Choosing between internet and mail survey modes for choice experiment surveys considering non-market goods

Søren Bøye Olsen
Institute of Food and Resource Economics, University of Copenhagen,
Rolighedsvej 25,
1958 Frederiksberg C
Denmark
E-mail: sobo@life.ku.dk , phone: +0045 35333643

Abstract:

In a choice experiment setup considering a non-market good, this paper adds to the literature on survey mode effects by providing evidence that internet surveys can be a viable alternative to traditional mail surveys when gathering feedback from a sample of respondents. The case study concerns preferences for protecting different types of landscape from road encroachment when building new motorways in Denmark. Two samples of respondents are surveyed – one by internet and one by mail. The performances of the two samples are compared over six different criteria; response rates, protest responses, demographics, preferences and WTP, estimation precision, and, finally, certainty in choice. Differences are observed for some of these criteria, implying that analysts should be aware that choosing internet over mail could be accompanied by a survey mode effect. However, the observed differences do not translate into significant differences in the unconditional WTP estimates. In most applied economic valuation studies of non-market goods, the main objective is in fact estimation of WTP. Hence, in the present case, the identified survey mode effects do not severely invalidate the applicability and continued use of the internet as a suitable means of collecting data for choice experiment economic valuation of non-market goods.

1. Introduction

Data collection in choice experiment surveys concerning economic valuation of non-marketed goods has often been carried out by mailing self-administered questionnaires to respondents randomly drawn from a survey population. The success of the mail survey approach is largely dependent on respondents filling in the questionnaire and returning it to the analyst. Thus, much research has focused on refining and improving the entire process of data collection via mailout self-administered questionnaires in order to achieve as high response rates as possible. Through years of research and intensive study, Dillman (1978) formulated *the total design method* and more recently *the tailored design method* (Dillman 2007) which is now widely accepted as a standard for conducting mail surveys. Other survey modes that have been used in stated preference (SP) studies over the years are various versions of telephone or in-person interviews.

With the emergence and surge of the internet up through the 1990's, a new and promising mode of data-collection appeared which has turned out to be one of the most profound developments in survey methodology (Dillman 2007). Internet surveys have gained widespread use particularly in market research, and according to Deutskens et al. (2006) internet surveys accounted for as much as 35% of the U.S. survey research market by the end of 2004. Within the marketing research literature much effort has gone into exploring and describing the potential advantages and disadvantages of using internet. However, virtually no surveys have addressed the desired *quality* of responses in terms of for instance the accuracy of answers (Deutskens et al. 2006). It seems reasonable to expect that most of the experience and findings from this branch of research can be more or less directly transferred to SP surveys.

In the context of economic valuation of non-market goods based on SP, data collection by means of internet surveys have gained widespread use in recent years. Berrens et al. (2004), Hudson et al. (2004), Tsuge and Washida (2003) and Ladenburg and Olsen (2008) are examples of internet-based

SP surveys, but none of these focus on comparability with other survey modes or the quality of the obtained data. In CVM studies concerning non-market goods Berrens et al. (2003) compare internet and telephone samples, while Marta-Pedroso et al. (2007) compare internet and in-person samples. To the best of the author's knowledge, Fleming and Bowden (2007) is the only paper published until now, which explicitly compares internet and mail samples in a non-market good case. However, their analysis is based on a revealed preference travel cost study.

Hence, there is an obvious gap in the literature with respect to comparisons of internet and mail sampling in the case of SP surveys considering non-market goods. Even though the existing literature can give valuable insights into many of the questions that arise when considering the choice of survey mode, some crucial questions have yet to be answered. Basically, the main question is: Can internet samples be used as an appropriate and reliable alternative to mail samples for administering SP surveys for non-market goods?

This paper provides a novel contribution to the literature by answering some of the emerging questions which need answering in order to assess the appropriateness and applicability of internet surveys as a means of conducting SP surveys in the context of non-market good valuation. Specifically, in a Choice Experiment survey setup this paper compares the performance of an internet and a mail sample[1] evaluated across six different criteria. The criteria are; 1) obtained response rates, 2) amount of protest zero bidders, 3) demographic distributions, 4) estimated preference parameters and unconditional WTP estimates, 5) estimation precision, and 6) certainty in choices.

The remainder of the paper is organized as follows. In the following section the existing literature on advantages and disadvantages of internet sampling is reviewed. Section 3 introduces the survey setup. In section 4 the results are presented along with a discussion and, finally, section 5 concludes the paper.

2. Advantages and disadvantages of internet sampling

One major advantage associated with internet sampling is reduced costs in relation to the eliminated expenses associated with printing and postage. Furthermore, the time and costs associated with the need to manually enter data in an electronic format is eliminated. Studies comparing the costs of mail surveys to internet based surveys generally find that the costs of conducting mail surveys are from 20 to 600 percent higher than the costs of internet based surveys (Cobanoglu et al. 2001; Forsman and Varedian 2002; Schleyer and Forrest 2000; Schmidt 1997; Weible and Wallace 1998). As the marginal costs of distributing questionnaires and receiving responses in an internet based survey can be exceptionally low, the saved costs can be spent on increasing the number of respondents markedly. In other words, at a given budget the internet survey offers much larger sample sizes than the conventional mail survey (Berrens et al. 2003).

Another advantage is the speed of response. Conventional mail surveys require a minimum of three to four days from sending the initial mail to receiving the filled-in questionnaire from the respondent. Internet surveys are faster to conduct. The minimum time required from e-mailing the invitation to participate in the survey to receiving a response from the respondent, is the time it takes for the respondent to read the e-mail, click the link to the web-based questionnaire, fill in the questionnaire and press the submit button. This can usually be done in as little as 20 to 25 minutes. Consequently, Cobanoglu et al. (2001) find an average response time of 4 days in an internet survey opposed to 16 days in a mail survey. Weible and Wallace (1998) find similar results.

Further advantages include improvements of the data quality in terms of less item non-responses, the ability to make questions conditional on previous answers in a (to the respondent) non-detectable

[1] This is not to say that other survey modes are not relevant, but these are the two modes of survey applied in the present empirical study. Furthermore, these two modes have been the most commonly applied in recent SP studies concerning non-market goods in Denmark. Hence, the choice to consider only these two methods even though two of the reviewers note that mail surveys are few and Computer Assisted Personal Interviews are more common in other countries. For a survey considering potential behavioral differences across CAPI and internet, see Bronner and Kuijlen (2007).

way, broader stimuli potential in terms of graphics and sound, as well as the avoidance of manual data entry mistakes (Ballantyne 2004; Muffo et al. 2003; Stanton 1998; Weible and Wallace 1998).

Increased response rates have been claimed to be another advantage of internet surveys over mail surveys. However, empirical studies in this respect show somewhat ambiguous results. Yun and Trumbo (2000), Cobanoglu et al. (2001), Wygant and Lindorf (1999), Ballantyne (2004) and McCabe et al. (2002) all find markedly higher response rates in internet surveys compared to mail surveys. But Forsman and Varedian (2002) and Weible and Wallace (1998) find similar response rates in the two modes, whereas Dommeyer et al. (2004), McMahon et al.(2003), Leece et al. (2004) and Kaplowitz et al. (2004) actually report lower response rates in internet surveys. As Leece et al. (2004, p.1) conclude: *"Researchers should not assume that the widespread availability and potential ease of internet-based surveys will translate into higher response rates"*. In a review of 65 studies Schonlau et al. (2002) find support for this statement and further suggest that the context of the individual study affects the response rate.

