Comparing the Conventional Stated Preference Valuation Technique with a Prediction Approach.

Lava Yadav\textsuperscript{1}, Tom van Rensburg\textsuperscript{1} and Hugh Kelley\textsuperscript{1}

\textsuperscript{1}Department of Economics, National University of Ireland, Galway, Ireland
\textsuperscript{*}lava.yadav@nuigalway.ie

Copyright 2010 by Lava Yadav. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

ABSTRACT:

Stated preference techniques have been used to place values on public goods by directly asking individuals to provide their personal values and opinions. This method has consistently resulted in the emergence of hypothetical bias. Several insights from the psychology literature suggest that social desirability bias, a contributor to hypothetical bias, occurs when individuals face such direct questions. However, replacing the direct questions with an indirect one that asks for their predictions about other’s values can potentially eliminate this bias. In this study we employ both questioning formats in a choice experiment to make comparisons between the stated responses. Predicted willingness to pay is 2.5 and 3.1 times smaller than hypothetical values indicating predictions to be a more accurate measure of actual values. The study further highlights the vulnerability of the conventional approach to a social desirability bias as it allows normative motives to distort respondents’ decisions, which in turn generates preferences for environmental attributes that are misleading.

Keywords: Stated Preference Techniques, Discrete Choice Experiments, Hypothetical Bias, Social Desirability Bias, Lake Wabegon Effect, False Consensus Effect, Willingness to Pay
I. INTRODUCTION

Conventional stated preference techniques typically request individuals to state their hypothetical payments for the hypothetical provision of a non-market good (Murphy, et al. 2005; Harrison and Rutstrom 2008). There is substantial evidence that these hypothetical values often exceed actual payments (List and Gallet 2001; Murphy, et al. 2005; Harrison and Rutstrom 2008). In recognition of this problem, the National Oceanic and Atmospheric Administration (NOAA) guidelines state that values elicited using stated preference techniques should be halved unless adjusted using some form of calibration technique (Arrow, et al. 1993). As a result several attempts to elicit unbiased responses through the use of calibration techniques have been made; however, none have succeeded in providing convincing results (Murphy and Stevens 2004; Murphy, et al. 2005).

Lusk and Norwood (2009a) attribute the hypothetical bias problem not to the elicitation format but to the direct questioning approach therein. When inquiring about one’s personal level of contribution to a public good, an individual is said to achieve additional utility simply by stating that they are willing to pay for it. As such, without having to pay for the good, if simply stating a higher verbal willingness-to-pay response increases your utility, it is likely that a respondent will have a tendency to do so (Harrison and Rutstrom 2008).

The misrepresentation of true opinions and values are carried out either because people are too optimistic about themselves (Winters and Neale 1985) or they want to present themselves in a morally or socially acceptable manner (Paulhus 1991). However, this tendency towards falsification, termed “Social Desirability Bias” has been observed to disappear when replacing the direct questions (for example “What do you think? or What are you willing to Pay?”) with indirect ones (for example “What would most people think? or What do you think others are willing to Pay?”) (Fisher 1993; Epley and Dunning 2000; Lusk and Norwood 2009a, b). The indirect questioning approach removes the ‘spotlight’ from the respondent which enables them to feel less conscious about their responses. Also, when making statements about other people, the additional utility obtained by envisioning oneself getting involved is removed.

With the indirect questioning approach, Epley and Dunning (2000) highlight the recurring phenomenon known as the “Lake Wabegon Effect”; where people tend to underestimate the generosity and ability of others in relation to their own. According to their study results, self assessments represent overly optimistic views of the self while the seemingly underestimated predictions more accurately reflect actual behavior of the public. The Lake Wabegon effect essentially relates to a self-serving prediction behavior that parallels the social desirability bias.

The estimation of the total value of a good involves the aggregation of individual values of the public. The use of the indirect questioning approach may be able to remove social desirability bias, but do they accurately represent those of the public? The theory related to the “False Consensus Effect” claims that the predicted values in fact manifest from personal opinions and

As a result of the false consensus effect, a positive correlation between the outcomes of the direct and indirect question is likely to exist (Van Boven, et al. 2000, 2005). The level of correlation however is dependent upon the nature/sensitivity of the question or the type of good that is being valued. Concerns about the social or moral consequences regarding one’s behavioral decision is said to be responsible for social desirability bias in direct questions. However, if the decision has no such normative affiliations, then there is no motivation for distorting true values and opinions (Miniard and Cohen 1983; Fisher 1993). Thus it is the level of normative and personal component perceived when making a decision that determines the disparity in the outcomes between the direct and indirect questions. As a result, the revealed values and opinions from direct questioning may be significantly misleading if the respondent perceives a large degree of normative consequences associated.

Through this study we would like to contribute to the area of research initiated by Lusk and Norwood (2009a) with a field experiment that employs a predicted/inferred valuation method. This paper reports the results of a choice based valuation study of the Irish Burren landscape incorporating both the conventional (direct questioning approach) as well as an inferred valuation technique (indirect questioning approach). The valuation question asked respondents to indicate their willingness to pay values for the conservation of rocky limestone pavements and orchid rich grasslands. A comparison of the values observed from the two techniques shows that the inferred valuation approach is capable of effectively reducing hypothetical bias. We find that the ratio between the marginal willingness to pay (WTP) values acquired through the direct and the indirect questioning approaches for the two environmental public goods in this study are 2.5 and 3.1 – which are very close to the average ratios between hypothetical and non-hypothetical values as reported on the Meta-analysis by Murphy, et al. (2005).

Direct and indirect questioning methods are also used to compare outcomes of several attitudinal questions. Significant differences are observed between the responses from the two formats confirming the presence of social desirability bias using direct questions. We also observe positive correlation between the responses from the two approaches indicating the presence of a false consensus effect – people making predictions about others based upon their own preferences.

