014) tailored-design method for mail-based surveys, which involves up to five separate mailings, including two that include the complete survey instrument.

To summarize, previous research findings reveal potential tradeoffs associated with choice of survey mode. Internet surveys offer potential gains in cost-effectiveness, which can lead to larger sample sizes and improved precision of estimates for a given budget. However, challenges reaching segments of the population without internet access may affect the ability to collect a representative sample, and generally lower response rates increase concerns of nonresponse error. Both of these issues may introduce sources of bias into data sets and affect the quality of estimates of

WTP. We used the results of a choice modeling experiment to examine these tradeoffs.

3. Methods

3.1. Case study background

Data to evaluate the impact of survey mode on sample characteristics, estimates of WTP and cost-effectiveness were provided by a previous study. Campbell et al. (2016, 2018) conducted a choice modeling survey to quantify preferences for woody biomass energy in Montana, Colorado and Arizona. The details of the study are beyond the scope of this paper and are covered thoroughly in these publications, but cost-effectiveness and mode comparison are not. The study design, survey modes and economic models used in the study are described here to provide context for the mode comparison.

There is significant interest in the interior western US in using forest biomass as a renewable energy source, especially when the biomass is generated by treatments to meet fire protection and forest restoration objectives (United States Forest Service, 2005). Such biomass can be used to produce heat and electricity, and potentially biofuels and bioproducts. Biomass generated from such treatments and used for these purposes is associated with higher renewable energy generation, lower risk of wildfire and better local air quality, but also higher energy bills. In order to determine which socioeconomic and environmental effects associated with woody biomass energy generation are most important to residents in this region, focus group meetings were held in Missoula, MT, Denver, CO, and Flagstaff, AZ in July through September of 2013. The meetings were attended by stakeholders from the United States Forest Service (USFS), state resource management agencies, universities, the forest industry, wildlife and land conservation groups, and local recreation groups. Members of these interest groups were familiar with the issues surrounding biomass energy and thought to be able to efficiently convey the primary concerns of their constituents and stakeholders, including citizens at large, in a facilitated discussion setting. In order gauge reaction and feedback from the general public, a pre-test of the survey instrument was conducted in which mall shoppers were intercepted and asked to complete the survey and provide their opinions. Based on feedback from the public, modifications were made to the survey to reduce complexity and increase clarity.

The five most important attributes associated with woody biomass energy identified at the focus group meetings were: the amount of woody biomass energy produced in the state (abbreviated to HOMES); unhealthy air days experienced locally (AIRDAYS); large wildfires in the state (WILDFIRES); forest health in the state (FORESTS); and household monthly energy bill (BILL). Each attribute was defined over a ten-year time horizon to provide a realistic time-frame in which to adopt and implement new forest management strategies, while also remaining relevant to respondents. The attributes are defined together with their status quo and alternative levels in Table 1. An example choice set is shown in Fig. 1. For detailed definitions of the attributes and description of the information presented in the survey, see Campbell (2016) and Campbell et al. (2016, 2018).

Attribute		Expected outcomes over 10 years		
		Strategy A	Current Strategy	Strategy B
Homes powered with wood in my state		30,000 homes	20,000 homes	10,000 homes
Unhealthy air days in my community		5 days per year	10 days per year	30 days per year
Large wildfires in my state		12 large wildfires per year	12 large wildfires per year	9 large wildfires per year
Forest health in my state		60% healthy forests	20% healthy forests	20% healthy forests
My household's monthly energy bill		\$200 (\$2,400 annually)	\$100 (\$1,200 annually)	\$120 (\$1,440 annually)
I would choose (select one only)		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Fig. 1. Example of choice set.

3.2. Description of survey modes

Data collection was conducted by the Bureau of Business and Economic Research at the University of Montana (BBER). A sample of 18,305 household addresses was obtained, with approximately one-third each from Arizona, Colorado and Montana. The study area was also stratified by air quality and forest ecoregion within each state. The sample was stratified to ensure coverage of people who live in forested areas and people who live in airsheds with a history of poor air quality because these characteristics were hypothesized to affect preferences toward the attributes of interest. Residents of forested areas were identified using US EPA level III Ecoregions (EPA, 2013). Poor air-quality airsheds were identified as EPA non-attainment airsheds, which have failed to meet national ambient air quality standards (EPA, 2013).

After being stratified by state, air quality and forest regions, entries in the sample frame were randomly assigned to one of three modes weighted by desired sample size and expected response rate, with 16,775 sent internet-only invitations, 511 sent mail-only invitations, and 1019 sent mixed-mode invitations. The mixed survey mode was administered as a potential method to alleviate sampling effects associated with the internet-based survey. The internet survey mode relied on mail contact to a physical mailing address, with an online response, and should not be confused with a completely web-based survey with an email-only sample frame and email-invitation. To be clear, all three modes used the same sample frame and all respondents, regardless of mode, were contacted by mail at a valid mailing address. Internet panels were not used, and were not evaluated in this study.

All potential respondents were contacted with an invitation

Table 1
Definitions of choice attributes and quadratic variables.

Variable	Definition	Levels	Units
HOMES	The amount of electric or thermal energy produced from woody biomass produced annually in the state, using residues from restoration treatments on public forests. Defined in terms of the number of homes that could be supplied with power.	10,000, 20,000 ^a , 30,000, 50,000	Homes per year
AIRDAYS	The number of days per year when air quality is unhealthy for sensitive groups in your community. ²	5, 10 ^a , 15, 30	Days per year
WILDFIRES	The number of wildfires per year that burn at least 1000 acres and threaten homes and watersheds in the state.	6, 9, 12 ^a , 15	Wildfires per year
FORESTS	The percent of healthy forestland in the state, across all forest ownership categories.	10, 20 ^a , 30, 60	Percent
BILL	Household average monthly energy bill in US dollars.	80, 100 ^a , 120, 150, 200, 400	US dollars

^a Indicates status quo attribute level.

letter mailed to their home explaining the purpose of the research and randomly presented with one of the following response options: (a) a web address and unique identification (ID) number that served as a required password to complete the survey online, (b) a notification that they would soon be receiving a physical survey packet in the mail, or (c) both a web address with ID number and the option to wait and receive a physical copy of the questionnaire in the mail if they did not respond online. Individuals in the online-only group (a) who had not completed the survey after about two weeks received a reminder post-card in the mail. Individuals in the other two survey groups (b and c) were contacted up to four times using the Dillman Tailored Design Method (Dillman et al., 2014). The four contacts proceeded with a second mailing that included the questionnaire, then a third mailing with a reminder postcard, and, if a response had still not been received, a fourth and final mailing that included a second hardcopy of the questionnaire. The layout and content of the survey was consistent across all three survey modes and each respondent was presented with 4 choice sets in each questionnaire. See Campbell (2016) and Campbell et al. (2016, 2018) for more details.

