University of Caen

University of Rennes 1

CREM UMR CNRS 6211

Centre de Recherche en Économie et Management
Center for Research in Economics and Management

Working Paper

Dealing with Internal Inconsistency in Double-Bounded Dichotomous Choice: An Application to Community-Based Health Insurance

Hermann Pythagore Pierre Donfouet
CREM UMR CNRS 6211, University of Rennes 1, France

Pierre-Alexandre Mahieu
LEMNA, University of Nantes, France

P. Wilner Jeanty
Rice University, Houston, Texas, US

December 2011 - WP 2011-30

université de Caen Basse-Normandie

Université de Rennes 1

Dealing with Internal Inconsistency in Double-Bounded Dichotomous Choice: An Application to Community-Based Health Insurance

Hermann Pythagore Pierre Donfouet

CREM, UMR 6211

University of Rennes I, France.

Corresponding author's email: donfouetz@yahoo.fr

Tel : +33 (0) 2 2323 3364

Fax: +(33) (0)2.23.23.35.09
7 Place Hoche, 35 065 Rennes Cedex.

Pierre-Alexandre Mahieu

LEMNA

University of Nantes

Chemin de la Censive du Tertre, 44 322 Nantes Cedex, France

P. Wilner Jeanty

Kinder Institute for Urban Research

Rice University

MS 208 6100 Main Street

Houston, TX 77005, US

Acknowledgements

A grant from the International Labor Organization (40052113/0) is gratefully acknowledged and the scientific support of the European Development Research Network (EUDN).We will also like to thank the African Doctoral Dissertation Research Fellowship offered by the African Population and Health Research Center (APHRC) in partnership with the International Development Research Centre (IDRC) for financial assistance. Any errors or omissions remain our sole responsibility.

No conflicts of interest.

Abstract

Contingent valuation method is commonly used in the field of health economics in an attempt to help policy maker in taking decisions. The use of the double-bounded dichotomous choice format results in a substantial gain in statistical efficiency over the single bounded dichotomous choice format. Yet, this efficiency gain comes at the cost of biasness known as internal inconsistency. This paper aims at reducing this internal inconsistency in double-bounded dichotomous choice by using the certainty calibration technique in a community-based health insurance study. Findings confirm the internal inconsistency between the initial and the follow-up responses and the statistical efficiency gains of the double-bounded dichotomous choice over the single-bounded dichotomous choice. Furthermore, the use of certainty calibration reduces this internal inconsistent pattern in responses and still maintains efficiency gain. We further discuss the policy implications.

Keywords: Contingent valuation; internal inconsistency; certainty calibration; community-based health insurance.

JEL Classification: C15; D6; I38.

1. Introduction

Contingent valuation (CV) method is used to assess the preferences of respondents for a specific good. The value obtained from CV survey is important to policy-makers since they may use this value in their decisions, such as when their decision to undertake a project depends on the results of a cost-benefit analysis. Thus, researchers must estimate without any bias the value that the respondents attach to a particular good or policy. In an attempt to reach this goal, the single bounded dichotomous choice and double-bounded dichotomous choice formats have been used over the past years.

The use of the open-ended format of the form "how much are you willing to pay for X (or for policy A)?" has been discarded in favor of the single bounded dichotomous choice (SBDC) of the form "are you willing to pay X dollars (or for policy A)?" Since the latter is incentive compatible and mimics behavior in regular markets where people usually purchase or decline purchase of a good at the posted price (Arrow et al. 1993; Bishop and Heberlein 1979). However, the SBDC provides less information about each respondent's willingness-to-pay (WTP) resulting in decreased efficiency in the estimates of WTP. In order to get more information about the WTP of respondents, Carson et al. (1986) developed the double-bounded dichotomous choice format (DBDC) which consists of asking another yes/no response to the individual, where a higher or a lower amount is presented to the individual depending on his first response. A few years later, Haneman et al. (1991) demonstrated that DBDC is more efficient than the SBDC. Further, empirical applications showed that WTP amounts from the first and second responses were not driven by the same underlying preferences, with the former being significantly lower than the latter (McFadden 1994; Cameron and Quiggin 1994; Kanninen 1995; Herriges and Shogren 1996; Bateman et al. 2001; Burton et al. 2003; Bateman et al. 2008). This is known as internal inconsistency emerging from the second responses. In fact, McFFaden (1994, pp705-706) stated that the double referendum elicitation format is internally inconsistent, causing some practitioners to abandon such elicitation format.