An inconsistency is observed between the self and predicted preferences for the two environmental attributes. This difference is found to be a result of additional normative motives perceived when contributing to the conservation of the orchid rich grasslands over the rocky limestone pavements. As such, this paper highlights the risks of being misled by outcomes of direct questioning when dealing with environmental goods; since they are generally associated with high levels of normative consequences.
II. BACKGROUND

Environmental Valuation and Hypothetical Bias

In spite of the overwhelming evidence of hypothetical bias in stated preference studies the quest for a solution still goes on. This pursuit for the true value is based upon the following two assumptions regarding human preferences (for environmental goods).

1. Individuals are well aware of their preferences and have the ability to forecast their abilities and tastes.
2. These preferences can be elicited accurately by directly requesting them from individuals through the means of surveys and interviews.

With the belief that hypothetical bias is an elicitation problem, the valuation literature has gone to great lengths in its attempt to develop the best technique in facilitating the elicitation of true values. More reliable techniques regarding incentive compatibility have been experimented with in order to combat possible strategic responses, which some believe as being responsible for hypothetical bias (Harrison and Rutstrom 2008). The refinement process has also witnessed the development of calibration techniques such as cheap talk scripts (Cummings and Taylor 1999), uncertainty adjustments (Champ and Bishop 2001; Poe, et al. 2002), and statistical bias functions (Hofler and List 2004). However, none have succeeded in providing convincing results (Murphy and Stevens 2004; Murphy, et al. 2005).

Although the problem of hypothetical bias is far from being solved, the refinements of valuation techniques have helped mitigate various types of biases. In recognition of the immense task a respondent faces in processing new information about public goods (which one is not used to valuing) and then formulating a value at the given time, newer techniques such as choice experiments have been utilized to simulate real word scenarios.

Today choice experiments (CE) are regarded as most reliable amongst stated preference techniques available due to the many advantages it has over other methods (see Hanley, et al. 1998). List, et al. (2006) claim that the greater levels of information provided in CEs are capable of mitigating many types of biases including hypothetical bias. Additionally, a study by Huber, et al. (2002) show that respondents find the CE technique to be more realistic and also feel more confident when making decisions. Regardless of these claims, the continual presence of hypothetical bias with CEs has been observed (Cameron, et al. 2002; Lusk and Schroeder 2003; Carlsson, et al. 2005; List, et al. 2006; Johansson-Stenman and Svedsater 2007; Ladenburg, et al. 2007; Carlsson, et al. 2008).
Conventional Valuation Technique: Direct Questions

The primary assumption that individuals are aware of their own preferences and are thus capable of predicting how they would behave when faced with making a decision in a real situation is challenged by studies that have observed the contrary. Individuals have been known to often fail in forecasting their own behavior and tastes (see for example Ausubel 1991; Loewenstein and Adler 1995; Tat, et al. 1998). Wilson and LaFleur (1995) find predictions of own behavior to be less accurate, and more so when allowed additional time to analyze reasons behind their predictions. Even when people are extremely confident about personal decisions it is still possible that they do not comply with their stated intentions. In a study by Vallone, et al. (1990) which focused on the confidence level of individuals, although some individuals were 100 percent certain about performing a particular behavior, it was performed only 77 percent of the time.

It may be the case that we are only capable of knowing our preferences for goods that are very familiar to us but have to construct new ones for the rest (Slovic 1995). In the context of being asked to value an environmental good (a task one is not familiar with), respondents struggle to understand the type of information that is presented and try to relate this information to previous personal encounters in order to construct a value for it (Vatn 2004). It is at this point that even without the intention of fabricating opinions and values, respondents may often provide a biased response by providing what is morally or socially acceptable. Moreover, respondents may state their opinion regarding whether the good being valued is worthy of support (Brown, et al. 2003). In such cases the results reveal attitudes rather than actual preferences (Kahneman and Sugden, 2005).

However, the disparities in actual and stated behavioral choices are not necessarily random; rather, there are a number of biases responsible for inflating and deflating stated choices. Lusk and Norwood (2009a) explore whether some of these biases are associated with the additional satisfaction individuals receive simply by stating that they are willing to pay. Similarly, Johansson-Stenman and Martinsson (2006) and Brekke, et al. (2003) suggest that utility is derived not only from the consumption of a good but also from the positive self image people perceive to have attained in the process.

This tendency to distort true values, known as social desirability bias is believed to be motivated by two factors (Paulhus 1984). The first factor relates to the portrayal of an overly optimistic view of the self which the individuals believe to be true. The second is driven by the propensity to present oneself in a socially conventional or optimal way in order impress observers.

Consequently evidence exists where individuals regularly distort their behavioral choices to confirm with societal norms (Streb, et al. 2008). Legget, et al. (2003) observe a social desirability bias to be responsible for higher WTP estimates in the presence of an interviewer. List et al. (2004) report a relation between the level of anonymity of the elicitation formats and willingness
to pay estimates. Similarly Levitt and List (2007) detail studies where an effect is observed on respondent behavior as a result of being watched. As such, the contemporary belief in the field of social psychology acknowledges that individuals have a biased view of themselves as a result of unconsciously or subconsciously engaging in self-deception (Taylor and Brown 1994).

The Prediction Approach: Indirect Questions

In recognition of the social desirability bias, several remedies have been proposed and experimented with (see Lusk and Norwood 2009b). In this study we examine one of the methods involving an indirect questioning approach, which has only recently been introduced for non-market valuation (Lusk and Norwood 2009a, Carlsson, et al. 2008).

Unlike conventional valuation techniques that directly inquire respondents to state their own values and opinions, this technique asks them to make predictions or inferences about other’s values. Essentially what this indirect questioning approach does is removes the ‘spotlight’ from the individual, which frees them from the belief that they are being judged. As a result they are able to make predictions about others without having the need to impress. Also, the additional utility respondents receive from overly generous or optimistic views about themselves do not exist when the inferences are made about someone else.