2.1. Sample- and self-selection problems in internet sampling?

The most severe disadvantage of using internet surveys is associated with the sampling procedure. As access to a computer and the internet is not typically available to every single individual in a population, problems concerning sample coverage and sample representativeness may be expected. In particular this holds true for surveys covering the population in general, whereas it poses a much lesser problem if the target population is a sub- population with very high internet usage percentages, e.g. university students or government agency staff members (Forsman and Varedian 2002; Kaplowitz et al. 2004).

Through the last decade online computing and internet access has become widely available in Denmark. In 2006 nearly 90% of the Danish population had access to the internet either at home or at work, and in total 73% of the Danish population reported using the internet at least once a week (Ottens 2006). On the face of it, the very high level of internet penetration in Denmark, being amongst the top three in the EU, holds promises for the suitability of internet surveys. However, there is still 14% of the population who have never used the internet. This constitutes a coverage problem as this group of people will be excluded from entering an internet survey of the general public. This calls for caution as this group of people consists mainly of people above 60 years of age, people who are unemployed, and people with relatively short educations.

Another drawback of using the internet is the uncertainty of whether or not it might introduce additional selection bias. Sample selection bias caused by missing observations is a well-known problem in traditional mail surveys (Heckman 1979; McFadden et al. 2005). However, the process of self-selection into internet surveys might differ from that of self-selection into mail surveys. If this is the case it might ultimately affect the elicited preferences in the sample. Whether this would be good or bad is difficult to say *a priori*. Preferably, this question should be addressed on a case by case basis. If internet sampling introduces *additional* self-selection processes on top of those already present in mail sampling, we would expect an overall increase in the bias. However, if internet sampling merely introduces *different* self-selection processes, it is less obvious whether this would actually increase, decrease or even at all affect the overall selection bias. In the current study, an increase in selection bias is expected due to the use of an internet panel which is subject to additional self-selection processes. Section 2.2 discusses this in more detail.

Even if there is no difference in sample representativeness, and self-selection mechanisms do not differ between the two modes of sampling, elicited preferences might still differ markedly. Possible explanations for such a pure survey mode effect could be; 1) differences in familiarity and technical issues might lead to more or less respondent "mistakes" (Knapp and Kirk 2003), 2) information assimilation and processing might differ across modes (Dillman and Christian 2005), and 3) perceived differences in anonymity might affect behavior (Knapp and Kirk 2003). Though not directly addressing SP surveys, Dillman (2007) finds that different survey modes generally lead to different answers to the same question, though the difference between internet and mail surveys is far less than for instance the difference between internet and telephone surveys.

3. Survey approach

In order to answer the questions above, this study compares two data sets obtained from a mail survey and an internet survey, respectively. Apart from minor differences due to technical limitations[2], identical questionnaires were used in the two surveys. The main working hypothesis for this survey is that preferences elicited in the two samples are identical. This entails the underlying assumption that no differences in elicited preferences and no differences in demographic background variables will materialize as a result of using the internet instead of paper-and-pencil questionnaires for data collection, i.e. there will be no survey mode effects. If this hypothesis cannot be rejected, the questions and concerns raised in the above section should not prevent us from increasing our use of internet sampling in future SP surveys, at least when the choice is between internet and mail.

The data sets used for analysing and testing this hypothesis originate from a choice experiment survey that examines the recreational merits associated with reducing the impact of new motorways on different types of landscape. The hypothetical scenario was based on the assumption that 100 kilometres of new motorways are to be built in Denmark during the next ten years. The scenario described how the exact location of these stretches of motorway through the countryside could be decided upon with more or less consideration for potential encroachment of different landscapes.

Three different landscape types of interest were identified and chosen as main attributes in the choice experiment design. The three attributes were 'forest', 'wetland', and 'heath/pastoral area'. The base case levels of the attributes were assigned on the basis of the area distribution of these landscape types in Denmark in general. Thus, the status quo alternative in the choice sets entailed 10 kilometres of new motorway located through forest areas, 5 kilometres through wetland areas, and 5 kilometres through heath/pastoral areas[3]. For each of the three main attributes, three levels of protection of the specific landscape type were used. To enable estimation of WTP, a price attribute with six levels was defined as an extra annual income tax on the household. The attributes and their assigned levels are summed up in table 1.

>>>Insert table 1 about here<<<

3.1. Design

A linear D-optimal fractional factorial design consisting of 18 choice sets (36 different alternatives) was identified using SAS software (Zwerina et al. 1996, Kuhfeld 2004). Each choice set consisted of three alternatives: the zero-priced status quo alternative and two improvement alternatives with an associated price. Respondents were randomly divided into a three-block structure, evaluating six choice sets each. Focus group interviews indicated that this was a suitable number of choice sets per respondent.

3.2. Sampling procedure

Respondents in both survey modes were sampled from the two municipalities Silkeborg and Gjern located in central Jutland. This survey area was chosen to define the target population, as a new motorway through this area is currently being planned[4].

---

[2] The technical limitations refer to the transferring of the printed version of the questionnaire to the screen, or rather making it fit to a standard screen size. The aim was to avoid respondents having to scroll down or sideways, as the internet survey agency recommended to avoid this.

[3] As the total stretch of motorway was fixed at 100 kilometres, a fourth supplementary attribute, 'arable land', was introduced to account for the location of the remaining 80 kilometres. Due to perfect correlation, it was not included in the experimental choice set design, and it is not included in the parametric modelling of preferences.

[4] As one reviewer notes, this target population may be more sensitive to the issue than the general public. Particularly, as the concern regarding motorway encroachment may be more widespread than it would be in other areas, self-selection might be less than it would be elsewhere. If so, it could influence inference. Unfortunately, it was not possible to assess whether this was the case in the present study. This stresses the need for further research in this area.

In June 2005, the mail survey was sent to 300 respondents aged 18 to 70, randomly selected from the 63,641 citizens registered in the Danish Civil Registration System for this area. Besides the questionnaire, a postage paid return envelope and an introductory letter were enclosed in the envelope. Approximately one week after dispatch, a short reminder notice was sent to those who had not yet responded. After one more week, a final reminder notice was sent to non-responders, this time with the questionnaire enclosed once again. Two weeks later, at the end of June, the collection of questionnaires ended.

During the same period the internet survey was carried out by the survey company ACNielsen AIM A/S. 253 respondents living in the survey area were sampled from ACNielsen's pre-recruited internet panel, called Panel.online. Internet surveys come in many varieties; see e.g. Couper (2000). Using a pre-recruited panel of internet users is quite common, and in the present case it was the most reasonable way of obtaining e-mail addresses for the population in the area. Respondents were sent an e-mail containing both information about the survey and a personalised link to the website where the web-based version of the questionnaire was located. Respondents not answering within a week were sent an e-mail reminder with the link to the website. Respondents still not answering after another week were sent a final e-mail reminder, and after yet another week the website was shut down.

It is obvious that the two different sampling procedures might have consequences. In the mail survey, respondents are randomly sampled directly from the survey population. Because people in the initial sample can choose whether or not they want to participate in the survey, self-selection could potentially bias results. Self-selection occurs when some specific group of people in the sample for some reason chooses not to respond. The decision to participate or not may be correlated with people's preferences for protection of different types of landscape thus making the respondents a non-random and non-representative sample with regard to preferences in the population, ultimately biasing results from the survey (Heckman 1979). For example, people who feel very strongly for protecting the landscape may be more willing to spend time answering the survey than those who do not.

The self-selection process in the internet survey is potentially quite different from that in the mail survey. At the time of the survey Panel.online had approximately 17,000 participants who had all been recruited in connection with ACNielsen's regular telephone or personal omnibus survey interviews[5]. Thus, participants in Panel.online go through additional major self-selection processes. One relates to the omnibus surveys which are (usually) based on a random sample of the entire population. The self-selection process related to the decision of participating in the omnibus survey or not, is similar to the one described for the mail sample above, but for two main reasons they are not identical. Firstly, the object of interest in the omnibus survey is most likely not the same as in the present survey. However, this problem might not be severe, as the omnibus surveys generally cover a wide variety of topics and, hence, a wide variety of people choose to participate. Secondly, the omnibus survey interviews are done by telephone or in-person. Compared to self-administered paper-and-pencil questionnaires this will typically lead to somewhat different self-selection processes (Dillman 2007).