Another important issue in CV is the so-called hypothetical bias, a tendency of respondents to state WTP amounts different from what they would pay in real settings. Different approaches have been proposed in the literature to address this issue. Among such approaches, the calibration approach proposed by Champ et al. (1997) has given good results. Respondents stating “yes” to the SDBC are asked how much they are sure about their answer on a numerical scale. “Yes” responses are recoded as “no” when the rating score is inferior to a given threshold. To the best of our knowledge, the certainty calibration approach has never been used in DBDC, due probably to the fear in internal inconsistency. This paper aims at testing whether the double calibration reduces the internal consistency based on a study dealing with community-based health insurance (CBHI).

The rest of the paper is organized as follows: Section 2 provides some background on internal inconsistency and hypothetical bias. Section 3 presents the methodology used, while section 4 describes the survey design and data, section 5 considers the empirical results of the study. Finally, section 6 discusses the results and section 7 concludes with some policy implications.

2. Background

2. 1 Internal inconsistency

Several reasons have been proposed in the literature to account for internal inconsistency in the double-bounded approach. The nature of the survey may account, give rise to this phenomenon. Since the results of the survey are not consequential, people may invest little effort in the valuation task and have a range of value in mind rather than a single point. Flachaire and Hollard (2007) showed that the existence of a range may be the culprit for the internal inconsistency. In their model which is based on the coherent arbitrariness principle, people are uncertain about their true WTP and are prone to anchoring effect.

The government wastage model was proposed by Carson et al. (1994) in order to explain the downward mean shifting in the second responses. In this model, the respondents saying “yes” to the initial bids for the provision of a public good might conceive of the higher follow-up bids as an attempt by the government to collect more funds than needed to cover the provision of the good and will say “no” to the follow-up bid since perceiving it as a waste. By the same token, respondents saying “no” to the initial bids might view the lower follow-up bids as an indication that the good being valued is of lower quality, and thus they will answer “no” to the follow-up bids. This model implies that the respondents will vote against the second follow-up bids regardless of whether or not they have accepted or rejected the initial bids. As Haab and McConnell (2002) stated, the aggregate proportion of “yes” to a given bid is lower and the DBDC will yield a smaller mean WTP. Another possible explanation is the strategic behavior model (Mitchell and Carson 1989) where the respondents answer the first questions truthful but answer the second ones strategically. They tend to lower the bids by rejecting any additional bids proposed by the researcher. To avoid this strategic behavior while gaining efficiency, Cooper et al. (2002) have proposed the one and one-half bound approach. Bateman et al. (2008) showed that the respondents were unfamiliar with the institutional procedures of the DBDC and they were surprised by follow-up questions.

2.2 Certainty calibration

Several techniques have been developed in an attempt to mitigate hypothetical bias. The most popular are certainly the cheap talk approach (Cummings and Taylor 1999) and the certainty calibration approach (Champ et al. 1997). In cheap talk, a script is placed just before the valuation question to inform the participants about the hypothetical bias and to remind them their budget constraint. In certainty calibration, “yes” responses are recoded as “no” responses when the individual is not sure enough about his response. Several studies have been conducted to compare the effectiveness of cheap talk and calibration at mitigating hypothetical bias. It turns out that the calibration technique is more effective than cheap talk (Blumenschein et al. 2008; Champ et al. 2009; Samnaliev et al. 2003; Loureiro et al. 2009). Furthermore, Mahieu and Rulleau (2011) argue that cheap talk may encourage the respondent to lower his WTP just to meet the expectation of the interviewer.

2.3 Double certainty calibration

In the current study, the follow-up certainty questions (FCQ) is used to calibrate the respondent’s WTP for the first and the follow-up bids. “Yes” respondents are asked how sure they are about their answers on a scale ranging from 1 to 10, where 1 means “very uncertain” and 10 “very certain”. Following Ethier et al. (2000), Poe et al. (2002), a threshold of 7 out of 10 is set. Then, all “yes” answers are not recoded if the answer to the FCQ is “equaled or greater than 7”; otherwise recoded as a “no”. This rule is also applied to the follow-up bid of the second question. The crux of the double calibration is that mitigating hypothetical bias may reduce the internal inconsistency in response patterns.

