Available online at www.sciencedirect.com
SCIENCE DIRECT®
Ecological Economics 57 (2006) 507–519
ECOLOGICAL ECONOMICS
www.elsevier.com/locate/ecolecon

Analysis

A comparison of alternative certainty calibration techniques in contingent valuation

Mihail Samnaliev [a,*], Thomas H. Stevens [b], Thomas More [c]

[a] Center for Health Policy and Research, University of Massachusetts Medical School, Chang Building, 222 Maple Ave., Shrewsbury, MA 01545, United States
[b] Department of Resource Economics, University of Massachusetts, 80 Campus Center Way, Amherst, MA 01003, United States
[c] USDA, Forest Service, Northeastern Research Station, P.O. Box 968, Burlington, VT 05402, United States

Received 25 August 2003; received in revised form 2 May 2005; accepted 2 May 2005
Available online 12 September 2005

Abstract

A field test of two types of certainty calibration techniques in contingent valuation of public lands indicated that a 10-point certainty scale reduced WTP estimates by about half. Adjusting for uncertainty via a 'Not Sure' option did not reduce WTP estimates but the variance increased. There are several differences between these two ways of accounting for respondents' uncertainty, which may suggest why they provide different WTP value estimates and variances.

Keywords: Certainty calibration; Contingent valuation; CVM; WTP

1. Introduction

Empirical evidence suggests that significant uncertainty often exists in responses to contingent valuation questions (Alberini et al., 2003). Since respondent uncertainty has often been related to the problem of hypothetical bias (see Harrison and Rustrom, in press; List and Gallet, 2001), several contingent valuation, CVM, formats that allow respondents to express uncertainty directly have been developed. Examples include the multiple-bounded question format (Welsh and Poe, 1998), a "random-valuation" model (Wang, 1997), various uncertainty scales (Champ et al., 1997; Ekstrand and Loomis, 1997) a polychotomous choice format (Ready and Navrud, 1999), and NOAA's well-known 'Don't Know' or 'Not Sure' option. However, agreement about the appropriate method for uncertainty adjustment is far from universal. For example, Wang (1997), Carson et al. (1994), and Alberini et al. (2003) present very different views about calibration for uncertainty.

A 10-point certainty scale following a dichotomous choice, DC, format and the inclusion of a 'Not Sure'

* Corresponding author.
E-mail addresses: Mihail.Samnaliev@umassmed.edu (M. Samnaliev), tstevens@resecon.umass.edu (T.H. Stevens).

0921-8009/$ - see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.ecolecon.2005.05.017

option are two common ways to incorporate uncertainty. Use of a certainty scale with a cut-off point of 8 and 10 (with 10 being very certain) has been shown to provide similar hypothetical and actual willingness to pay, WTP, estimates (Champ et al., 1997). The treatment of 'Not Sure' responses has been more controversial (Wang, 1997), but a common approach has been to treat them as either 'No' or missing (Alberini et al., 2003; Carson et al., 1994).

This study compares the effect of these two types of certainty adjustment on WTP estimates in a randomized split sample mail survey. We find that treatment of 'Yes' responses with certainty of less than 8 (or 10) as 'No' provide different willingness to pay estimates than treatment of 'Not Sure' responses as either 'No' or as missing. We then contribute to the discussion on the motivation underlying uncertain responses and argue that the two calibration methods may be conceptually different.

2. Previous studies

The motivation behind uncertain responses is not well understood. After the NOAA panel suggested that a 'Don't Know' option should be added to the DC CVM format, a body of literature has explored respondent motivation underlying 'Not Sure' responses. Alberini et al. (2003) suggest three interpretations of responses to this option. One possibility is that 'Don't Know' respondents are not in the market for the good being valued. A second interpretation is that 'Don't Know' respondents have not yet made up their mind. The third possibility is that these responses reflect uncertainty. Moreover, Alberini et al. define two types of uncertainty: (a) "true" uncertainty wherein respondents have insufficient experience and (b) "false" uncertainty wherein respondents do not want to spend time thinking about the valuation question or would like to indicate some support for the item being valued, but would not pay the amount asked. Carson et al. (1994) recommend that 'Not Sure' responses be treated as missing, because respondents who choose the 'Not Sure' option would say 'No' if actually forced to choose. In addition, Champ et al. (2003) find that respondents may choose the 'Not Sure' option because they are uncertain about their income, ability to commit to spending money, or about the benefits of the program. Other hypotheses include the notion that uncertainty may arise because of lack of knowledge, interest, or inability to make a quick decision.

Wang (1997) presented an alternative interpretation of 'Don't Know' responses. He argued that 'Don't Know' (or 'Not Sure') answers represent the point of indifference to the offered bid. As the price of the commodity increases, a typical respondent would switch her answer from 'Yes' to 'Don't Know' and from 'Don't Know' to 'No'. Wang included the 'Don't Know' answers in a multinomial probit model estimation and concluded that they provide useful information about preferences.

On the other hand, certainty scale calibration has become quite popular in dichotomous choice (DC) CV studies. In this approach, people are asked how certain they are of their response on a 10-point scale. A common application of the certainty scale is to treat positive answers as 'Yes' only when certainty levels are at least 8 on a 10-point scale with 10 indicating 'Very Certain' (for example, see Champ et al., 1997). The effectiveness of this method has been established by comparing hypothetical payments to actual donations (Champ et al., 1997; Polasky et al., 1996). These, as well as other recent studies, suggest that uncertainty scale calibration can reduce hypothetical bias and/or so called 'Yea-Saying' effects. However, Ekstrand and Loomis (1997) reported that the effect of this method depends on how the scale is used. Bias reduction was reported when certainty levels of at least 8 were used to calibrate only 'Yes' answers, but reduction of bias was questionable when 'No' answers were also calibrated. In addition, the authors found that certainty calibration reduced the goodness of fit (of the logit WTP model) and increased the variance in responses.[1]

Taken together, these arguments demonstrate the complexity of the issue of uncertainty calibration. Uncertainty is not a precise or single condition and may be caused by a range of factors. Further, little is known about the separate or confounding effect of each factor and this presents a methodological prob-

[1] However, Welsh and Bishop (1993) reported that certainty calibration reduced the variance in responses. Several other studies have also applied certainty scales to calibrate 'Yes' and 'No' responses (Li and Mattsson, 1995).