Another additional self-selection process that the internet panellists go through is related to the decision of participating in Panel.online or not. It seems quite likely, that those accepting to join Panel.online are not a random sample of those who have initially chosen to participate in the omnibus survey. Furthermore, Panel.online participants receiving the internet version of the questionnaire are subject to a third self-selection process in terms of the decision to participate in this specific survey or not. Even though this final self-selection corresponds to that of the mail sample, it still potentially differs as it is conditional on the first two self-selection processes. Hence, it seems reasonable to suspect that the keyboard-and-screen option will be more biased by self-selection processes than the paper-and-pencil option[6].

[5] Having finished the omnibus survey interview, respondents are asked if they would like to join Panel.online where they would occasionally receive e-mails linking to questionnaires. Answering a questionnaire enters the respondent in a draw for various prizes.

[6] As the considerations concerning additional self-selection bias in the internet sample was known ex ante, propensity weighted sampling based on quotas regarding gender and age was applied in order to account for

3.3. Econometric specification

The model applied in the parametric analysis of responses is a mixed logit model which can be derived in a number of different ways (see Train 2003; Hensher and Greene 2003). In the present case, a model formulation which incorporates random parameters as well as an error component was found suitable. This model specification avoids major limitations of the multinomial logit model. Importantly, it allows explicitly for unobserved taste heterogeneity, i.e. random taste variations across respondents, and it is not restricted by the Independence of Irrelevant Alternatives (IIA) property (Hensher and Greene 2003; Revelt and Train 1998; Train 2003). Furthermore, it is a computationally practical and flexible model which can approximate any random utility model (McFadden and Train 2000).

Following Scarpa et al. (2005) an Alternative Specific Constant (ASC) is specified for the status quo alternative in order to capture the systematic component of a potential status quo effect. Furthermore, an error component additional to the usual Gumbel-distributed error term is incorporated in the model to capture any remaining status quo effects in the stochastic part of utility. The error component which is implemented as an individual-specific zero-mean normally distributed random parameter is assigned exclusively to the two non-status quo alternatives. By specifying a common error component across these two alternatives, correlation patterns in utility over these alternatives is induced. Thus, it captures any additional variance associated with the cognitive effort of evaluating experimentally designed hypothetical alternatives (Hensher and Greene 2007; Scarpa et al. 2007; Scarpa et al. 2008). This results in the following general utility structure:

$$U_{ntj} = \begin{cases} V(x_{ntj}, \widetilde{\beta}_n, \beta, \mu_n) + \varepsilon_{ntj}, & j = 1,2; \\ V(ASC, x_{ntj}, \widetilde{\beta}_n, \beta) + \varepsilon_{ntj}, & j = SQ \end{cases} \quad (1)$$

where the indirect utility, V, is a function of the vector of explanatory variables, x_{njt}, as well as the vectors of individual-specific random parameters, $\widetilde{\beta}_n$, and fixed parameters, β. For the two experimentally designed policy alternatives, the common individual-specific error component μ_n enters the indirect utility function, while it is replaced by the ASC for the status quo alternative. The unobserved error term ε_{ntj} is assumed Gumbel-distributed. The individuals are denoted by n, while j is the alternative and t is the choice set. The $\widetilde{\beta}_n$ varies over individuals in the population with density $f(\beta|\theta)$, where matrix θ is a vector of the true parameters of the taste variation, e.g. representing the mean and standard deviation of the β's in the population. Assumptions concerning the distribution of each of the random parameters, i.e. the density function $f(\beta|\theta)$ are necessary. The true distribution is unknown, so in principle any distribution could be applied (Carlsson et al. 2003; Hensher and Greene 2003). The normal is the most easily applied distribution (Train and Sonnier 2005).

In the present paper, parameters associated with the three landscape type attributes are specified as normally distributed random parameters to allow for both negative and positive preferences for the different types of landscape. Focus group interviews and a pilot test indicated that this could be expected. The multivariate normal distribution of individual tastes across individuals can be written as $\widetilde{\beta}_n \sim N(\beta, \Omega)$. The variance-covariance matrix Ω is specified so as to allow for correlation across random parameters, i.e. diagonal as well as off-diagonal values are estimated in the model (Train and Weeks 2005; Scarpa et al. 2008).

The price parameter is treated as a fixed rather than a random parameter, even though it implies fixed marginal utility of money. This approach is chosen for two reasons. First, it results in a

---

demographic biases in the internet panel sampling frame. The propensity weighting procedure was only conducted for respondents in the internet sample.

behaviourally plausible negative sign for all respondents. Second, and more importantly, it avoids a number of severe problems associated with specifying a random price parameter (see Meijer and Rouwendal 2006; Hensher et al. 2005; Hess et al. 2005; Train and Sonnier 2005; Campbell et al. 2006; Hensher and Greene 2003; Rigby and Burton 2006; Train 2003; Train and Weeks 2005, Train 2001). Hence, to avoid these problems, the commonly applied approach of modelling the price parameter as a fixed parameter is used in this paper[7]. Consequently, mean WTP can be calculated simply by dividing the mean estimate of the random parameter of interest with the estimate of the fixed price parameter (Louviere et al. 2000). The distribution of WTP follows the distribution of the random parameter. Hence, the standard deviation of the now normally distributed WTP can be calculated by dividing the estimated standard deviation for the random parameter by the estimated price parameter (Revelt and Train 1998). However, this approach might be too simple as it implies taking the point estimates as given and ignoring the sampling variance in these (Hensher and Greene 2003; Hensher et al. 2005). To take sampling variance into account and use all the information on the distribution of the random parameters instead of just the mean and the standard deviation, the Krinsky-Robb parametric bootstrap method is utilized to generate full out-of-sample unconditional WTP distributions for each landscape type (Krinsky and Robb 1986, 1990). From these distributions, confidence intervals are derived by identifying the relevant percentiles (Hole 2007). This approach is similar to that recommended by Hensher et al. (2005).

4. Results and discussion

In the following, the internet and mail samples are compared according to the six criteria listed in the introduction.

4.1. Response rates

Obtained rates of response are reported in table 2. With a total of 161 responses, the internet survey yielded a response rate of 63.6 percent, whereas the mail survey obtained 181 responses. Due to an initially larger sample, this resulted in a slightly lower response rate at 60.3 percent. Nevertheless, both modes of survey produce an initial response rate above 60% which is generally considered good for analysis and reporting (Babbie 1990).The similarity of response rates would suggest that the choice of survey mode does not affect the initial rate of response much. Recalling that previous studies have shown ambiguous results with regard to response rates, it is not surprising that the initial response rates differ slightly between the two modes of survey in the present study. Looking at the effective samples, the response rates differ somewhat more than in the initial sample, still with the internet obtaining the highest response rate. This is in support of the group of studies that find internet sampling to lead to increased rates of response (Schonlau et al. 2002).