3. Econometric methods

3.1 Single Bounded Dichotomous Choice

Following Hanemann (1984), we assume that $v\ \ p, q, y, s, \varepsilon$ is the indirect utility function of the individual, where p represents the prices of the market goods, q the non-market good, y

the respondent's income, s sociodemographic characteristics such as age, income, gender, and ε the stochastic component of preferences. Via the questionnaire, the respondent is confronted with the possibility of a change from initial situation to the proposed alternative (that is from q^0 *to* $q^1 \succ q^0$). In the survey, the researcher will inform the respondent that this change will cost him a certain amount A and he is then asked whether he would be in favor of it at that price. The respondent will answer a "yes" if only $v(p,q^1,y-A,s,\varepsilon) \geq v(p,q^0,y,s,\varepsilon)$ and "no" otherwise. Hence,

$$\Pr(response\ is\ "yes") = \Pr\{v(p,q^1,y-A,s,\varepsilon) \geq v(p,q^0,y,s,\varepsilon)\} \qquad (1)$$

By using the compensating variation measure, the quantity C satisfies:

$$v(p,q^1,y-C,s,\varepsilon) = v(p,q^0,y,s,\varepsilon),$$

Thus, $C = C(p,q^0,q^1,y,s,\varepsilon)$ is his maximum WTP for the change from q^0 *to* q^1. It follows that he answers "yes" if the stated price is *less* than his WTP and "no" otherwise.

Hence, an equivalent condition to (1) is:

$$\Pr(response\ is\ "yes") = \Pr\{C(p,q^0,q^1,y,s,\varepsilon) \geq A\}, \qquad (2)$$

In other words, the respondent will say "yes" when his maximum willingness to pay for the change from q^0 *to* q^1 is larger than the proposed bid *A*. For instance, when the respondent is asked whether he would pay *A* monetary units for a health policy which aims at improving his health status from q^0 *to* q^1, he will answer with a "yes" if his willingness to pay is larger to *A*.

Besides, it is assumed that $C(p,q^0,q^1,y,s,\varepsilon)$ is a random variable, while the respondent's WTP for the change in q is something that he himself knows, it is something that the researcher does not know but treats as a random variable. Let $G_c(\bullet)$ be what the investigator assumes is the cumulative distribution function (CDF) of C, and $g_c(\bullet)$ the corresponding density function. Then (2) becomes:

$$\Pr(response\ is\ "yes") = 1 - G_c(A), \tag{3}$$

The form of the function $G_c(A)$ determines the econometric model to be used. If the $G_c(A)$ follows a probit standard distribution and the model to estimate is linear, then the expected mean WTP is:

$$\mu_{SBDC} = -\frac{\alpha}{\beta}, \tag{4}$$

Where α is the intercept and β the estimated marginal utility of income.

The standard errors for μ_{SBDC} is obtained from the variance of $Var(\mu_{SBDC}) = Var\left(\frac{\alpha}{\beta}\right)$, which is calculated by the Delta method (Taylor series expansion). However, because confidence intervals obtained from the Delta method are symmetric around the mean, hence not appropriate (Park et al. 1991) , the 95% confidence intervals for mean WTP estimates are constructed using the Krinsky and Robb (1986)'s Monte Carlo simulation and implemented in Stata using the *wtpcikr* command (Jeanty 2007).

3.2 Double Bounded Dichotomous Choice

The DBDC has the advantage of higher statistical efficiency in welfare estimates over the SBDC. In the DBDC, two sequences of bids are offered to the respondents. First, a respondent is asked whether he would be willing to accept or reject an initial bid; thereafter a second bid is offered; depending on the respondent's answer to the first bid, the second bid could be iterated downwards or upwards. In other words, a respondent is asked if he will be willing to pay an initial bid $A for perceived improved access to health care services via the introduction of a newly proposed community-based health insurance (CBHI). If he accepts the initial bid, a second higher bid A_n^h (the double of the first bid) will be offered. If he rejects, a second lower bid A_n^l (half of the first bid) will be offered. Therefore there are four possible responses: “yes-yes”; “yes-no”; “no-yes” and “no-no”.