Two other characteristics of the survey mode must be considered. First, the mail-only respondents received a \$2 incentive (consistent with the incentivized version of the Dillman method), but this incentive was not included in the two other modes due to budget constraints. Therefore, the results produced by the mail and mixed survey modes should not be compared as two variations of the same method, with the only variation being whether or not an internet response option is provided. Rather, they should be viewed as two distinct survey contact modes with multiple differences. Second, because of the large number of Spanish speaking residents in Arizona and Colorado, for census tracts with at least 50% Hispanic population, respondents were provided with Spanish and English language versions of all mail and internet materials, including the option of completing a Spanish language version of the survey.

3.3. Econometric model

The choice modeling data in this study was analyzed using the multinomial logit model (MNL). With the MNL, the probability that an individual will select alternative i over alternative j , can be expressed as

$$P(i|C) = \frac{\exp(\mu V_i)}{\sum \exp(\mu V_j)} \quad (1)$$

where μ is a scale parameter inversely proportional to the variance of the error term. By assuming constant error variance, this parameter can be set to equal one (Ben-Akiva and Lerman, 1985). If however, error variance differs between data sets, respondents, or alternatives, scale heterogeneity may exist which can result in biased parameter estimates (Louviere et al., 2000). If scale heterogeneity is suspected, models such as the scale heterogeneity model (SMNL) or the generalized multinomial logit model (GMNL) can be used to account for the differences in error variance between groups or individuals (Greene and Hensher, 2010). Because the choice sets used in the study did not differ in complexity, scale heterogeneity was not expected to be a complication, supporting the use of MNL. If the scale parameter (μ) is assumed to be one, equation (1) can be expressed as,

$$P_n(i|C_n) = \frac{\exp(\beta_{ni}X_{ni} + \alpha C_{ni} + \tau Q_{ni} + \gamma R_n X_i)}{\sum \exp(\beta_{nj}X_{nj} + \alpha C_{nj} + \tau Q_{nj} + \gamma R_n X_j)} \quad (2)$$

X_n is a vector of terms for the attribute levels encountered by

individual n ; β_n is a vector of associated estimated coefficients; C_n is the cost attribute associated with each alternative and α is the associated coefficient; Q_n is an alternative specific constant (ASC), taking a value of 1 for status quo alternatives and zero otherwise, with an associated coefficient of τ ; R_n is a vector of case-specific socioeconomic characteristics and attitudinal variables, included to account for heterogeneity in preferences across respondents, and have an associated coefficient of γ ; and i and j are as previously defined. Socioeconomic characteristics and attitudinal variables were selected based on preliminary models and data exploration that revealed them to be significant predictors of respondent preferences toward the woody biomass energy attributes, which are discussed in the next section.³

In order to obtain policy relevant interpretations of the estimated coefficients, the marginal effects of each attribute must be calculated. Based on the model represented by equation (2), for attributes 1 through K the average household marginal willingness to pay (MWTP) for a one-unit improvement in the k th attribute can be estimated by equation (3)

$$MWTP = \left(\frac{\beta_n + \sum_{m=1}^M \gamma_{nm} G_m}{\alpha + \sum_{m=1}^M \theta_{nm} G_m} \right) \quad (3)$$

where G represents the fraction of the study area population that falls into each of the m socioeconomic or attitudinal categories and all other parameters are defined as above. Based on the method used by Han et al. (2008), equation (3) produces adjusted average household MWTP that corrects for the potential that survey respondents were not representative of the demographic characteristics of the study area as a whole.

The magnitude of WTP estimates is a common metric of comparison in the survey mode literature, and is used here to compare modes (Fleming and Bowden, 2009; Olsen, 2009; Covey et al., 2010; Bell et al., 2011; Lindhjem and Navrud, 2011a; and Mjelde et al., 2016). Without knowing the true value of WTP, comparison of the magnitude of WTP estimates produced using the different modes can provide an indication of whether or not there are mode-effects associated with the different survey modes, that is, whether or not researchers can expect different survey modes to provide different estimates of WTP. In addition, differences in magnitude of WTP estimates have policy implications when they are used to analyze prospective policy changes. Because one of the goals of this research is provide information that will inform decision making by practitioners, effects on the magnitude of welfare estimates that can have policy implications are a relevant metric of comparison.

4. Results

4.1. Response rates and sociodemographic characteristics

A total of 18,305 survey invitations were sent out, including 16,775 internet-mode invitations, 1019 mixed-mode invitations, and 511 mail-only invitations. The names and addresses for invitations were all drawn from the same sample frame in a stratified random sample, and randomly assigned to one of the modes. The survey effort yielded 1226 total complete returned surveys. As shown in Table 2, at 42% the mail-only survey mode had the highest effective response rate. The response rate for the mixed-mode was

³ Individual characteristics like sociodemographic and attitudinal characteristics are included in the model as interaction terms, multiplied by the levels of the attributes in each alternative. This is because they do not vary across alternatives like attribute levels do, and as a result would drop out of the model if included on their own.

Table 2
Response rates by mode.

	Internet	Mail	Mixed
Invitations sent	16,775	511	1019
Undeliverable invitations	1451	57	125
Delivered invitations	15,324	454	894
Complete responses	692	189	345
Overall response rate	4.5%	42%	39%
MT response rate	5.9%	54%	50%
CO response rate	4.5%	35%	36%
AZ response rate	3.1%	35%	29%
Urban response rate ^a	3.9%	34%	34%
Rural response rate ^a	4.3%	39%	29%

Notes:

^a Response rates for urban and rural residents cannot be compared to overall response rates or state response rates because urban and rural response rates are calculated using the total number of sent invitations, rather than the number of delivered invitations. The number of undeliverable invitations was not recorded across urban and rural residents.

39%, while the internet-mode had the lowest response rate of 4.5%. Of the 345 total responses to the mixed-mode invitations, 291 were completed with the mail hard-copy (84% of mixed mode responses) and 54 were completed on the internet (16%).