Several studies have shown significant difference in responses between the direct and indirect questioning approaches (Fisher 1993; Epley and Dunning 2000, Johansson-Stenman and Martinsson 2006; Lusk and Norwood 2009a, b, c). According to the model specified by Miniard and Cohen (1983), the disparity between the outcomes from the two questioning formats is characterized by the level of personal and normative reasons for making a behavioral choice. The personal motivations relate to those expectations that are dependent upon the intrinsic worth of the outcome to the person. Normative consequences on the other hand are governed by what the individual believes are expected by other people or the society.

The influence of normative vs. personal outcomes on direct and indirect questions was studied by Fisher (1993) through a number of experiments. According to his results personal outcomes were not influenced by the type of question asked, however, a significant difference was observed between the two when inquiring about normative outcomes. Respondent anonymity was found to have an effect on the responses to normative outcomes when direct questioning was used, but not with indirect questions. Through several other experiments Fisher (1993) concludes that while the indirect questioning approach is potentially immune to social desirability bias, direct questions remain vulnerable.

A common observation revealed by studies analyzing responses to indirect questions show that individuals consistently believe others to behave in a manner that is less kindly, less generous, or even less reasonable (Goethal 1986; Goethals, et al. 1991; Epley and Dunning 2000; Johansson-Stenman and Martinsson 2006; Lusk and Norwood 2009b). This phenomenon is commonly termed as the ‘Lake Wabegon Effect’ or the feeling of being ‘holier than thou’. Goethals, et al.
(1991) showed that people believed they were more likely than their peers to give up their seat on a crowded bus for a pregnant woman. Johansson-Stenman and Martinsson (2006) found individuals to believe that they were highly concerned about the environmental performance of a car they were purchasing but not as concerned about its status value (Johansson-Stenman and Martinsson 2006). On the contrary when making predictions about others, they expected relatively less people to be concerned about the environmental performance and instead be more concerned about the status value. A recent study on animal welfare showed that only 15.6% of the public considered low meat prices to be more important than the well-being of the farm animals but predicted 67.5% of the American public to think otherwise (Lusk and Norwood 2009b).

Epley and Dunning (2000) state that while self assessments represent overly optimistic views of the self, the seemingly underestimated predictions more accurately reflect actual behavior of the public. In one of their experiments they observed 83% of students stating that they would buy at least one daffodil flower as a contribution to the annual campus day charity while they expected only 56% of their peers to do the same. In reality only 43% of students ended up buying a daffodil flower indicating a higher level of accuracy for the predicted values. A similar result was found in a prisoner’s dilemma type game where 84% of students predicted they would cooperate and expected only 64% of their peers to do the same. The end result saw only 61% cooperating.

According to Epley and Dunning (2000), individuals are excessively optimistic about their own behavior and consider themselves to be a member of ‘an elite moral minority’. Although we have an idea of how other people normally behave in certain conditions and situations, when estimating our own behavior we tend to believe that such rules and patterns do not apply to the self. Additionally, we fail to recognize in others the various self-conscious emotions that influence our own decisions. As such, the anomalous outcomes as a result of the endowment effect is observed to diminish when individuals are asked to act on behalf of others rather than on their own interests (Marshall, et al. 1986; Van Boven, et al. 2000)

The prediction approach is not a new concept; rather, its potential has been well recognized and employed in various areas. Psychologists and marketers have been well aware of its potential benefits and have used them since the 1950s (Maccoby and Maccoby 1954; Westfall, et al. 1957). Since its implementation in Iowa Electronic Markets it has constantly outperformed the results of polls (Berg, et al. forthcoming; Forsythe, et al. 1992). The mechanism behind opinion polls are similar to that of conventional stated preference techniques wherein they both ask individuals to make self assessments (how they themselves would act or respond). As prediction markets seem to have done better than opinion polls, it shouldn’t be a surprise that the prediction mechanism might also outperform the conventional stated preference approach.
When estimating a value for a non-market good, we are essentially interested in the personal preferences held by individuals for the commodity in question. The total value is acquired by aggregating these personal preferences across the diverse range of individuals that represent the public. Hence, it is necessary to ensure that by altering the elicitation format to one that requests for predictions, personal preferences of the respondents are not compromised.

A direct link between personal preferences and those inferred about others is drawn by the “False Consensus Theory”. According to Mullen, et al. (1985), “False Consensus refers to an egocentric bias that occurs when people estimate consensus for their own behaviors.” As a result, people have a false notion that more people are like themselves (Ross, et al. 1977). Personal values and opinions play an important role when making predictions about others (Dawes 1989); so when no information is provided regarding others’ behavior, personal preferences are regarded as being a strong representation for the general public. As such when making predictions about others, individuals are believed to be “…describing their own feelings behind a façade of impersonality” (Simon and Simon 1975, p. 586).

A Meta analysis of 115 studies by Mullen, et al. (1985) demonstrates the occurrence of this phenomenon for a wide range of preferences and beliefs. A study by Katz and Allport (1931) demonstrated that the students who admitted to cheating predicted a higher percentage of other students that cheated. Fisher (1993) also observed respondents to project their personal beliefs when stating their responses to indirect questions.

To our knowledge the study by Lusk and Norwood (2009a) is the only published work that has incorporated the prediction based technique for non-market valuation. They conduct lab experiments to test whether predictions (inferred values) are capable or removing social desirability bias in addition to a hypothetical bias. In their experiments subjects are asked to vote on a majority rule referendum, whether or not to pay $3.00 to save a rare plant (a “bright lights” Swiss Chard plant). According to their results, predicted values are very similar to those revealed in the non-hypothetical treatment; thus successfully eliminating hypothetical bias. However they claim that actual consumption values are lower than non-hypothetical values as even non-hypothetical values suffer from a social desirability bias.