>>>Insert table 2 about here<<<

4.2. Protest zero responses

In line with Morrison et al. (2000), protest zero bidders and irrational bidders are removed from the initial samples to yield the effective samples used in the parametric modelling of preferences in section 3.4. Protest bidders are respondents who, in their answers to the debriefing questions, reveal that their evaluations of the choice sets do not reflect their true preferences for the environmental good

[7] In a previous version of this paper (Olsen 2007), two additional models applying a random coefficient for the price variable were tested. Allowing the price parameter to be random (normally and lognormally distributed) showed some structural advantages leading to statistically significant model improvements compared to using a fixed price parameter, However, severe problems emerged in establishing reliable WTP distributions. Even though these models may be more appealing from a theoretical point of view, they are consequently less relevant for policy advice as well as for comparing WTP estimates across samples. Interestingly, the overall conclusions concerning potential survey mode effects in the previous version of the paper were the same across the different model specifications tested. This suggests that the conclusions in the present paper do not rest on the choice of using a fixed price parameter in the model.

in question (e.g. Mitchell and Carson (1989) or Jorgensen et al. (1999)). According to Lindsey (1994), protest responses may occur in two forms. Respondents may state a zero WTP even though their true WTP is actually higher, or they may deliberately state a WTP which is higher than their true WTP. The three main reasons for stating a protest response are 1) failing to understand the valuation question and answering anyway, 2) acting strategically, or 3) protesting against some specific component of the study (Boyle 2003). However there exists no well established or generally agreed upon protocol for identifying protest respondents, so this is typically done on an ad-hoc basis (Jorgensen et al. 1999; Boyle and Bergstrom 1999). Several authors note that this is unfortunate, as using different approaches to identifying protest answers might lead to significantly different WTP estimates (Boyle and Bergstrom 1999; Jorgensen and Syme 2000; Meyerhoff and Liebe 2006))

In the present study, protesters were identified as respondents who chose the zero-priced opt-out alternative in all six choice sets and reasoned this with one of the following statements; "I pay enough taxes as it is", "The questions were too difficult", "No more motorways should be built in Denmark", and "I don't know why I chose as I did". This is the typical way of identifying protest zeros (Meyerhoff and Liebe 2006). Unfortunately it was not possible in the present study to identify the other type of protest bids in terms of bids overstating the true WTP, as there were no debriefing questions addressing this issue. Irrational bidders were identified as those who failed to pass a built-in test for transitivity in terms of a choice set containing a perfectly dominating alternative.

As table 2 reports, the number of protest zero bidders differ markedly in the two samples. Considering that Marta-Pedroso et al. (2007) found no survey mode effects on protest responses in a study comparing an internet survey to an in-person survey, it is quite interesting that the mail survey entails more than double the amount of protest bidders than the internet survey. With a χ^2-value of 5.68 and a p-value of 0.017, a χ^2-test comparing the number of protest bidders and irrational respondents in the two samples confirm that the difference is significant. Taken at face value, this finding would imply that internet surveys do in fact have an advantage over mail surveys in terms of receiving more valid replies, which translates into higher effective response rates, other things being equal. On the other hand, this might simply reflect a larger degree of self-selection in the internet sample which could potentially bias results, as described in section 2.2. If this is the case, it might be argued that the additional self-selection processes in the internet sample result in respondents being more willing to accept the hypothetical scenario and play along in the valuation exercise. There is no *a priori* expectation as to whether this will affect stated preferences or not.

4.3. Socio-demographic characteristics of the samples

Table 3 displays the distributions of the samples with regard to certain socio-demographic variables. Comparing the distributions in the two samples reveals the following.

Gender and age

With regard to gender, there is no significant difference between the two samples at any of the three different sample levels. Furthermore, the samples' gender distributions do not differ significantly from that of the survey population in the study area in general.

When looking at the age distribution in the initial samples (columns 2-4), the internet and mail samples differ significantly not only from each other but also from the survey population. The internet sample exhibits an overrepresentation of people in the middle age classes from 35 to 54 years of age. This is counterbalanced by an underrepresentation of respondents in the two oldest age classes, whereas the youngest age classes are fairly well represented. Comparing to the expected distribution, the initial mail sample, on the other hand, exhibits an overrepresentation of people in the age classes from 45 years and up. Comparing the age distribution in the initial mail sample to the internet sample, it is likely that the highly significant difference between the two is caused by the additional self-selection processes in the internet sample as well as the rather ineffective propensity weighting procedure.

Interestingly, when looking at the respondent samples (columns 5-7), the difference between the age distributions in the two samples is slightly evened out. Exclusion of protesters further reduces the

difference to an extent where the difference is no longer significant in the effective samples (columns 8-10). However, comparing the effective samples to the survey population, the internet sample is still significantly skewed, whereas for the mail sample the difference is no longer significant.

>>>Insert table 3 about here<<<

Household gross income and education

At all three sample levels, both the internet and the mail sample exhibit highly significant deviations from the distribution of household income in the survey population. However, the distribution is similar in the two samples for the respondent samples as well as the effective samples, so the skewness is similar in the two modes of sampling. Thus, both samples exhibit a much larger share of high income households than is the case in the survey population. This coincides with a strong overrepresentation of university educated in the two samples compared to the survey population. This is not surprising as income is positively correlated with the level of education. The share of university educated is however significantly larger in the internet sample than in the mail sample for both respondent and effective samples. Unfortunately, information about income and education was not available for the initial mail sample. Nevertheless, comparing the initial internet sample to the survey population further confirms the strong overrepresentation of highly educated people with high income. These differences could be interpreted as further evidence of the different self-selection processes in the two samples.

4.4. Parametric analysis of stated preferences and estimation of WTP

Using Biogeme version 1.7 (Bierlaire 2003) the mixed logit models are estimated by use of simulated maximum likelihood estimation. All results are obtained using Halton draws with 300 replications as this was found to be sufficient for results to stabilize. In the dataset, all of the attributes are linearized such that the parameter estimates represent the shift in utility associated with a one-unit increase. Results are summarized in table 4.

>>>Insert table 4 about here<<<

Across the two samples, all but the mean estimates for the heath attribute are significant on a 95% significance level. Of the three landscape type attributes, wetland is clearly valued highest, followed by forest, and heath is the least valued landscape, not being significantly different from zero. This internal ranking is similar in both samples. The landscape attribute estimates are all of a negative sign. This indicates that on average respondents experience a diminishing utility when one additional kilometre of motorway is placed through the specific types of landscape[8]. Likewise, the price parameters are of the expected negative sign. The estimated standard deviations for the landscape types reveal significant heterogeneity in respondents' preferences for the three landscape type attributes. However, there are differences between the two samples. While there is significant taste variation in the internet sample with regard to wetland and heath, this is not the case in the mail sample. In other words, preferences elicited in the mail sample are generally more homogeneous than those elicited in the internet sample. Disturbingly, this suggests that the choice of survey mode might indeed affect the preferences elicited – a result which might be ascribed to the different self-selection processes leading to behavioural differences in the two samples.

Another difference is evident when focus is turned to the correlation in utility between the two experimentally designed policy alternatives. The *a priori* expectation was that even though the ASC captures some of the status quo effects, there would be significantly different covariance structures across the utilities of the experimentally designed alternatives and those of the status quo alternative

[8] As one reviewer notes, the applied coding of variables implies that the parameter estimates reflect the (dis)utility of increases in the number of kilometers of motorway going through the landscape types, relative to going through agricultural land. Hence, the negative estimates.

(Scarpa et al. 2005; Scarpa et al. 2008). Even though a systematic component of the status quo effect is present in both samples in terms of significant ASCs, the stochastic component, μ_{12}, is only significant in the mail sample. This suggests that the internet sample is less susceptible to status quo effects.

In both samples, there are some significant correlations between random parameters. Heath is positively correlated with forest, while heath and wetland are uncorrelated. With regard to the correlation between wetland and forest, this is positive and significant at the 90% level in the internet sample, while it is insignificant in the mail sample. This again suggests that there are structural differences in the preferences elicited in the two modes of survey.