Contrary to expectations, the internet mode did not perform worse than the other modes at collecting responses from rural residents. People who live in rural areas responded at a slightly higher rate than urban residents to both the internet-only and mail-only survey modes. For the mixed-mode, urban residents responded at a higher rate than rural residents. Response rates were highest in Montana and lowest in Arizona, across all survey modes.

Overall, the survey respondents were on average older, better educated, wealthier, and more likely to be male than residents of the study area as a whole (Table 3). This was true for all three of the survey modes considered individually.

Using Pearson's chi-squared (χ^2) tests, significant differences were found between the sociodemographic characteristics of respondents to alternative survey modes. Not surprisingly, internet access amongst respondents in the internet survey mode was statistically significantly higher than amongst respondents to either the mail or mixed survey modes. Chi-squared (χ^2) tests revealed that the internet-only sample also contained a significantly larger proportion of high income earners and college educated individuals than the other survey modes. Although still substantially above the value of 14% seniors for the general population including children, at 32% seniors the internet sample was significantly younger than the mail and mixed modes, which were 39% and 40% seniors respectively. The sample from the mail survey was the most likely to over-represent men and climate change skeptics. The mixed-mode generated a sample similar to the mail-only sample, and did not provide a significantly more representative sample of the population compared to the internet, despite offering the ability to sample people without access to the internet.

4.2. Willingness to pay

The focus of WTP analysis in this paper is comparison of estimates across survey modes. For additional interpretation and policy analysis of WTP estimates, including aggregation for the population of the study area, see Campbell et al. (2016, 2018). Table 4 presents parameter estimates from three MNL models, estimated with data from each survey mode separately. Socio-demographic and attitudinal characteristics that vary across the survey modes account for some of the heterogeneity in choice that is not explained by the attribute levels. A reduced form of the

model, containing only the attribute levels and the ASC, was run on the dataset for each survey mode but those models are not presented here. The full models including all interaction terms were deemed preferable because they account for potential sources of heterogeneity arising from differences in sociodemographic characteristics between the samples collected with each survey mode, and provide a better statistical fit of the data (Likelihood Ratio test $p < .01$, for a pooled dataset and each survey mode individually). Using identical models for all three modes also facilitates comparison of the estimated coefficients.

Because of the interaction terms in the model, the coefficients on the attributes represent base-case preferences. The base case in these models are people who are younger than 65 years old, do not have a college education, make less than \$100 k per year and believe in man-made climate change. The attribute coefficients have the expected signs for all of the survey modes, but the statistical significance varies from mode to mode, possibly as a result of differences in standard errors, arising from differences in sample sizes. For the internet mode, all attributes except WILDFIRES are statistically significant. For the mail mode, FORESTS and BILL are statistically significant, while HOMES, AIRDAYS, and WILDFIRES are all statistically insignificant for the base case. Positive coefficients on FORESTS and HOMES, and the negative coefficients on AIRDAYS, WILDFIRES, and BILL are consistent with expectations that increases in the level of HOMES and FORESTS increase likelihood of an alternative being selected, while increases in AIRDAYS, WILDFIRES, and BILL, decrease the likelihood of an alternative being selected. Although no external test of scope was conducted, these results suggest that respondents were willing to pay more for larger quantities than they were for smaller quantities, thus exhibiting sensitivity to scope.⁴

For the mixed mode, AIRDAYS, FORESTS, and BILL are statistically significant, while HOMES and WILDFIRES are statistically insignificant. The interaction terms revealed that, regardless of survey mode, college educated respondents had statistically significantly higher WTP to avoid unhealthy air days. College educated respondents also had higher WTP to increase the proportion of healthy forestland in their state and climate change skeptics had lower WTP in general than other respondents, although both of these were not always significant.

Despite the differences in the collected samples between the survey modes, there are no significant differences in MWTP between the survey modes. Table 5 reports the average monthly household MWTP for each attribute across survey mode, estimated using equation (3). A 95% confidence interval for each choice attribute was estimated with 500 bootstrap repetitions using the method described by Efron and Tibshirani (1986). While the mean values of MWTP vary somewhat between the survey modes, in all cases the 95% confidence intervals overlap, providing evidence that the mean values are not significantly different. As a statistical test of differences in MWTP estimates across survey modes t-tests were conducted for each attribute and the ASC. The independent null hypotheses were: $H_0: MWTP_{\text{internet}} = MWTP_{\text{mail}}$, $MWTP_{\text{internet}} = MWTP_{\text{mixed}}$, and $MWTP_{\text{mail}} = MWTP_{\text{mixed}}$. All test

⁴ A criticism sometimes made of stated preference valuation is that results do not show sensitivity to the magnitude or scope of the good being valued (for more on this, see Haab et al. (2013) and Whitehead (2016)). Although not conducted in this study, external tests of scope can be conducted by splitting the sample, differing the quantity of change in the environmental good presented to the two groups, and comparing to see if WTP is sensitive to changes in quantity. For an example of a split-sample scope test employed in a choice experiment, see Lew and Wallmo (2011). Scope tests are one type of validity test, which are used to assess how successfully the estimated value measures the theoretical construct under investigation (Brown, 2003).

Table 3
Mean value of sociodemographic characteristics by mode and study area population.

Characteristic	Definition	Internet	Mail	Mixed	Population ^b	Test statistic
MALE	Male respondents	62%	65% ^a	62%	50%	$\chi^2 = 9.73, p < .01$
SENIOR ^c	Individuals who are 65 years old or older	32% ^a	39%	40%	14%	$\chi^2 = 87.53, p < .01$
HIGHINC	Individuals with annual household income > \$100 k	25% ^a	24%	24%	20%	$\chi^2 = 33.03, p < .01$
COLLEGE	Individuals with at least a bachelor's degree	61% ^a	47%	47%	31%	$\chi^2 = 287.84, p < .01$
INTERNET ACCESS ^e	Individuals that have internet access	98% ^a	91%	90%	74% ^d	$\chi^2 = 456.00, p < .01$
SKEPTIC	Individuals who do not believe in man-made climate change	47%	55% ^a	47%	49% ^f	$\chi^2 = 41.54, p < .01$

Notes:

- ^a Indicates statistically significant difference from sample mean of the other survey modes.
^b Based on the weighted average of the populations of Arizona, Colorado, and Montana. Source: [Census Bureau \(2010\)](#).
^c Senior citizens are defined as age 65 and older.
^d State-specific census data was only available for high speed internet access, while the survey did not specify high-speed or not. Nationally, the rate of high-speed internet access is only one percentage point lower than the rate of access to any type of internet access, so these numbers should be comparable. Source: [File and Ryan \(2014\)](#).
^e Proportion of household that own a computer and have internet access.
^f Source: [Yale Project on Climate Change Communication \(2014\)](#).

results fail to reject the null hypotheses of equality of means at a 10% confidence level (for all tests, $p < .10$), providing no evidence of statistically significant differences in the MWTP among the three survey modes. Although no formal tests of precision were performed, the confidence intervals are generally tighter for the estimates from the internet sample than for the other survey models, which is likely as a result of the larger sample collected with the internet survey. Across all survey modes, the magnitude of the coefficients on the ASC is large relative to the other attributes, indicating a strong preference for the status quo (Table 5). However, the coefficients are not significantly different from zero for the mail and mixed-mode surveys.