Carlsson, et al. (2008) use choice experiments to compare donation levels between a non-hypothetical, hypothetical and prediction treatments. A strong hypothetical bias is reported with an average stated WTP value of 1.64 SEK and revealed WTP value of 0.427SEK. However with a low predicted average WTP value of 0.155 SEK they conclude that individuals considerably underestimate other’s generosity in order to bolster their own self image.
III. STUDY AREA

The Burren

The Burren is regarded as one of Europe’s most important and most widely recognized landscapes. Located in the west of Ireland in counties Clare and Galway, it spans across an area approximately 720 km$^2$. Two of the most prominent habitats of the Burren region include the rocky uplands with its mysterious landscape formed by karst limestone pavements and the orchid rich grasslands. In recent decades these habitats, which have been maintained through traditional farming practices have come under threat as a result of evolving market and social structures which have altered farming practices. As such the revival of farming practices that resemble the traditional ones are necessary to protect these habitats.

IV. STUDY DESIGN AND METHODOLOGY

Experimental Design

Image manipulation software was used to create photomontages to aid the written descriptions of the potential outcomes resulting from proper management (and lack of management) for the two landscape features. Starting with a ‘control’ photograph for each of these habitats, various features were manipulated to provide respondents with a visual representation of changes that potentially occur to the landscapes with and without management$^1$. The cost attribute was described as the ‘Expected Annual Cost’ to the respondent of implementing the respective management practices as shown in the choice sets.

The objective of the survey and the method of implementation were discussed with several members of the public in a focus group that was conducted at the National University of Ireland, Galway. During this session the magnitude and range of the cost attribute were also tested using open ended willingness to pay questions. A few interviewers were trained before conducting a pilot survey in the city of Galway, Ireland. Following the pilot survey, the questionnaire was further refined to minimize confusions and shortened to limit the interview time to between 12 and 15 minutes.

The Choice Experiment

The three attributes included in each choice set were referred to as Landscape, Biodiversity and Expected Annual Cost. As shown in Table 1, the Landscape and Biodiversity attributes had two levels each ‘With Management’ and ‘No Management’ while there were four levels for the Expected Annual Cost (€5, €10, €20, and €40).

---

$^1$ See Appendix for a description of the landscape attributes that were provided to the respondents.
Each choice set consisted of three alternatives. The first two alternatives labeled Option A and Option B were experimentally designed while the third alternative labeled ‘Status Quo’ was fixed in every choice set. The Status Quo alternative represented a scenario with no management in either of the attributes and was associated with ‘zero’ expected annual cost. While this was the case for the Status Quo alternative in every single choice set, the other two alternatives were allowed to vary.

Before the choices were made, respondents were familiarized with the two attributes and their likely conditions with and without management. They were then provided with a sample choice task and were told that the alternatives represented the Government’s available environmental policy options. The respondents were made aware of the additional costs associated with maintaining good environmental standards through the following statements.

“Maintaining good environmental standards and keeping the management practices in place requires financial support. So each of the management options also has a particular cost involved.”

Respondents were reminded that the Expected Annual Cost attribute represented a monetary value that the respondent would personally have to pay per year through increased Income Tax and Value Added Tax. The respondents were then provided with a sequence of four different choice tasks and asked to choose his/her preferred alternative in each case. Upon the completion of the four choice tasks, they were given four more choice tasks in which they were asked to make predictions of what they believed others would choose when given those very choices. The following statement was provided to the respondents requesting their predictions.

“I will now present to you another series of choices just like the ones you were shown earlier. This time instead of making your own choices, I would like you to predict the choices you think most people would make. [On average what would the general public choose?]”

About 50% percent of the respondents were asked to make predictions before making decisions for themselves in order to account for order effects.

Model

A standard random utility model is employed in the analysis of the choice data.

\[ U_{ij} = \theta_i + \alpha_j \text{Rock}_i + \beta_j \text{Grass}_i + \gamma_j \text{Cost}_i + \varepsilon_{ij} \]

The equation above represents the utility individual \( k \) receives from alternative \( i \). \( \Theta_i \) represents an alternative specific constant. As our experiment was unlabelled we only include this constant for the status quo alternative. \( \text{Rock}_i \) and \( \text{Grass}_i \) indicate whether or not the karst limestone pavements and the orchid rich grasslands are managed, \( \text{Cost}_i \) represents the expected annual cost of

---

2 See Appendix for the Sample Choice Task shown to respondents
implementing the management schemes. $\alpha_j$, $\beta_j$, and $\gamma_j$ represent the corresponding parameters and $\varepsilon_{ij}$, the error term. The utility model is identical for individuals acting on their own interests as well as when making predictions about others. The model is estimated using a mixed logit model with simulated maximum likelihood using Nlogit 4. Using the utility specification above, the marginal willingness to pay for managing the karst limestone pavements and the orchid rich grasslands are provided by the ratio of the corresponding coefficients and the cost coefficient.

Data and Sampling Method

A total of four trained interviewers administered the survey between July 2009 and October 2009 in six counties. Most of the 300 interviews were conducted in county Galway followed by Limerick and Clare. The in-person interviews were conducted at the respondents’ homes that were randomly chosen.

Of all the respondents surveyed, 27 percent resided in the countryside. Almost all the respondents were aware of the Burren region and slightly over 70 percent had visited the area within the last five years. The average respondent age was 39.3 years with an average wage of €37,399. The average attained education of 4.65 signified a level between a high school degree and a college degree.

V. RESULTS

Attitudinal Questions

We first report the results of some attitudinal questions the respondents were asked regarding their opinions towards the rural environment and willingness to pay for them. After respondents were presented with the set of four choice tasks where they were asked to indicate their preferred alternative, they were inquired about a number of attitudinal questions to which they responded by rating on a five point likert scale. They were asked to state whether or not they agreed with the statements provided by indicating 1 if they strongly disagreed and 5 if they strongly agreed. Consequently, after making predictions about others’ choices on an additional four choice tasks, they were presented with the same attitudinal questions, but were asked to indicate what they thought about others’ opinions and views.