The econometric procedure used follows Hanemann et al. (1991). This model assumes that the WTP distribution from the first answer and the WTP from the second one are identical $\mu = \mu_1 = \mu_2$. The interval data probit model is estimated. In this model, the mean/median WTP estimates and the dispersion parameters are assumed to be the same across equations or questions.

Let A^1 denote the first bid and A^2 the second bid. The bounds on the WTP are:

$$
\begin{aligned}
&A^1 \leq WTP \prec A^2 \;\; for\ the\ "yes-no"\ responses;\\
&A^1 \succ WTP \geq A^2 \;\; for\ the\ "no-yes"\ responses;\\
&WTP \geq A^2 \;\; for\ the\ "yes-yes"\ responses;\\
&WTP \prec A^2 \;\; for\ the\ "no-no"\ responses;
\end{aligned}
$$

The general form the double-bounded model is:

$$WTP_{ij} = \mu_i + \varepsilon_{ij} \tag{5}$$

WTP_{ij} represents the j^{th} the respondent’s willingness to pay , and *i= 1, 2* represents the first and the second answers. μ_1, μ_2 are the means for the first and the second responses. To construct the likelihood function, Hanemann et al. (1991) assumed that $\mu = \mu_1 = \mu_2$, where μ is a parameter. Furthermore, they assumed that the model in all its parts is the same for each question that is for the j^{th} individual:

$$WTP_j = \mu + \varepsilon_j, \tag{6}$$

written with the error as normal, the j^{th} contribution to the likelihood function is:

$$L_j\left(\mu|\mathrm{A}\right)=\Pr\left(A^2-\mu\succ\varepsilon_j\succ A^1-\mu\right)^{YN}.\Pr\left(\mu+\varepsilon_j\succ A^2\right)^{YY}\times\Pr\left(\mu+\varepsilon_j\prec A^2\right)^{NN}.\Pr\left(A^1-\mu\succ\varepsilon_j\succ A^2-\mu\right)^{NY},$$

where:

$$\begin{aligned}
&YY=1\ \mathit{for\ a}\ "yes-yes"\ \mathit{answer},\ 0\ \mathit{otherwise};\\
&NY=1\ \mathit{for\ a}\ "no-yes"\ \mathit{anwser},\ 0\ \mathit{otherwise}\\
&YN=1\ \mathit{for\ a}\ "yes-no"\ \mathit{anwser},\ 0\ \mathit{otherwise}\\
&NN=1\ \mathit{for\ a}\ "no-no"\ \mathit{anwser},\ 0\ \mathit{otherwise}
\end{aligned}$$

The standard error of the mean WTP for DBDC is also calculated using the Delta method and confidence interval using the Krinsky and Robb's (1986) Monte Carlo simulation.

When comparing the mean WTP of SBDC $\left(\mu_{SBDC}\right)$ and mean WTP of the DBDC $\left(\mu_{DBDC}\right)$, it is expected in this current study that the mean WTP of DBDC $\left(\mu_{DBDC}\right)$ will be lesser than the mean WTP of the SBDC $\left(\mu_{SBDC}\right)$. According to the internal inconsistency, one expects a downward shift from the μ_{SBDC} to μ_{DBDC}. The one-tailed test of difference in mean WTP of the SBDC $\left(\mu_{SBDC}\right)$ and DBDC $\left(\mu_{DBDC}\right)$ can be undertaken. The test is constructed as follows:

$$\begin{aligned}
&H_0:\quad \Delta=\mu_{SBDC}-\mu_{DBDC}>0\\
&H_1:\quad \Delta=\mu_{SBDC}-\mu_{DBDC}\leq 0
\end{aligned}\qquad (7)$$

This test is a bit complex given that there is correlation between the first answer and the second answer which yields to non-independence of the values obtained for the two elicitations questions. Hence, the covariance between the responses from the first initial questions and follow-up questions are different to zero. In other words, it is not possible to use paired t-test.