4.3. Sources of survey error

In this study, all three survey modes relied on the same sample frame and probability-based sampling procedures, thus controlling

for differences in noncoverage error or sampling error. As a test for the presence of a survey mode effect on preferences, the full model with an added survey mode variable was run on the pooled data set including all three modes. Results from this model did not reveal any statistically significant effect of survey mode on preferences toward the attributes. The low response rate expected from the internet survey relative to the mail and mixed modes suggests elevated potential for nonresponse bias. Although we were unable to quantify nonresponse error (through a follow-up survey of actual nonrespondents, for example), to assess the potential for nonresponse error, the characteristics of late-responders were compared with those who responded earlier, based on the assumption that late-responders are more similar to non-responders than they are to early responders (Armstrong and Overton, 1977). This is not equivalent to quantifying nonresponse error, but does provide some useful information. Late-responders were identified as those who responded after receiving the reminder post card, and they

Table 4
Regression results, MNL interactions model.

	Internet		Mail		Mixed	
	Estimate	SE	Estimate	SE	Estimate	SE
HOMES	0.00987**	0.00459	0.0106	0.00673	0.00618	0.00592
AIRDAYS	-0.0505***	0.00966	-0.0275	0.0174	-0.0307**	0.0120
WILDFIRES	-0.0254	0.0220	-0.0258	0.0472	-0.0340	0.0284
FORESTS	0.0379***	0.00363	0.0277***	0.00697	0.0285***	0.00453
BILL	-0.00366***	0.000890	-0.00355**	0.00156	-0.00388***	0.00121
ASC	0.298***	0.0587	0.468***	0.104	0.287***	0.0902
SKEPTIC X HOMES	-0.00933**	0.00412	0.00566	0.00782	-0.00282	0.00659
SKEPTIC X AIRDAYS	0.0379***	0.00908	0.00478	0.0176	-0.0185	0.0144
SKEPTIC X WILDFIRES	0.0288	0.0222	0.0603	0.0460	-0.0172	0.0341
SKEPTIC X FORESTS	-0.0149***	0.00341	0.00923	0.00710	-0.0145***	0.00523
SKEPTIC X BILL	-0.00218**	0.000939	-0.00248	0.00189	-0.00160	0.00146
HIGHINC X HOMES	0.00847*	0.00463	-0.0164	0.0117	0.00508	0.00776
HIGHINC X AIRDAYS	-0.00581	0.0107	-0.0121	0.0322	-0.0122	0.0208
HIGHINC X WILDFIRES	-0.0274	0.0262	-0.106*	0.0551	0.00939	0.0398
HIGHINC X FORESTS	-0.000636	0.00391	0.00268	0.0102	0.0183**	0.00812
HIGHINC X BILL	0.000401	0.00104	0.000185	0.000279	-0.000720	0.00185
COLLEGE X HOMES	0.00137	0.00467	0.0100	0.00906	0.00535	0.00655
COLLEGE X AIRDAYS	-0.0266***	0.00936	-0.0401*	0.0209	-0.0264*	0.0150
COLLEGE X WILDFIRES	-0.00583	0.0232	0.00518	0.0538	0.0120	0.0357
COLLEGE X FORESTS	0.00291	0.00359	0.0153*	0.00845	0.0133**	0.00519
COLLEGE X BILL	-0.00104	0.000954	-0.00248	0.00224	-0.000902	0.00148
SENIOR X HOMES	-0.000234	0.00447	-0.0107	0.00834	-0.00771	0.00688
SENIOR X AIRDAYS	0.00645	0.00953	-0.0106	0.0156	0.0156	0.0140
SENIOR X WILDFIRES	-0.0430*	0.0240	0.00587	0.0475	-0.0200	0.0348
SENIOR X FORESTS	-0.00334	0.00362	-0.00688	0.00712	0.000420	0.00533
SENIOR X BILL	-0.000945	0.000999	0.00266	0.00187	-0.00135	0.00151
N	7620		1956		3492	
Log-likelihood	-2367.1		-557.6		-1042.3	

Note: * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 5
WTP by Survey mode, MNL Interactions model.

Attribute	Internet		Mail		Mixed	
	Average household MWTP (\$)	95% confidence interval (\$)	Average household MWTP (\$)	95% confidence interval (\$)	Average household MWTP (\$)	95% confidence interval (\$)
HOMES	1.47	0.38 2.56	2.29	0.31 4.27	1.22	-0.28 2.72
AIRDAYS	-8.03	-10.53 -5.52	-8.15	-14.27 -2.02	-9.05	-12.98 -5.12
WILDFIRES	-4.88	-9.89 0.13	-3.08	-14.75 8.59	-7.49	-14.21 -0.77
FORESTS	6.11	4.62 7.60	7.14	3.35 10.93	5.59	3.71 7.48
ASC	81.39	19.55 143.23	131.88	-960.5 1224.2	73.96	-23.05 170.96

Table 6
Mean value of sociodemographic characteristics, late-responders and non-late responders.

Characteristic	Late- responders	Not-late responders	Test statistic
MALE	52%	69%	$\chi^2 = 438.4, p < .01$
SENIOR	31%	38%	$\chi^2 = 77.2, p < .01$
HIGHINC	22%	26%	$\chi^2 = 39.2, p < .01$
COLLEGE	48%	59%	$\chi^2 = 156.9, p < .01$
SKEPTIC	50%	47%	$\chi^2 = 18.0, p < .01$

represent 40% of respondents. Late-responders were compared to the 60% of the respondents who responded before the reminder post-card. Comparison of sociodemographic and attitudinal characteristics revealed that late-responders differed significantly from other respondents in some ways (Table 6). Late-responders were significantly less likely to be male, senior citizens, high income earners, or have a college degree. They were also significantly less likely to believe in anthropogenic climate change, or believe that public forests are in need of restoration, which suggests that the survey topic may have resonated less with late-responders.