---

3 In order to account for distance decay effects (Hanley, et al. 2003) counties Dublin, Westmeath and Sligo (counties that do not border the Burren region) were also included as survey sites.
4 The income and age questions had classes. For the estimation, the midpoints were used
5 Education (Primary = 1, Junior Certificate = 2, Leaving Certificate = 3, On the job training/professional qualification of degree level = 4, College/University Degree (B.Sc., B.A., etc) = 5, Post graduate (M.Sc., Ph.D., etc.) = 6)
The analysis was conducted on the responses provided by a total of 292 respondents. Table 2 shows the different statements that were presented to the respondents and the average scores provided for each of them with regard to their personal opinions as well as their beliefs about others. According to the $\chi^2$ tests the scores stated for the self are significantly different from the scores predicted for others.

The predicted scores are significantly higher than personal scores for the first two statements. This difference suggests that people assume money to be less of a concern for them but more of a priority for others. Consequently the predicted scores are significantly lower for the last three statements indicating that people consider themselves to care more about the environment than others. These results are in accordance with previous studies on social desirability bias and the Lake Wabegon effect.

Spearman’s tests of correlations are conducted to see if the scores from the two questioning formats are related. We find a significant positive correlation for all of the questions which confirms the existence of the false consensus effect. This shows that in spite of the underestimated predictions, respondents do identify themselves with others.

Additionally we test to see the presence of any order effects on the predicted scores. Only the second statement “Most people object to paying higher taxes” revealed a statistically significant disparity in terms of the likert scale ratings. People expected others to be less objectionable to paying taxes if they were first asked to state their personal response to a similar question. This is likely due to a further dominance of the false consensus effect wherein explicitly stating one’s personal values allows for a more direct transfer of personal opinions for making predictions about others.

**Willingness to Pay Values**

Next, we examine the stated choices by employing discrete choice models. We start off with a standard multinomial logit model to analyze the data. Separate status quo alternative specific constants are included to allow for different status quo effects for the self and for those predicted. To capture observed heterogeneity, socio-economic variables are included by interacting them with the status quo alternative specific constants. The validity of the IIA assumption of the MNL model is assessed using the Hausmann-McFadden tests. According to the results, the IIA

---

6 Eight respondents were removed from the analysis as they genuinely preferred the Burren region to be stripped of any forms of management. They claimed to prefer a landscape that existed before any farming practices existed (a landscape reforested by pine and dominant shrub species, hazel and blackthorn). A preference of no management was stated by choosing the status quo option for each choice task; however, they believed others to respond in favor of management; acknowledging their unique preferences. For purposes of consistency in comparing decisions for the self versus predictions of others we chose to remove them from the analysis.

7 The average predicted score for “Most people object to paying higher taxes” was 4.8 when respondents were asked to make predictions first. This score was lowered to an average of 4.55 when the predictions were preceded by direct questions. With a $\chi^2$ value of 13.14 we find this difference to be significant at the 1% level of significance.
assumption holds when dropping Option A but this is not the case when Option B is dropped. Due to the lack of sufficient evidence supporting the IIA assumption, we explore an alternative discrete choice model that does not rely on the IIA property.

Following Train (1998) we allow for unobserved heterogeneity by allowing the experiment specific parameters to have a distribution. We assign a triangular distribution to these parameters and additionally impose the spread to be equal to the mean to assure the same sign for the parameter across the distribution (Hensher, et al. 2005). This ensures that the ‘Expected Annual Cost’ parameter is always negative while the ‘Karst limestone pavement’ and ‘Orchid-rich grassland’ are positive. The data enters the models in a panel structure which allows for correlation amongst repeated choices from a given individual. Finally, we allow for correlations between the two alternatives, Option A and Option B. As proposed by Scarpa, et al. (2005) we account for a difference in preferences between the status quo (taking no action) and the other alternatives (taking action) with an error component specification. Separate error components are allowed for choices made for the self and those predicted.

Table 3 reports the results of the mixed logit model estimated with simulated maximum likelihood using 1000 Halton draws. The model is statistically significant with a $\chi^2$ value of 5390.81, 18 degrees of freedom and a p-value equal to zero. A pseudo-R$^2$ value of 0.362 suggests that the overall model fit is adequate (Domencich and McFadden 1975). All the experimental specific parameters, karst limestone pavements, orchid rich grasslands, and expected annual cost have the expected signs for both the self and predicted choices. The estimated standard deviations of the random parameters are highly significant which suggests considerable unobserved heterogeneity. The positive and statistically significant coefficients for error components show that the status quo alternative is perceived to be different from the experimentally designed alternatives. The implicit willingness to pay values are estimated as the marginal rates of substitution between the two environmental attributes and the expected annual cost attribute. The mean of these values are estimated at €53.66 per year for the karst limestone pavements and €49.39 per year for the orchid rich grasslands. Similarly, the predicted values for the karst limestone pavements and the orchid rich grasslands are €21.53 and €15.89 respectively.

It is evident that the predicted values are significantly lower than the values estimated for the self with respondents consistently estimating higher WTP values for themselves in comparison to others. While we can practically assume that the self assessments are overstated (given the overwhelming evidence of hypothetical bias in stated preference studies) it is the accuracy of the predicted values that is of greater interest. Based upon the results of our model, the average WTP

---

8 The Hausman-McFadden test statistic associated with dropping Option A is 13.76. Comparing this value to the $\chi^2_{(10)}$ Critical Statistic at $\alpha = 0.05$ of 26.296 ($p=0.62$), we fail to reject the null hypothesis of IIA assumption. Similarly the test statistic associated with dropping Option B is 27.12 ($p=0.04$) which suggests the rejection of the null hypothesis of IIA assumption.