To formally test the hypothesis of identical preferences in the two samples, a likelihood ratio test for nested models is conducted. To account for potential differences in scale variance (Swait and Louviere 1993), the ratio λ between scale parameters in the two samples is estimated and presented in the fourth column in table 4. The estimated scale ratio of 1.33 is significantly different from 1 evaluated at a 90% significance level. This indicates that scale does indeed differ across the two samples, and with a LR-test value of 5.79 obtained for the pooled and scaled sample, the hypothesis of identical preferences across the two samples cannot be rejected. Thus, even though some differences are suggested when comparing individual parameter signs and confidence levels, the LR test indicates that these differences do not cause significant differences at the overall model level.

Due to the differing scale parameters in the two datasets, parameter estimates in table 4 are not directly comparable between the two samples (Louviere et al. 2000). This is overcome by instead turning to the unconditional WTP estimates reported in table 5. Direct comparison of WTP estimates across the two samples is possible as the scale parameter cancels out when calculating WTP (Train 2003).

>>>Insert table 5 about here<<<

Table 5 reveals an internal preference ordering consistent with the findings in table 4. In both samples, the WTP for protecting wetlands is higher than for protecting forests, and the WTP for protecting heaths is lower than for forests and wetlands. Looking at the 95% confidence intervals, it is evident that the WTP estimates are quite similar across the two samples. This is confirmed by the more formal p-values. These are obtained from a t-test testing the null hypothesis of equality in means (H_0: $WTP_{internet}=WTP_{mail}$), as well as from the Complete Combinatorial approach suggested by Poe et al. (1994; 2005). This approach tests the null hypothesis of the difference between the two being equal to zero (H_0: $WTP_{internet} - WTP_{mail} = 0$). Both tests fail to reject their null hypothesis for all three landscape type attributes. This suggests that the different modes of survey do not cause any significant differences in WTP.

4.5. Estimation precision

According to Louviere et al. (2003), the scale parameter is inversely proportional to the standard deviation of the unobserved effects, i.e. the error term in the utility function. As such, the scale parameter can be interpreted as an expression of estimation precision; a relatively high scale parameter implies a relatively low variance of the error term, and, consequently, a relatively higher degree of precision in the estimates (Adamowicz et al. 1998). Hence, the estimated scale ratio, λ, presented in table 4, expresses the variance of the unobserved factors in the mail sample relative to that in the internet sample. As already noted above, the estimated scale ratio at a value of 1.33 is significantly different from 1. Recalling that this estimate is the scale parameter of the mail sample relative to that of the internet sample which is normalized to 1, it implies that the variance[9] of unobserved effects is 43% lower in the mail sample than in the internet sample (Train 2003). In other words, the mail sample induces a markedly higher degree of estimation precision and reliability in choices than the internet sample. This result corresponds to the difference observed in overall model fit for the two

[9] Due to the normalization, the relative error term variance is calculated as $\sigma^2 = 1 / \lambda^2$.

samples in table 4. Even though the pseudo-R^2 values reveal that both models generally fit the data very well, the mail sample obtains a markedly higher fit than the internet sample.

4.6. Certainty in choice

To assess the degree of preference uncertainty, respondents where asked to state their level of certainty in choice after each completed choice set. Replies were given on a five-point Likert scale ranging from "very uncertain" (value 1) to "very certain" (value 5). The averages of the self-reported certainty levels for each choice set are illustrated in figure 1.

>>>Insert figure 1 about here<<<

The figure reveals a tendency of internet-respondents generally reporting a higher level of certainty in choice than the mail-respondents. This tendency becomes slightly more pronounced throughout the sequence of the six choice sets. Confirming this tendency, χ^2-tests comparing the distributions of replies across the two samples rejected the null of identical distributions of certainty in choice for all six choice sets. This result is somewhat surprising considering the above results concerning the lower degree of estimation precision in the internet sample. Lundhede et al. (2009) find evidence that as the certainty in choice increases, scale parameters increase, i.e. error variance decreases, and, hence, precision increases. A straight-forward interpretation of the counter results found in the present paper could be that internet panellists are relatively more experienced in terms of answering questionnaires. As a consequence, they feel more confident in their choices even though their stated preferences actually indicate the opposite. This suggests that the internet respondents tend not to consider their preferences for the non-market good to the same degree as the mail respondents before answering the preference eliciting choice sets. Thus, for this type of research, using an internet panel instead of mailing respondents could imply running the risk of obtaining somewhat more "quick-and-dirty" responses.

5. Conclusion

By now, the internet has gained widespread use as a tool for data collection in applications of non-market good stated preference surveys. However, the applicability of this rather new mode of survey is not yet well established. This paper adds to the existing literature on survey mode effects by comparing the performance of an internet sample to a mail sample in a choice experiment survey considering the preferred location of new motorways through different types of landscape in Denmark. Over a range of different criteria, the results suggests that a non-market good choice experiment survey conducted using a pre-recruited internet panel can indeed be a viable alternative to the traditional mail surveys. Even though there are some noteworthy differences across the two modes of survey, they are considered to be of minor importance when the goal of the survey is estimation of WTP.

The internet sample yields higher effective response rates than the mail sample, one of the reasons being that the internet sample contains significantly fewer protest bidders. This is likely to be caused by additional self-selection processes in the internet sample, but more research into this area is clearly needed. With regard to the ability to represent the demographics of the intended survey population, neither survey mode performs satisfactorily, though the internet sample is slightly more off target. This could be interpreted as a further indication of the different self-selection processes in the two samples. The parametric modeling of preferences suggests that there are minor structural differences in preferences across the two samples, but a Likelihood Ratio test cannot reject that, on the overall, the preferences elicited by the two modes of sampling are identical when differences in scale are accounted for. A comparison of unconditional Willingness-To-Pay estimates supports this conclusion as no significant differences are found here either. Interestingly, a significant difference in scale parameters is identified, implying that the mail sample obtains a higher degree of precision in the estimation than the internet sample. Considering the positive correlation between certainty in choice

and estimation precision established in Lundhede et al. (2009), this result is at odds with the fact that, when asked directly, internet respondents state a generally higher level of certainty in their choices.

Considering the advantages as well as the continuing increase in internet access in the general population, internet sampling appears to be a valid replacement of the traditional mail sampling approach in SP surveys considering valuation of non-market goods. However, as this paper shows, analysts conducting such surveys should be aware, that a survey mode effect could very well be present. Even though choosing internet over mail is subject to a survey mode effect in terms of a larger degree of self-selection bias, minor changes in stated preferences, and a decrease in estimation precision, this does not translate into marked differences in the WTP estimates. Thus, considering the fact that the main objective in most applied economic valuation studies of non-market goods is in fact estimation of WTP, these minor drawbacks do not severely invalidate the applicability and continued use of the internet as a suitable means of collecting data for choice experiment economic valuation of non-market goods. Of course, unequivocal conclusions should not be drawn from one single study, and it might be that the identified survey mode effects could actually affect WTP estimates in other cases. Consequently, more research on the topic is needed before we can fully endorse internet surveys as a way forward for non-market valuation.

6. Acknowledgements

The author gratefully acknowledges the helpful and valuable comments given by Morten Mørkbak, Mogens Fosgerau, Danny Campbell, Eirik Schrøder Amundsen, Bo Jellesmark Thorsen, Jørgen Dejgård Jensen and Jacob Ladenburg. Financial support from the Environmental Assessment Institute, Copenhagen, is appreciated. As always, any errors remaining are the sole responsibility of the author.

7. References

Adamowicz W L, Louviere J, Swait J (1998) Introduction to attribute-based stated choice methods. National Oceanic and Atmospheric Administration, Washington, USA.