Based on this information, there is evidence that non-respondents could potentially differ from respondents in some ways, but no statistically significant differences were found between the preferences of late-responders and others in terms of MWTP (95% confidence intervals overlap for all attributes, Table 7). Therefore, if late-respondents share similarities with non-respondents, nonresponse error may not have a significant effect on estimates of MWTP in this case. However, a post hoc analysis of nonresponse error was not carried out to confirm this possibility.

With regards to protest responses, the number of respondents who selected the status quo option for every choice set (which may be a true preference for the status quo, but sometimes represents a protest response) were compared across survey modes, and represented similar proportions of total responses. The number of respondents who always selected the status quo for mail-only, internet-only and mixed mode were 13 (6.9%), 46 (6.6%), and 22 (6.4%), respectively. Results of a χ^2 test reveal no statistically significant difference in the proportion of respondents who always selected the status quo between the survey modes.

4.4. Cost-effectiveness

In order to allow a level cost comparison between the survey modes, we applied a common response target of 400 responses per survey mode and estimated the cost per-response for each mode at the 400 responses level.⁵ The number of invitations that would

need to be sent to obtain 400 responses was estimated based on the actual response rate achieved by each survey mode in the choice modeling survey.

All unit costs of materials and labor adopted in the cost-effectiveness analysis, as well as proportions of respondents receiving the second mailing of the questionnaire, are actual costs and proportions associated with the choice modeling survey. Costs are categorized as either survey costs, which could potentially be contracted out to a survey company or other organization (such as BBER in this study) or analytical costs (associated with research functions). Analytical costs are assumed to be constant across all survey modes and are omitted from the analysis. Survey costs were classified as: 1) printing and mailing costs, 2) labor costs, or 3) purchase of the sample address list. Printing and mailing costs include four contact mailings for both the mail-only and mixed modes, and two contacts for the internet-only mode. Costs in the second and fourth mailings for mail-only and mixed modes include postage for the mail packet, copies of the 16-page color survey, and the return envelope and return. Mailing costs have been adjusted to account for actual costs and do not assume all people received the maximum number of mailings. One hundred percent of people in the mixed and mail-only modes received the first three contacts, and 92% received a second questionnaire in the fourth mailing. One hundred percent of people in the internet-only group received two contacts. The cost of including the \$2 bill cash incentive for the mail-only mode was included in the cost of the second mailing.

Three categories of labor costs were included: a) administrative and clerical costs associated with data collection, including creation of a plan to administer the survey, assembly and mailing of contact materials, and collection of completed questionnaires and associated data entry; b) development of the online survey for the internet-only and mixed survey modes and associated web hosting; and c) Spanish language translation costs (Table 8). Labor costs associated with data collection are variable and increase with the number of invitations sent out. Online survey development and Spanish translation costs, on the other hand, are fixed because there is zero marginal cost associated with an increase in the number of invitations.

Conducting a bi-lingual survey incurred significant added costs. Packets sent to census tracts that were 50% Hispanic or higher included both an English and Spanish language version of the questionnaire. Therefore, double the number of questionnaires were printed and higher postage was paid for each household that received both version in the mail-only and mixed modes. In contrast, accommodating two languages with the internet-only mode invitation required only that contact materials be printed double-sided with a different language on each side. In addition, a Spanish language web-based survey was developed to serve the mixed and internet-only modes.

The purchase of the address list is fixed at \$500 for up to 1200 addresses, and then costs increase at the marginal rate of \$0.09 for each additional address beyond 1200. The cost of researcher time to design the study, including the statistical design, and to develop the

⁵ The number of responses needed to achieve a desired level of confidence in parameter estimates is a function of the acceptable level of sample error and the size of the population being sampled. For populations of more than 1 million people and a target sample error of $\pm 5\%$, 95% confidence levels should be attainable with 384 responses (Dillman, 2007), which was rounded to 400 for this study.

Table 7
Marginal Willingness to Pay for late-responders and non-late responders.

Attribute	Late-Responders			Non-Late Responders		
	Average household MWTP (\$)	95% confidence interval (\$)		Average household MWTP (\$)	95% confidence interval (\$)	
HOMES	0.92	-0.19	2.02	1.77	1.01	2.52
AIR DAYS	-8.36	-11.21	-5.52	-8.84	-10.74	-6.93
WILDFIRES	-7.69	-13.09	-2.30	-6.02	-9.98	-2.06
FORESTS	5.39	3.97	6.80	6.21	5.17	7.24
ASC	63.73	29.61	97.85	54.73	33.00	76.46

Table 8
Summary of cost components included in each mode. Analytical costs, such as study design, sample selection, and data analysis are assumed to be equal for all modes and are excluded from the cost comparison.

Cost category	Sub-category	Mode		
		Internet	Mail	Mixed
Printing and mailing		2 contacts	4 contacts	4 contacts
Cash incentive		None	\$2	None
Purchased sample frame (contacts)		18,000	1200	1200
Labor costs	Administration	Yes	Yes	Yes
	Translation	Yes	Yes	Yes
	Programming	Yes	No	Yes
	Web hosting	Yes	No	Yes

survey materials is assumed to be the same for all modes and has been excluded from the analysis. These costs are distinct from survey costs because they are generally less likely to be contracted out to a survey company, especially for research purposes as opposed to marketing and other applications.

Sensitivity analyses of survey costs were conducted to test the robustness of the results to changes in response rates, and target number of respondents. Sensitivity analysis was conducted for each survey mode by changing one of these factors at a time ($\pm 20\%$ and $\pm 50\%$ from base case levels), while holding other factors fixed, to analyze the effect that these important survey parameters have on

cost-effectiveness. A third sensitivity analysis was performed on the proportion of households that received both English and Spanish language versions of the materials.

Based on detailed survey cost records from the case study, [Table 9](#) reports the cost to achieve 400 completed survey responses for each survey mode. Results from the cost comparison reveal internet-only to be the most cost-effective of the survey modes, with a cost of \$61 per response. Mail-only was the second most cost-effective survey mode, with a cost of \$70 per response. At \$90 per response, the mixed survey mode was the least cost-effective option.