9 To account for heterogeneity, a number of different interactions between the socio-demographic variables and the attributes were examined. However, significant heterogeneity was still observed within all random parameters.
values for the self are 2.49 and 3.11 times larger than predicted values for the karst limestone and orchid rich grasslands respectively. These differences are very close to the average magnitude of 2.5 - 3 between hypothetical and non-hypothetical values for public goods as reported on the Meta-analysis by Murphy, et al. (2005). We cannot know for sure what the results associated with an identical non-hypothetical treatment would have provided us with. However assuming them to be in the range reported by Murphy, et al. (2005), our study may have successfully eliminated hypothetical bias.

**Order Effects**

Carlsson, et al. (2008) in their study report predicted donation levels to be significantly lower than those observed in both hypothetical and non-hypothetical treatments. Although the respondents making predictions of others’ values were not asked to state their own preferences, Carlsson, et al. (2008) claim that the predictions may have been heavily anchored on personal preferences. In our survey we asked individuals for both values. By comparing the predictions that were made before and after making choices for themselves, we test whether explicitly stating one’s personal preferences has an impact on the subsequent predictions.\(^\text{10}\)

Similar to the effects observed in the attitudinal questions related to the ‘Objection in paying taxes’, we examine whether allowing for individuals to explicitly state their personal choices before making predictions has an impact. We split the sample in two with respect to the order in which they were conducted and estimate two separate models using identical specifications as used in the mixed logit model in the previous section. Comparing these two models (where the parameters are not restricted by order) with the restricted model, we find a considerable improvement in the explanatory power of the model when allowing for order effects.

According to the unrestricted models the average predicted WTP values for the karst limestone pavements and the orchid rich grasslands are €18.05 and €14.32 respectively when respondents are asked to make predictions first. However these values increase to an average of €25.60 and €18.60 when followed by explicit choices for the self. The difference in predicted WTP values for both karst limestone pavements and orchid rich grasslands are statistically significant.

Unlike what was expected by Carlsson, et al. (2008) we find that explicitly stating one’s own opinions or values first does not necessarily cause respondents to increasingly underestimate the values and opinions of others. On the contrary we find that the predicted values and opinions are even closer to those of their own. This is likely due to a stronger false consensus effect that results from realizing or calculating one’s personal values which are used as legitimate information to make predictions about others. However, it is unclear whether explicitly stating one’s own behavior before making predictions allows respondents to clearly assess their own values and make appropriate adjustments for making unbiased predictions, or whether it introduces additional bias.

\(^\text{10}\) Some level of order effects were also reported in the study by Lusk and Norwood (2009c)
Self vs. Predicted Preferences for the Environmental Attributes

The strong normative consequences linked to making a decision regarding one’s WTP value for environmental goods may have led to the difference in WTP values as shown earlier. Here we compare the responses from the two questioning formats to explore consistencies in preferences between the two environmental goods.

The WTP estimates of €53.66 and €49.39 suggests that respondents favour the karst limestone pavements over the orchid rich grasslands. However this difference seems to be very minor with individuals willing to pay only 8.7% more for the karst limestone pavements compared to the orchid rich grasslands. Analyzing the predicted values of €21.53 and €15.89, we observe a consistency in terms of a higher preference for the karst limestone pavements over the orchid rich grasslands. However, respondents expect others to value the karst limestone pavements significantly more than the orchid rich grasslands and thus expect them to pay 35.5% higher for them.

This differential is further explored through the rankings between the attributes which were inquired right after the respondents made their choices. Respondents were asked to rank between the three attributes (including expected annual cost) to indicate which attribute was most important or most influential when making their decisions. In Table 4 we report the number of times each of the attributes was ranked as number one (of the highest priority). Forty three percent of the respondents claimed that the conservation of the karst limestone pavements was the most important factor when making choices for themselves. Thirty seven percent claimed it was the orchid rich grasslands while only 20% admitted ‘cost’ as having the most significant impact upon their decision. Consequently, when asked to rank the level of importance for others, the respondents predicted 18% to be primarily influenced by the karst limestone pavements, 11% by the orchid rich grasslands and 71% by cost.11 The large difference in rankings for the cost attribute further confirms the social desirability bias/Lake Wabegon effect where individuals consistently overestimate their own generosity compared to others.

Ignoring the rankings for the cost attribute we examine the relative importance between the karst limestone pavements and the orchid rich grasslands. While 57% of respondents stated that they preferred the karst limestone pavements over the orchid rich grasslands, they predicted 70% of others to have such preferences.12 Before the results for the survey were examined, it was expected that the karst limestone pavements would be preferred over the orchid rich grasslands

---

11 Ignoring or assigning less importance to the cost attribute has also been observed in other studies involving choice experiments. Instead of taking into account all the attributes in the choice set, respondents tend to base their choices only on attributes they consider relevant (Campbell, et al. 2008). As such, Campbell, et al. (2008) observed only 69% of respondents to consider the cost attribute when making their decisions. We find similar results in our study where most people claim to be least influenced by the cost attribute when making their decisions. The inconsequential nature of stated preference surveys are generally attributed to such behavior, however, we find that respondents did believe the cost attribute to have a significant effect on decisions of other people. Hence it is very likely that this behavior is another result of the social desirability bias.

12 No order effects were found for the predicted rankings or the rankings for the self.
as they are more visible and stand out to represent the Burren. While this higher preference was observed in both cases, it was significantly higher in terms of predictions. It was expected that respondents would overestimate their generosity compared to others in accordance with the social desirability bias/Lake Wabegon effect, however, contrary to the false consensus effect, an inconsistency in preferences for the karst limestone pavements over the orchid rich grasslands was not expected.

An insight into this discrepancy was obtained through a common statement that was made by respondents when providing us with the rankings.

“I would be more concerned about the grasslands as they contain the unique flowers which need to be protected. So personally, I would give more priority to the management of the grasslands. However, most people would care for the Karst limestone as that is what they see and know about Burren. They wouldn’t understand the value of the unique flowers.”