Babbie E (1990) Survey Research Methods. Wadsworth Publishing Company, Belmont, California

Ballantyne C (2004) Online or on paper: An examination of the differences in response and respondents to a survey administered in two modes. Paper presented at the Australasian Evaluation Society 2004 International Conference, Adelaide, South Australia, 13-15 Oct 2004

Berrens R P, Bohara A K, Jenkins-Smith H C, Silva C L, Weimer L (2003) The Advent of Internet Surveys for Political Research: A Comparison of Telephone and Internet Samples. Polit Anal 11(1): 1-22

Berrens R P, Bohara A K, Jenkins-Smith H C, Silva C L, Weimer D L (2004) Information and effort in contingent valuation surveys: application to global climate change using national internet samples. J Environ Econ Manage 47(2): 331-363

Bierlaire M (2003) BIOGEME: A free package for the estimation of discrete choice models. In: Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland, 19-21 Mar 2003

Boyle K J (2003) Contingent valuation in practice. In: Champ P A, Boyle K J, Brown T C (eds.) A Primer on Nonmarket Valuation. Kluwer, Dordrecht, Netherlands, pp 111 –169

Boyle K J, Bergstrom J C (1999) Doubt, doubt, and doubters: the genesis of a new research agenda? In: Bateman I J, Willis K G (eds.) Valuing Environmental Preferences. Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries. Oxford University Press, Oxford, pp 183– 206

Bronner F, Kuijlen T (2007) The live or digital interviewer, Int J Market Res 49(2): 167-190

Campbell D, Hutchinson W G, Scarpa R (2006) Using mixed logit models to derive individual-specific WTP estimates for landscape improvements under agri-environmental schemes: evidence from the Rural Environment Protection Scheme in Ireland. In: Koundouri P and Birol E (eds) Choice Experiments Informing European Environmental Policy. Edward Elgar Publishing (Forthcoming)

Carlsson F, Frykblom P, Liljenstolpe C (2003) Valuing wetland attributes: an application of choice experiments Ecol Econ 47(1):95-103
Cobanoglu C, Warde B, Moreo P J (2001) A comparison of mail, fax and web-based survey methods. Int J Market Res 43(4):441-452
Couper M P (2000) Web Surveys: A Review of Issues and Approaches. Publ Opin Q 64: 464-494
Deutskens E, de Jong A, de Ruyter K, Wetzels M (2006) Comparing the generalizability of online and mail surveys in cross-national service quality research. Market Lett 17(2):119-136
Dillman D A (1978) Mail and Telephone Surveys. The Total Design Method. John Wiley & Sons, New York
Dillman D A (2007) Mail and Internet Surveys: The Tailored Design Method - 2007 Update with New Internet, Visual, and Mixed-Mode Guide. John Wiley & Sons, New York
Dillman D A, Christian L M (2005) Survey Mode as a Source of Instability in Responses across Surveys. Field Meth 17(1):30-52
Dommeyer C J, Baum P, Hanna R W, Chapman K S (2004) Gathering faculty evaluations by in-class and online surveys: Their effects on response rates and evaluations. Assess Eval High Edu 29(5):611-623
Fleming C M, Bowden M (2009) Web-based surveys as an alternative to traditional mail methods. J Environ Manag 90(1):284-292
Forsman G, Varedian M (2002) Mail and Web Surveys: A Cost and Response Rate Comparison in a Study of Students Housing Conditions. Paper presented at The International Conference on Improving Surveys, 25 - 28 Aug 2002, Copenhagen, Denmark
Greene W H, Hensher D A (2007) Heteroscedastic control for random coefficients and error components in mixed logit. Transport Res E Logist Transport Rev 43(5):610-623
Heckman J J (1979) Sample Selection Bias as a Specification Error. Econometrica 47(1):153-161
Hensher D A, Greene W H (2003) The mixed logit model: The state of practice. Transportation 30(2):133-176
Hensher D A, Rose J M, Greene W H (2005) Applied Choice Analysis: A Primer. Cambridge University Press, Cambridge, UK
Hess S, Bierlaire M, Polak J (2005) Estimation of value of travel-time savings using Mixed Logit models. Transport Res Pol Pract 39(3):221-236
Hole A R (2007): A comparison of approaches to estimating confidence intervals for willingness to pay measures. Health Econ 16(8):827-840
Hudson D, Seah L-H, Hite D, Haab T (2004) Telephone presurveys, self-selection, and non-response bias to mail and Internet surveys in economic research Appl Econ Lett 11(4): 237-240
Jorgensen B S, Syme G J (2000) Protest responses and willingness to pay: attitude toward paying for stormwater pollution abatement. Ecol Econ 33(2): 251-265
Jorgensen B S, Syme G J, Bishop B J, Nancarrow B E (1999) Protest Responses in Contingent Valuation. Environ Resource Econ 14(1): 131-150
Kaplowitz M D, Hadlock T D, Levine R (2004) A Comparison of Web and Mail Survey Response Rates. Publ Opin Q 68(1):94-101
Knapp H, Kirk S A (2003) Using pencil and paper, Internet and touch-tone phones for self-administered surveys: does methodology matter? Comput Hum Behav 19(1):117-134
Krinsky I, Robb A L (1986) On approximating the statistical properties of elasticities. Rev Econ Stat 68(4):715–719
Krinsky I, Robb A L (1990) On approximating the statistical properties of elasticities: a correction. Rev Econ Stat 72(1):189–190
Kuhfeld W (2005) Marketing Research Methods in SAS. Experimental Design, Choice, Conjoint and Graphical Techniques. SAS Institute Inc. http://support.sas.com/techsup/technote/ts722.pdf. Cited 12 November 2008.
Ladenburg J, Olsen S B (2008) Gender Specific Starting Point Bias in Choice Experiments: Evidence from an Empirical Study. J Environ Econ Manage 56(3):275-285

Leece P, Bhandari M, Sprague S, Swiontkowski M F, Schemitsch E H, Tornetta P, Devereaux P J, Guyatt G H (2004) Internet Versus Mailed Questionnaires: A Controlled Comparison. J Med Internet Res 6(4):e39
Lindsey G (1994) Market Models, Protest Bids, and Outliers in Contingent Valuation. J Water Resour Plann Manag 120(1):121-129
Louviere J, Hensher D A, Swait J (2000) Stated Choice Methods. Analysis and Applications. Cambridge University Press, Cambridge, UK
Lundhede T H, Olsen S B, Jacobsen J B, Thorsen B J (2009) Tough and easy choices: Testing the influence of utility difference on stated certainty-in-choice in choice experiments. Paper presented at the International Choice Modelling Conference 2009, 30 March – 1 April 2009, Harrogate, United Kingdom
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
McCabe S E, Boyd C J, Couper M P, Crawford S, Darcy H (2002) Mode Effects for Collecting Alcohol and Other Drug Use Data: Web and U.S. mail. J Stud Alcohol 63(6):755-761
McFadden D, Train K (2000) Mixed MNL models for discrete response. J Appl Econometrics 15(5):447-470
McFadden D, Bemmaor A C, Caro F G, Dominitz J, Jun B H, Lewbel A, Matzkin R L, Molinari F, Schwarz N, Willis R J, Winter J K (2005) Statistical Analysis of Choice Experiments and Surveys. Market Lett 16(3):183-196
McMahon S R, Iwamoto M, Massoudi M S, Yusuf H R, Stevenson J M, David F, Chu S Y, Pickering L K (2003) Comparison of e-mail, fax, and postal surveys of pediatricians. Pediatrics 111(4):299-303
Meijer E, Rouwendal J (2006) Measuring welfare effects in models with random coefficients. J Appl Econometrics 21(2):227-244
Meyerhoff J, Liebe U (2006) Protest beliefs in contingent valuation: Explaining their motivation. Ecol Econ 57(4): 583-594
Mitchell R C, Carson R T (1989) Using Surveys to Value Public Goods: The Contingent Valuation Method. Resources for the Future, Washington DC, USA
Mitofsky W J (1999) Pollsters.com. Publ Perspect 10: 24-26
Morrison M D, Blamey R K, Bennett J W (2000) Minimising Payment Vehicle Bias in Contingent Valuation Studies. Environ Resource Econ 16(4): 407-422
Muffo J M, Sinclair A, Robson V (2003) A Comparison of Web Versus Paper Alumni Surveys. Paper presented at the Annual Forum of the Association for Institutional Research, Tampa, Florida, 20 May, 2003
Olsen S B (2007) Internet vs mail – Are stated preferences affected by the mode of sampling in a choice experiment. Paper presented at the EAERE 15th Annual Conference 2007, 27-30 June, Thessaloniki, Greece
Ottens M (2006) Use of the Internet among individuals and enterprises. In: Statistics in focus. Available via Eurostat. http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NP-06-012/EN/KS-NP-06-012-EN.PDF. Cited 26 July 2007
Poe G L, Severance-Lossin E K, Welsh M P (1994) Measuring the Difference (X–Y) of Simulated Distributions: A Convolutions Approach. Am J Agr Econ 76(4):904–915
Poe G L, Giraud K L, Loomis J B (2005) Computational Methods for Measuring the Difference of Empirical Distributions. Am J Agr Econ 87(2):353-365
Revelt D, Train K (1998) Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level. Rev Econ Stat 80(4):647-657
Rigby D, Burton M (2006) Modeling Disinterest and Dislike: A Bounded Bayesian Mixed Logit Model of the UK Market for GM Food. Environ Resource Econ 33(4):485-509
Scarpa R, Ferrini S, Willis K (2005) Performance of Error Component Models for Status-Quo Effects in Choice Experiments. In: Scarpa R, Alberini A (eds) Applications of Simulation Methods in