Table 9
Survey implementation costs by mode, target sample of 400.

	Internet	Mail	Mixed
Response Information			
Response rate	4.5%	42%	39%
Number of Invitations, to achieve 400 usable responses	8889	962	1026
Mailing Costs			
1st Contact mailing	\$5764	\$624	\$665
2nd Contact mailing ^a	n/a	\$11,051	\$9736
3rd Contact mailing	\$3771	\$408	\$435
4th Contact mailing	n/a	\$8397	\$8957
Total Mailing Costs	\$9535	\$20,479	\$19,793
Labor Costs			
Sample design	\$716	\$716	\$716
Admin & clerical ^b	\$2528	\$5141	\$4716
Website design & hosting	\$9410	n/a	\$9410
Spanish Translation	\$1000	\$1000	\$1000
Total Labor Costs	\$13,648	\$6857	\$15,837
Sample Address Costs			
First 1200	\$500	\$500	\$500
After the first 1200	\$688	0	0
Total Sample Address Costs	\$1188	\$500	\$500
Total Costs	\$24,371	\$27,836	\$36,130
Cost per Response ^c	\$61	\$70	\$90

Notes:

^a Second mailing contact costs were lower for the internet because some people completed the internet version of the survey before the second mailing.

^b Differences in admin and clerical costs between survey modes arise as a result of more labor being required to assemble and mail survey packets, and manually enter data from returned mail surveys, with the most being required for the mail-only survey mode.

^c Cost per response is total costs divided by the target response number of 400.

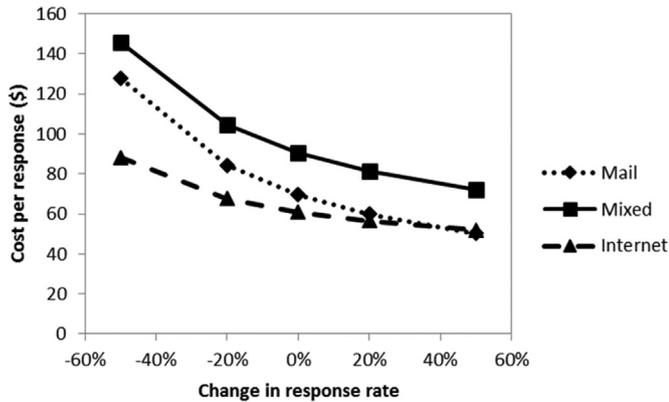


Fig. 2. Sensitivity of cost per response to response rate.

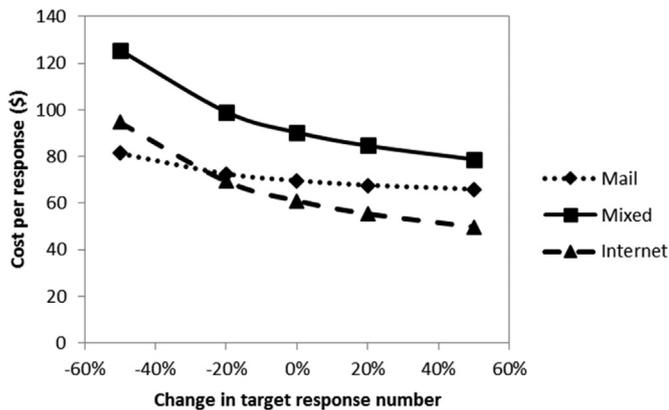


Fig. 3. Sensitivity of cost per response to target response number (base case is 400 for all survey modes).

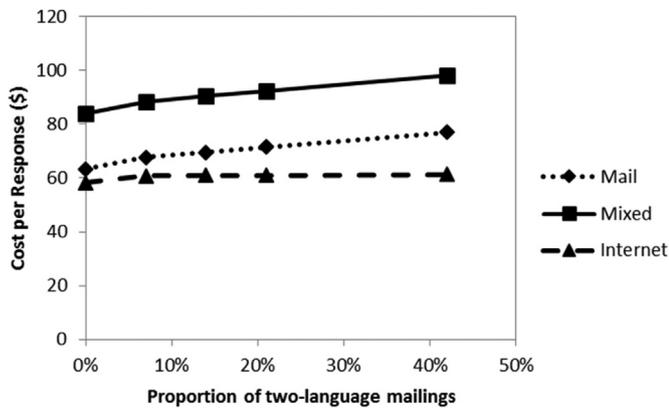


Fig. 4. Sensitivity of cost per response to proportion of two-language mailings (base case is 14% for all survey modes).

Results from analysis of the sensitivity of cost per usable response to key survey parameters are displayed in Figs. 2–4. The base case parameter values for each survey mode in Fig. 2, can be found in Table 9. The base case parameter value in Figs. 3 and 4 are the same for all survey modes, namely 400 responses and 14% of invitations with a Spanish language option, respectively.

The sensitivity analyses revealed that the finding that the mixed-mode is the least cost-effective is robust against changes in the levels of parameters on which the sensitivity analyses were conducted. This is unsurprising given the combination of high fixed website design labor costs and high mailing costs. In addition, the response rate was lower than the mail-only mode (39% versus 42%,

respectively).

With respect to both response rate and target response number, there is a point at which mail contact becomes more cost-effective than internet-only (Figs. 2 and 3). For response rates 50% higher than the base case for each survey mode,⁶ the cost per response achieved by the mail-only mode becomes smaller than for internet-only; \$50 versus \$52, respectively (Fig. 2). Given a target number of responses of 400, Fig. 2 does indicate that the cost per response for the mail-only mode would be equivalent to the base case costs for internet-only with a 20% improvement in response rate to 50.4%. At a target of 200 responses, 50% below the base case, the cost per response for mail-only is lower than for the internet-only mode (Fig. 3). However, Fig. 3 also reveals that as the target response number increases, the cost advantage of internet over mail becomes greater.

Sensitivity analysis with respect to the proportion of households receiving both English language and Spanish language materials revealed that the cost advantage of the internet mode over the mail mode becomes larger as the proportion of households that receive both language versions of the survey increases (Fig. 4). The inclusion of a Spanish language option increases the costs of printing and mailing much more for mail than for internet surveys.

5. Discussion

This study compared internet-only, mail-only and mixed (internet and mail) survey modes for a survey estimating public preferences toward woody biomass energy in Arizona, Colorado and Montana. The evaluation was made on the basis of: (a) how representative the sociodemographic characteristics of the sample are relative to the population, (b) whether there are differences in MWTP between the survey modes; and (c) the cost-effectiveness of alternative modes in terms of total cost and cost per usable response.