Consistent comments of this nature from several respondents indicate that the disparity in preferences between the attributes is due to additional normative motives present in the valuation of the orchid rich grasslands. Just as people overestimate their generosity with higher willingness to pay values, they also seem to believe that they are less superficial than others.

VI. DISCUSSION AND CONCLUSIONS

In this paper we compare the conventional stated preference technique with a prediction based approach to highlight some of the advantages of the latter for use in environmental valuation. Through the results of this study we wish to contribute to the on-going literature on non-market valuation by providing further evidence of the potential of the prediction based technique proposed by Lusk and Norwood (2009a). As claimed by various theories in the psychology literature, by switching a direct questioning approach to one that asks for predictions about others, we are able to successfully eliminate social desirability bias which may be at the heart of the hypothetical bias problem.

We observe the mean willingness to pay values for the environmental attributes to be 2.5 and 3.1 times the predicted values. These numbers are close to the average range between hypothetical and non-hypothetical values as observed in the Meta-analysis by Murphy, et al. (2005). Thus, it is likely that our predicted values quite effectively reflect non-hypothetical values. However, because no non-hypothetical treatments were included in our study, we are not able to make a direct comparison.

\[ \chi^2 \text{ value of } 78.20 (p\text{-value } = 0.00), \text{ revealed this difference to be statistically significant at the one percent level of significance.} \]
Although non-hypothetical values have been viewed as unbiased estimates of true WTP, Lusk and Norwood (2009a) indicate that actual consumption values are lower than non-hypothetical values as they too are susceptible to social desirability bias. List, et al. (2004) observed that varying degrees of anonymity was capable of influencing both hypothetical and non-hypothetical outcomes. As such, a better chance of obtaining these unbiased values may be through prediction based questions.14

Similar to the WTP estimates, the results of the attitudinal questions are also observed to be consistent also with the Lake Wabegon effect and social desirability bias where individuals are cynical about others’ generosity and moral goodness in comparison to their own. We found that people generally assume money to be less of a concern for them but more of a priority for others. On the other hand they consider themselves to care more about the environment than others. However, in spite of the underestimated predictions, we observe respondents to identify themselves with others.

What was not expected initially was the inconsistency observed between the personal and predicted preferences for the two environmental attributes. However, the remarks made by some of the respondents enabled us to see that the difference was in fact characterized by additional normative motives for stating a higher preference for the orchid rich grasslands. As such, the interpretation of results of direct questions may be misleading if the researcher is unaware of the normative reasons people may have for making a behavioral choice.

Lusk and Norwood (2009b) advise careful interpretation of survey results when dealing with topics that contain strong social norms. The normative motives associated with such issues are more apparent than others and so the researcher can expect the responses to have a strong social desirability bias. From our study we see that the interpretation of survey responses is even more challenging when the normative motives are subtle. With certain issues one may not be able to anticipate the level of vulnerability to social desirability bias unless responses from both direct and indirect questions are compared. Although psychometric scale measures such as the Crowne-Marlowe (CM) have been employed extensively, they are not able to indicate the inclination of respondents to distort actual values (Johnson, et al. 2002).

As such, the employment of both direct and indirect questioning approaches may be beneficial as it not only provides responses that are free of social desirability bias but can also provide information regarding its level of vulnerability to the bias. Additionally, the value estimates from the two separate questioning formats can be compared to get an idea of their relative magnitudes. Although no formal scales to compare the accuracy of the predicted values exist, the ratio of the

14 Although Murphy, et al. (2005) observe the average calibration factor to be 2.6, they also acknowledge severe positive skewness with their data which cause the calibration factor evaluated at the median value to drop down to 1.35. By assuming the true calibration factor to lie somewhere in between, we can expect our predicted estimates to be lower than what would have been observed in a similar but non-hypothetical situation.
two values could be compared with calibration factors such as those developed by Murphy, et al. (2005) as crude measures. We believe this combined practice will be able to provide additional information for making more accurate estimates.

Theoretically the indirect questioning approach is effective in eliminating social desirability bias. The implementation is very simple and direct; and the few studies employing it have demonstrated its strong potential for environmental valuation. As such we highly recommend more studies incorporating and experimenting with this technique.

---

15 Compared to the different methods available for mitigating social desirability bias, Lusk and Norwood (2009b) consider the indirect questioning approach to be much simpler to employ and to analyze and interpret.
References:


http://www.burrenlife.com/
### Table 1: Attributes and Attribute levels of the Choice Sets