Environmental and Resource Economics. The Economics of Non-Market Goods and Resources, vol. 6. Springer, Dordrecht, The Netherlands, pp247-273
Scarpa R, Thiene M, Marangon F (2008) Using flexible taste distributions to value collective reputation for environmentally friendly production methods Can J Agr Econ 56(2):145-162
Scarpa R, Willis K, Acutt M (2007) Valuing externalities from water supply: status quo, choice complexity, and individual random effects in panel kernel logit analysis of choice experiments. J Environ Plann Manag 50(4):449-466
Schleyer T K L, Forrest J L (2000) Methods for the design and administration of web-based surveys. J Am Med Informat Assoc 7:416-425
Schmidt W C (1997) World-Wide Web survey research: Benefits, potential problems, and solutions. Behav Res Meth Instrum Comput 29:274-279
Schonlau M, Fricker Jr. R D, Elliott M N (2002) Conducting Research Surveys via E-mail and the Web. RAND Publications, Santa Monica, CA
Stanton J (1998) An empirical assessment of data collection using the Internet. Person Psychol 51:709-725
Swait J, Louviere J (1993) The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. J Market Res 30(3):305-314
Train, K (2001) A comparison of hierarchical Bayes and maximum simulated likelihood for mixed logit. Working paper. Department of Economics, University of California, Berkeley, USA.
Train K (2003) Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge, UK
Train K, Sonnier G (2005) Mixed Logit with Bounded Distributions of Correlated Partworths. In: Scarpa R, Alberini A (eds) Applications of Simulation Methods in Environmental and Resource Economics. The Economics of Non-Market Goods and Resources, vol. 6. Springer, Dordrecht, The Netherlands, pp117-134
Train K, Weeks M (2005) Discrete Choice Models in Preference Space and Willing-to-pay Space. In: Scarpa R, Alberini A (eds) Applications of Simulation Methods in Environmental and Resource Economics. The Economics of Non-Market Goods and Resources, vol. 6. Springer, Dordrecht, The Netherlands, pp1-16
Tsuge T, Washida T (2003) Economic valuation of the Seto Inland Sea by using an Internet CV survey. Mar Pollut Bull 47(1-6): 230-236
Weible R, Wallace J (1998) Cyber research: The impact of the internet on data collection. Market Res 10(3):19-24
Wygant S, Lindorf R (1999) Surveying collegiate Net surfers - Web methodology or mythology. In: Quirk's Marketing Research Review. http://www.quirks.com/articles/a1999/19990706.aspx?searchID=955999. Cited 26 July 2007
Yun G W, Trumbo C (2000) Comparative responses to a survey executed by post, email and web form. J Comput-Mediated Comm 6(1)
Zwerina K, Huber J, Kuhfeld W (1996) A General Method for Constructing Efficient Choice Designs. SAS Institute Inc. http://support.sas.com/techsup/technote/ts650/ts650D.pdf. Cited 26 July 2007

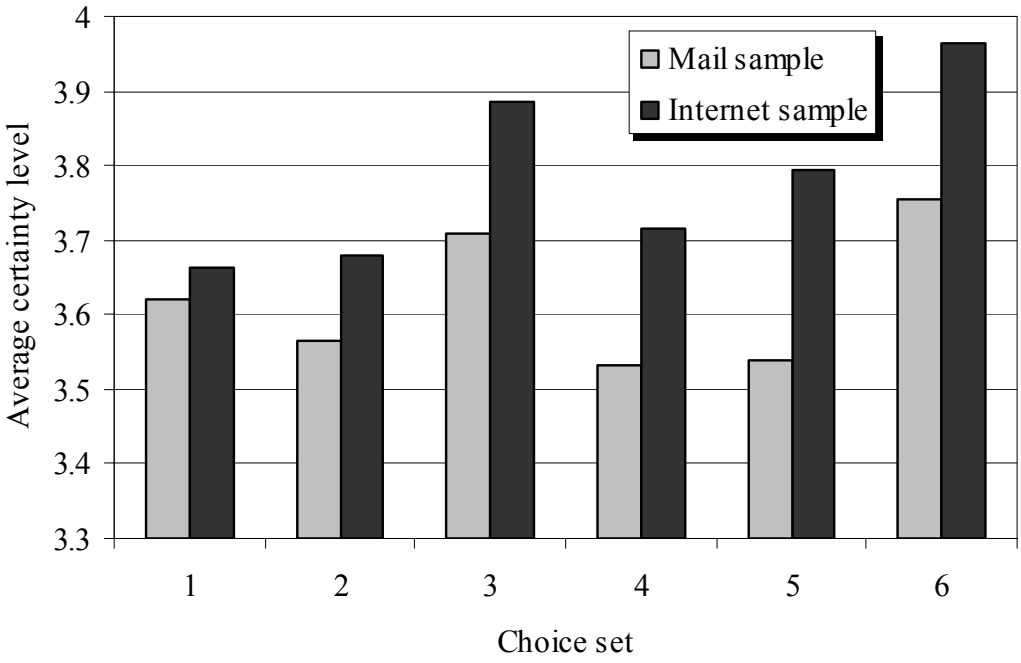

Figure 1. Average levels of self-reported certainty in choice. Note that the origin of the vertical axis is not 0.

Table 1
Attribute levels used in the choice experiment survey

Attribute (type of landscape)	Level (km new motorway through landscape)
Forest	0 km, 5 km, 10 km
Wetland	0 km, 2.5 km, 5 km
Heath/pastoral area	0 km, 2.5 km, 5 km
Arable land	80 km, 82.5 km, 85 km, 87.5 km, 90 km, 92.5 km, 95 km, 97.5 km, 100 km
Annual extra tax payment per household	(0 DKK) [a], 100 DKK, 200 DKK, 400 DKK, 700 DKK, 1100 DKK, 1600 DKK

Note: 100 DKK ≈ €13.4 ≈ $16.8
[a] The status quo cost at 0 DKK was only used in the status quo alternative.