Comparison of the sociodemographic profiles of the samples collected by the different survey modes reveals that the internet mode collected a sample that was significantly wealthier and more highly educated than the mail and mixed-mode samples, and was farther from the mean value of the study area population for these characteristics. However, the internet sample was more representative than the other modes in terms of age by having a significantly lower proportion of seniors. These findings are consistent with other studies that have found internet samples to be younger, more highly educated, and wealthier than mail samples (Olsen, 2009; MacDonald et al., 2010). Based on these results, it is not clear that one survey mode produced a sample that is more representative of the population than the samples collected by the other survey modes.

Despite some statistically significant differences in demographics (Table 3), based on qualitative comparison of 95% confidence intervals and associated t-tests, there was no evidence of statistically significant differences in MWTP for choice attributes between the survey modes. The lack of significant differences in MWTP between the survey modes means that no evidence was found to suggest the presence of a survey mode effect on estimates of MWTP from differences in measurement error between the internet and other survey modes. Furthermore, concerns over nonresponse error associated with the relatively low response rate of the internet mode appear to be alleviated by a lack of significant differences in MWTP estimates between modes, and between late-responders and other respondents; however, the survey did not

⁶ This is equivalent to a mail-only survey response rate of 63% and an internet-only response rate of 6.8%.

explicitly quantify nonresponse error using follow-up survey of true nonrespondents. Because all internet responses were solicited by mail invitations to valid home addresses from the same sample frame as the other modes, concerns about differences in coverage and sampling error were minimized. Because there were significant differences between the samples collected by each of the three survey modes, standard approaches were used to account for potential differences in preferences between the collected samples and the population (weighting by population characteristics in the calculation of MWTP) and among the three collected samples (inclusion of sociodemographic control variables in the MNL model). However, findings by previous studies on accounting for preference heterogeneity between survey modes are mixed. Olsen (2009), who accounted for preference heterogeneity with a random parameters model specification, found no significant differences in WTP between internet and mail survey modes. Bell et al. (2011), however, did find significant differences in economic measures between internet and mail survey modes, even when accounting for sociodemographic characteristics in a two-tailed Tobit regression analysis. Our case study provided no evidence in favor of one survey mode over another on the basis of sample error.

The internet survey mode was found to be the most cost-effective of the three modes examined. This facilitates collection of a larger sample for a given budget constraint, which may result in more precise estimates of MWTP compared to smaller samples. It is worth reiterating here that we did not evaluate the use of internet panels, nor surveys conducted fully online in which solicitation is made via email or text message to mobile devices.

The mixed survey mode has some inherent attraction in that it provides respondents with a choice of how to take the survey, potentially increasing response rates, but it was the least cost-effective and, given the lower response rate than mail-only mode and the statistically insignificant differences in MWTP between survey modes (based on comparison of 95% confidence intervals and t-tests), the potential benefits of using a mixed survey mode were not found to make it a preferable option in this study. A caveat is required in comparing the mail-only and mixed modes. The mail-only survey mode garnered the highest response rate, followed by the mixed-mode, with the internet survey producing the lowest response rate. However, the lower response rate for mixed-mode than the mail-only may be a result of the \$2 incentive that was provided in the mail-only contact material, and not a result of some inherent attribute of mixed mode surveys. Incentives have been shown to produce higher response rates (Mooney et al., 1993).

As a result of the low marginal cost of extending additional invitations once the fixed costs of setting up the internet survey have been incurred, the cost savings for the internet mode increase as the target number of responses increases. The cost advantages of the internet survey mode are even larger if a multi-lingual approach is required. Sensitivity analyses highlighted that mail-only surveys are more cost-effective than internet surveys when the target number of respondents is small and when the response rate is high. This is due to the relatively high fixed costs associated with setting up internet survey web pages for a small number of responses, versus the relatively low fixed costs but relatively high marginal costs of additional invitations for the mail mode. Fig. 2 suggests the response rate would have to be about 63% (+50% from the base case) at 400 responses for the mail-only survey mode to be more cost-effective than internet-only.

6. Conclusions

Based on a comparison of response rates and sociodemographic characteristics of respondents, willingness to pay estimates, and the cost-effectiveness of sample collection, the internet survey

mode was found to be the preferred survey mode for collecting stated preference nonmarket valuation data. The internet mode is the most cost-effective of the three survey modes, offering the ability to collect a larger sample for a given budget so long as the target sample is at least 300. Although some significant differences in the characteristics of the collected sample were found between the survey modes, estimates of MWTP from samples collected using these three alternative modes were not significantly different, supporting the use of self-administered internet surveys in choice modeling and potentially other nonmarket valuation research.

Acknowledgements

This project was supported by the Biomass Research and Development Initiative, Competitive Grant no. 2010-05325, from the U.S. Department of Agriculture (USDA), National Institute of Food and Agriculture (NIFA). Additional support was provided by the USDA NIFA through AFRI-CAP Competitive Grant 2013-68005-21298 Bioenergy Alliance Network of the Rockies, and by the U.S. Forest Service.