<table>
<thead>
<tr>
<th>Label</th>
<th>Attribute</th>
<th>Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landscape</td>
<td>Karst limestone pavements</td>
<td>‘With Management’ ‘No Management’</td>
</tr>
<tr>
<td>Biodiversity</td>
<td>Orchid rich grasslands</td>
<td>‘With Management’ ‘No Management’</td>
</tr>
<tr>
<td>Expected Annual Cost</td>
<td>Expected annual cost of implementing the chosen alternative</td>
<td>€5, €10, €20, €40</td>
</tr>
<tr>
<td>Statements (Self/Prediction)</td>
<td>Mean/ (Std Dev)</td>
<td>$\chi^2$ / (p-value)</td>
</tr>
<tr>
<td>---------------------------------------------------------------------------------------------</td>
<td>-----------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Strongly Disagree</td>
<td>Strongly Agree</td>
<td></td>
</tr>
<tr>
<td>1 The options are too expensive</td>
<td>3.29 (1.28)</td>
<td>167.39 (0.00)</td>
</tr>
<tr>
<td>2 The options are too expensive for most people</td>
<td>3.86 (1.17)</td>
<td></td>
</tr>
<tr>
<td>1 I object to paying higher taxes</td>
<td>3.89 (1.18)</td>
<td>117.27 (0.00)</td>
</tr>
<tr>
<td>2 Most people object to paying higher taxes</td>
<td>4.68 (0.65)</td>
<td></td>
</tr>
<tr>
<td>3 The Burren landscape is an important Irish phenomenon that needs to be protected</td>
<td>4.75 (0.56)</td>
<td>57.94 (0.00)</td>
</tr>
<tr>
<td>3 Most people would consider the Burren landscape to be an important Irish phenomenon that needs to be protected</td>
<td>4.27 (0.85)</td>
<td></td>
</tr>
<tr>
<td>4 I’d like to improve the rural environment and protect wildlife</td>
<td>4.70 (0.54)</td>
<td>51.82 (0.00)</td>
</tr>
<tr>
<td>4 Most people would like to improve the rural environment and protect wildlife</td>
<td>4.23 (0.80)</td>
<td></td>
</tr>
<tr>
<td>5 I’d like to improve the appearance of the countryside</td>
<td>4.67 (0.62)</td>
<td>33.35 (0.00)</td>
</tr>
<tr>
<td>5 Most people would like to improve the appearance of the countryside</td>
<td>4.24 (0.79)</td>
<td></td>
</tr>
</tbody>
</table>
Table 3: Results of the Mixed Logit Model.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean Estimates</th>
<th>Standard Deviation Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock_Self</td>
<td>5.61 (0.564)***</td>
<td>5.61 (0.564)***</td>
</tr>
<tr>
<td>Grass_Self</td>
<td>5.13 (0.514)***</td>
<td>5.13 (0.514)***</td>
</tr>
<tr>
<td>Cost_Self</td>
<td>-0.15 (0.017)***</td>
<td>-0.15 (0.017)***</td>
</tr>
<tr>
<td>Rock_Pred</td>
<td>3.05 (0.326)***</td>
<td>3.05 (0.326)***</td>
</tr>
<tr>
<td>Grass_Pred</td>
<td>2.21 (0.264)***</td>
<td>2.21 (0.264)***</td>
</tr>
<tr>
<td>Cost_Pred</td>
<td>-0.15 (0.023)***</td>
<td>-0.15 (0.023)***</td>
</tr>
<tr>
<td>Error_Self</td>
<td>6.24 (1.286)***</td>
<td></td>
</tr>
<tr>
<td>Error_Pred</td>
<td>8.14 (1.165)***</td>
<td></td>
</tr>
</tbody>
</table>

Number of Observations: 2336
Number of individuals: 292
Pseudo R²: 0.362
Mean WTP Rock - Self: €53.66
Mean WTP Grass – Self: €49.39
Mean WTP Rock – Prediction: €21.53
Mean WTP Grass – Prediction: €15.89
Log Likelihood function: -1490.145
χ²: 5390.81
Degrees of Freedom: 18

*** Significant at less than 1% level of significance

Table 4: Ranking between Attributes

<table>
<thead>
<tr>
<th>Attribute regarded as most important</th>
<th>ROCK</th>
<th>GRASS</th>
<th>COST</th>
<th>ROCK preferred over GRASS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self</td>
<td>43%</td>
<td>37%</td>
<td>20%</td>
<td>57%</td>
</tr>
<tr>
<td>Prediction</td>
<td>18%</td>
<td>11%</td>
<td>71%</td>
<td>70%</td>
</tr>
<tr>
<td>Total Respondents</td>
<td>289</td>
<td>289</td>
<td>289</td>
<td>289</td>
</tr>
</tbody>
</table>
Description of the Burren provided to Respondents:

The Burren

The Burren is located in the West of Ireland in county Clare.

The most distinctive feature of the Burren would be the rocky landscape with its exposed limestone pavements.

Despite the barren appearance, the limestone pavement is rich in flora and fauna. The lime-rich rock supports a diverse range of grassland and woodland plants.

With over 700 different types of flowering plants and ferns, this region hosts over 70% of Ireland’s native floral species.
Description of the karst limestone pavements with and without management:

With Management:

The picture depicts a section of the Burren landscape with proper management. The removal of shrubs (mainly hazel and blackthorn) improves the visibility of the limestone features and the controlled grazing of livestock prevents excessive growth of the shrub.

No Management:

The picture depicts a section of the Burren landscape that lacks any farm management practice. The spread of shrub (mainly hazel and blackthorn) reduces the visibility of the limestone features and access to them. The continued absence of management leads to shrub encroachment at the rate of 1% every year.

No Management:

Eventually over a number of years a few dominant shrub species will completely take over the landscape resulting in very poor visibility of the limestone pavements.
Description of the Orchid Rich Grasslands with and without management:

No Management:

The picture shows a section of the orchid-rich grassland that results when no management practice is in place. The lack of livestock grazing on the fields leads to undergrazing, which results in fewer species of plants. Shrub taking over at the expense of orchids and herbs brings about less flowering plants which in turn reduces the number of insect species such as butterflies and bees. There is an overall decline in the

With Management:

The picture depicts a section of the orchid-rich grasslands unique to the Burren as a result of proper management practice. The grasslands are grazed by livestock in a controlled manner which helps maintain high levels of plant and insect diversity. The grasslands are of very high quality, harboring many types of rare species of plants (orchids and herbs) and insects (butterflies/moths and bumble bees).

No Management:

Eventually over a number of years a few dominant shrub species will completely take over the grasslands.
Sample Choice Task:

<table>
<thead>
<tr>
<th>Landscape</th>
<th>Option A</th>
<th>Option B</th>
<th>No Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landscape</td>
<td><img src="image1" alt="Landscape with Management" /></td>
<td><img src="image2" alt="Landscape with Management" /></td>
<td><img src="image3" alt="Landscape No Management" /></td>
</tr>
<tr>
<td>Biodiversity</td>
<td><img src="image4" alt="Biodiversity No Management" /></td>
<td><img src="image5" alt="Biodiversity with Management" /></td>
<td><img src="image6" alt="Biodiversity No Management" /></td>
</tr>
<tr>
<td>Expected Annual Cost</td>
<td>€ 10</td>
<td>€ 20</td>
<td>€ 0</td>
</tr>
</tbody>
</table>