Table 2
Response rates

	Internet		Mail	
	#	%	#	%
E-mails / Questionnaires sent out	253	100	300	100
- No reply (unreturned)	92	36.4	111	37.0
- Undeliverable / returned but completely unanswered	- [(a)]	-	5	1.6
- Returned but choice sets unanswered	-	-	3	1.0
Respondent sample	161	63.6	181	60.3
- Identified protest zero bidders	9	3.6	26	8.7
- Identified irrational bidders	4	1.6	3	1.0
Effective sample (used for analyses)	148	58.5	152	50.7

[a] It was not possible to determine how many respondents refused to answer or only partially answered the internet questionnaire, as the final datasets supplied by ACNielsen did not contain this information.

Table 3
Distributional comparison of socio-demographic variables

	Silkeborg area in general %	Initial sample[a] %			Respondent sample %			Effective sample%		
		Internet	Mail	χ^2-value	Internet	Mail	χ^2-value	Internet	Mail	χ^2-value
Gender										
Male	49.4	44.7	49.7	$3.0^{ns\ (b)}$	44.7	48.0	0.8^{ns}	44.6	46.0	0.1^{ns}
Female	50.6	55.3	50.3	$2.3^{ns\ (c)}$	55.3	52.0	1.4^{ns}	55.4	54.0	1.4^{ns}
				$0.0^{ns\ (d)}$			0.1^{ns}			0.7^{ns}
Age										
18 - 24	10.8	7.9	8.7		5.0	6.7		5.4	7.3	
25 - 34	19.7	20.2	13.0	56.9^{***}	17.5	14.0	19.5^{**}	17.7	14.7	11.0^{ns}
35 - 44	22.5	34.1	23.0	25.4^{***}	36.3	26.8	24.5^{***}	34.7	28.0	19.4^{***}
45 - 54	20.5	20.2	22.3	29.9^{***}	23.8	22.3	12.7^{*}	24.5	22.0	6.7^{ns}
55 - 64	19.2	13.9	18.7		13.1	18.4		13.6	18.7	
65 - 70	7.2	3.6	14.3		4.4	11.7		4.1	9.3	
Household gross income (DKK)										
< 150,000	14.6	1.8	-	-	2.0	6.9	6.10^{ns}	1.5	8.2	8.4^{*}
150,000 - 299,999	28.1	15.2	-	65.2^{***}	14.9	13.1	45.2^{***}	14.1	11.9	46.2^{***}
300,000 - 499,999	22.1	30.5	-	-	27.0	26.9	35.6^{***}	26.7	23.9	32.8^{***}
> 500,000	35.2	52.5	-		56.1	53.1		57.8	56.0	
Education										
Primary school	32.1	7.2	-		6.7	11.7		6.6	10.7	
Vocational	37.1	31.2	-		30.2	36.3		28.5	32.7	
High school	7.4	4.2	-	-	4.0	3.9	19.29^{**}	4.4	3.3	11.1^{*}
University, <3 yrs	4.8	18.1	-	200.1^{***}	20.8	12.3	157.4^{***}	21.2	14.0	157.0^{***}
University, 3-4 yrs	14.6	27.8	-	-	24.8	26.3	77.2^{***}	24.8	28.0	90.7^{***}
University, >4 yrs	4.0	11.4	-		13.4	9.5		14.6	11.3	

[a] The demographics of the initial sample can be fully assessed for the internet sample as internet panellists have previously answered a range of socio-demographic questions when accepting to join the panel. For the initial mail sample, the Danish Civil Registration System was able to inform about gender and age of the randomly drawn respondents, whereas household gross income and educational status were not available.

[b] The first χ^2-value concerns the comparison of distributions in the internet and the mail sample. It is calculated on the basis of the actual number of respondents in the mail sample compared to the number of respondents in the internet sample adjusted to reflect the total number of respondents in the mail sample. . '***' indicates significantly different distributions on the 99.9% level, '**' on the 99% level, '*' on the 95% level, and 'ns' indicates no significant difference between the two samples.

[c] Similarly, the second χ^2-value concerns the internet sample compared to the area in general,

[d] and the third χ^2-value concerns the mail sample compared to the area in general.

Table 4
Parameter estimates with robust standard errors in parentheses.

Attribute		Internet	Mail	Pooled	Pooled (scaled)
Fixed parameters					
ASC (Status quo)		-0.578 (0.301)	-0.885 (0.453)	-0.892 (0.229)	-0.413 (0.226)
Price [a]		-0.272 (0.033)	-0.287 (0.052)	-0.292 (0.026)	-0.223 (0.027)
Random parameters					
Heath, β_1	Mean	0.008 (0.058)	-0.053 (0.046)	0.032 (0.044)	-0.052 (0.025)
	St.dev.	0.489 (0.080)	0.147 (0.126)	0.498 (0.092)	0.106 (0.044)
Forest, β_2	Mean	-0.134 (0.022)	-0.163 (0.041)	-0.139 (0.027)	-0.131 (0.023)
	St.dev.	0.202 (0.068)	0.378 (0.097)	0.299 (0.044)	0.215 (0.047)
Wetland, β_3	Mean	-0.267 (0.052)	-0.279 (0.051)	-0.272 (0.045)	-0.222 (0.035)
	St.dev.	0.251 (0.097)	0.021 (0.106)	0.364 (0.081)	0.043 (0.060)
Error component					
Policy alternatives, μ_{12}		0.166 (0.332)	3.560 (0.978)	0.635 (0.653)	2.380 (0.340)
Random parameter correlations [b]					
$\beta_1\beta_2$		0.319 (0.061)	0.239 (0.084)	0.315 (0.067)	0.164 (0.027)
$\beta_1\beta_3$		0.149 (0.122)	0.331 (0.218)	0.410 (0.093)	0.257 (0.068)
$B_2\beta_3$		0.256 (0.135)	-0.072 (0.098)	0.118 (0.077)	-0.115 (0.081)
Scale ratio, λ [c]		-	-	-	1.33 (0.179)
No. of observations		888	906	1794	1794
McF pseudo-R^2		0.336	0.435	0.372	0.384
LL at convergence		-648.0	-562.6	-1237.3	-1213.6
Likelihood Ratio test statistic				53.19	5.79

[a] For computational reasons all prices entered in the dataset have been rescaled with a factor 1/100, hence price parameter estimates obtained from the parametric models should accordingly be divided by 100.
[b] Below-diagonal elements from the Cholesky decomposition matrix are reported to identify cross-correlation patterns among the random parameters which are unconfounded with the standard deviation parameter estimates of each of the random parameters (Hensher et al. 2005).
[c] The scale ratio is the scale factor of the mail sample relative to that of the internet sample which is normalized to 1. Testing against this value, the t-test value returned is 1.83 which is significant at the 90% significance level.

Table 5
Mean unconditional Willingness-To-Pay estimates and 95% Confidence Intervals for avoiding 1 km of motorway through landscape types (DKK per household per year).

	Internet		Mail		P-value[b]	
	WTP	95% CI[a]	WTP	95% CI	t-test	CC
Heath	-2.7	[-44.3;41.1]	18.6	[-11.7;60.3]	0.444	0.213
Forest	49.3	[32.8;70.2]	56.8	[35.0;79.5]	0.607	0.322
Wetland	98.2	[63.5;138.8]	97.2	[57.1;157.8]	0.977	0.503

[a] The confidence intervals are obtained using the Krinsky-Robb procedure (Krinsky and Robb 1986;1990) with 10000 replications. This method is also referred to as the parametric bootstrap.
[b] P-values reported for an asymptotic t-test of equality in means as well as the Complete Combinatorial test proposed by Poe et al. (1994; 2005). While the former reports the two-sided significance level, the latter reports a one-sided approximate significance level.