References

- Armstrong, J., Overton, T., 1977. Estimating nonresponse bias in mail surveys. *J. Market. Res.* 14, 396–402.
- Bell, J., Huber, J., Viscusi, W.K., 2011. Survey mode effects on valuation of environmental goods. *Int. J. Environ. Res. Publ. Health* 8 (4), 1222–1243.
- Ben-Akiva, M., Lerman, S.R., 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press, Cambridge, MA.
- Berrens, R.P., Bohara, A.K., Jenkins-Smith, H., Silva, C., Weimer, D.L., 2003. The advent of internet surveys for political research: a comparison of telephone and internet samples. *Polit. Anal.* 11 (1), 1–22.
- Brown, T.C., 2003. In: Champ, P., Boyle, K., Brown, T. (Eds.), *Introduction to Stated Preference Methods*. Kluwer Academic Publishing, The Netherlands.
- Campbell, R.M., 2016. *Evaluation of Social Preferences for Woody Biomass Energy in the U.S. Mountain West*. University of Montana, Missoula, MT, p. 10883. PhD Dissertation. <http://scholarworks.umt.edu/etd/10883>.
- Campbell, R.M., Venn, T.J., Anderson, N.M., 2016. Social preferences toward energy generation with woody biomass from public forests in Montana, USA. *For. Pol. Econ.* 73, 58–67.
- Campbell, R.M., Venn, T.J., Anderson, N.M., 2018. Heterogeneity in preferences for woody biomass energy in the US Mountain West. *Ecol. Econ.* 145, 27–37.
- Census Bureau, 2010. *United States Census 2010 Data*. United States Census Bureau.
- Champ, P.A., 2003. Collecting survey data for nonmarket valuation. A primer on nonmarket valuation. In: Champ, P., Boyle, K., Brown, T. (Eds.), *Kluwer Academic Publishing, The Netherlands*.
- Cobanoglu, C., Warde, B., Moreo, P., 2001. A comparison of mail, fax and web-based survey methods. *Int. J. Market Res.* 43 (4).
- Covey, J., Robinson, A., Jones-Lee, M., Loomes, G., 2010. Responsibility, scale and the valuation of rail safety. *J. Risk Uncertain.* 40 (1), 85–108.
- Dillman, D., Smyth, J., Christian, L., 2014. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. John Wiley & Sons, Hoboken, NJ.
- Dillman, D., 2007. *Mail and Internet Surveys: The Tailored Design Method*. John Wiley & Sons, Hoboken, NJ.
- Efron, B., Tibshirani, R., 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statist. Sci.* 1 (1), 54–77.
- EPA, 2013a. *Ecoregions of North America*. United States Environmental Protection Agency.
- EPA, 2013b. *SIP Status and Information*. United States Environmental Protection Agency.
- File, T., Ryan, C., 2014. *Computer and Internet Use in the United States: 2013*. United States Census Bureau, American Community Survey.
- Fleming, C.M., Bowden, M., 2009. Web-based surveys as an alternative to traditional mail methods. *J. Environ. Manag.* 90 (1), 284–292.
- Greene, W., Hensher, D., 2010. Does scale heterogeneity across individuals matter? An empirical assessment of alternative logit models. *Transportation* 37, 413–428.
- Haab, T., Interis, M., Petrolia, D., Whitehead, J., 2013. From hopeless to CUrious? Thoughts on Hausman's "dubious to hopeless" critique of contingent valuation. *Appl. Econ. Perspect. Pol.* 35 (4), 593–612.
- Han, S.-Y., Kwak, S.-J., Yoo, S.-H., 2008. Valuing environmental impacts of large dam construction in Korea: an application of choice experiments. *Environ. Impact Assess. Rev.* 28 (4–5), 256–266.
- Hays, Ron D., Liu, Honghu, Kapteyn, Arie, 2015. Use of internet panels to conduct surveys. *Behav. Res. Methods* 47 (3), 685–690.
- Lew, D., Wallmo, K., 2011. External tests of scope and embedding in stated preference choice experiments: an application to endangered species valuation. *Environ. Resour. Econ.* 48 (1), 1–23.

- Lindhjem, H., Navrud, S., 2011a. Are Internet surveys an alternative to face-to-face interviews in contingent valuation? *Ecol. Econ.* 70 (9), 1628–1637.
- Lindhjem, H., Navrud, S., 2011b. Using internet in stated preference surveys: a review and comparison of survey modes. *Int. Rev. Environ. Resour. Econ.* 5.
- Louviere, J., Hensher, D., Swait, J., 2000. *Stated Choice Methods: Analysis and Application*. Cambridge University Press, Cambridge.
- MacDonald, D.H., Morrison, M.D., Rose, J., Boyle, K., 2010. Untangling differences in values from internet and mail stated preference studies. In: *World Congress of Environmental and Resource Economics*. Montreal, Canada.
- Marta-Pedroso, C., Freitas, H., Domingos, T., 2007. Testing for the survey mode effect on contingent valuation data quality: a case study of web based versus in-person interviews. *Ecol. Econ.* 62 (3–4), 388–398.
- Mjelde, J., Kim, T.-K., Lee, C.-K., 2016. Comparison of internet and interview survey modes when estimating willingness to pay using choice experiments. *Appl. Econ. Lett.* 23 (1).
- Mooney, G., Giesbrecht, L., Shettle, C., 1993. To pay or not to pay: that is the question. In: *Meeting of the American Association for Public Opinion Research*. St. Charles, IL.
- Olsen, S., 2009. Choosing between internet and mail survey modes for choice experiment surveys considering non-market goods. *Environ. Resour. Econ.* 44 (4), 591–610.
- Perrin, A., Duggan, M., 2015. *Americans' Internet Access: 2000–2015*. Pew Research Center.
- Pew Research Center, 2015. *Coverage Error in Internet Surveys*.
- Pew Research Center, 2016. *Internet Surveys*. (Accessed 15 June 2016).
- Schleyer, T., Forrest, J., 2000. Methods for the design and administration of web-based surveys. *J. Am. Med. Inf. Assoc.* 7.
- Schonlau, Matthias, Fricker, Ronald D., Elliott, Marc N., 2002. *Conducting Research Surveys via E-mail and the Web*. RAND Corporation, Santa Monica, CA. http://www.rand.org/pubs/monograph_reports/MR1480.html.
- Sinclair, M., O'Toole, J., Malawaraarachchi, M., Leder, K., 2012. Comparison of response rates and cost-effectiveness for a community-based survey: postal, internet and telephone modes with generic or personalised recruitment approaches. *BMC Med. Res. Methodol.* 12, 132–132.
- United States Environmental Protection Agency, 2017. *Air Quality Index Basics*. Available online at: <https://airnow.gov/index.cfm?action=aqibasics.aqi>. (Accessed 14 December 2017).
- United States Forest Service, 2005. *A Strategic Assessment of Forest Biomass and Fuel Reduction Treatments in Western States*. General Technical Report RMRS-GTR-149. Rocky Mountain Research Station, Fort Collins, CO, 17p.
- Weible, R., Wallace, J., 1998. Cyber research: the impact of the internet on data collection. *Market. Rev.* 10 (3).
- Whitehead, J., 2016. Plausible responsiveness to scope in contingent valuation. *Ecol. Econ.* 128, 17–22.
- Windle, J., Rolfe, J., 2011. Comparing responses from internet and paper-based collection methods in more complex stated preference environmental valuation surveys. *Econ. Anal. Pol.* 41 (1), 83–97.
- Yale Project on Climate Change Communication, 2014. *Yale Climate Opinion Maps*. <https://environment.yale.edu/poe/v2014/>.