Bootstrap technique is an effective way to undertake this test (Efron and Tibshirani 1993). With *wtpcikr* command in Stata software, developed by Jeanty (2007) , we have saved the replications data for both the mean WTP of SBDC and mean WTP of DBDC in a dataset. Then, we have loaded the dataset for the SBDC, merged it with the dataset for the DBDC, and calculated the difference between mean WTP of SBDC and mean WTP of DBDC. Once we have calculated the difference, the achieved significance level (ASL) is then calculated. The greater the ASL (greater than 5%), the more likely the internal inconsistent patterns in responses. The same procedure was employed for both responses with and without calibration.

Given that we expect a downward mean shifting in the second responses ($\mu_{SBDC} > \mu_{DBDC}$), the use of certainty calibration could be used to reduce the discrepancy between μ_{SBDC} and μ_{DBDC} , thus producing internal coherent patterns in responses.

4. Survey design and data

The good being valued in the study is the provision of CBHI to the rural households in Bandjoun, a province located in West of Cameroon. Given that most rural households are excluded from formal insurance, CBHI has emerged as a concept and strategy to reach the poor in rural areas with adequate health care service. CBHIs are small scale, voluntary health insurance programs, organized and managed in a participatory manner (Tabor 2005). CBHI is now adopted in many developing countries (see for instance Dong et al. 2003; Dror et al. 2007; Ataguba et al. 2008; Asenso-Okyere et al. 1997). Recently policymakers in Cameroon have adopted a health strategic plan for the promotion of CBHI. It aims at: (a) putting in place

CBHI per health district by 2015 and (b) covering at least 40% of the population by the CBHI by 2015. A face-to-face interview was conducted in six villages on a sample of 369 rural households heads selected by a two-stage cluster sampling technique. In an attempt to conduct a state-of-the art contingent valuation, guidance provided in Arrow et al. (1993), Carson (2000), Carson et al. (2000), and Whittington (2002) were followed. The scenario explained to the respondents the concept of CBHI, the operation of CBHI, the benefits associated to CBHI, and the premium that they have to pay to receive such benefit. Focus groups and pre-test conducted helped to determine these initial bids: 250, 350, 450, 550, 650 and 800 CFA francs. The follow-up bids were the double of the initial bids A_n^h if the respondent answered "yes" to the first valuation question and half of the initial bids A_n^l if he has answered "no". Furthermore, the follow-up certainty questions (FCQ) were included after the initial bids and the follow-up bids as well. The FCQ asked the respondents to rate on a 10-point numerical likert scale ranging from 1 "very uncertain" to 10 "very certain", how sure they felt that they would actually pay for the CBHI if they answer "yes" to the valuation question. This self-reported certainty level is used to re-code responses to the WTP question and to provide an estimate of mean WTP similar to the actual WTP. Parallel to Ethier et al. (2000), Poe et al. (2002), a threshold 7 out of 10 is set. Then, all "yes" answers are recoded as "no" if the score is strictly inferior to 7.

Figure 1 clearly indicates that the percentage of "yes" responses based on the first bids is downward sloping. This is without calibration. This suggests a downward sloping Hicksian demand function. This figure demonstrates that the responses of households are in conformity with economic theory. In fact, as the premiums increase, the households are less willing to pay for CBHI. This shows that the insurance is a normal good.

[Insert Figure 1 here]

5. Results

Table 1 and table 2 provide mean WTP of the SBDC and DBDC. As can be seen, there is downward mean shifting in the second responses ($\mu_{SBDC} > \mu_{DBDC}$). Indeed, there is clear pattern of internal inconsistency in responses of the respondents. This finding is in conformity with previous researches (Hanemann et al. 1991; Cameron and Quiggin 1994; Herriges and Shogren 1996; DeShazo 2002; Bateman et al. 2008).

Following Loomis and Ekstrand (1998), we compare the efficiency gain of the DBDC over the SDBC. The ratio of the confidence interval to the mean WTP is used as a relative measure of efficiency of WTP estimates (CI/mean = (Upper bound – lower bound)/meanWTP). The lower the ratio, the higher the efficiency. A close look to the estimates in table 1 and table 2 confirm that the ratio of the confidence interval to the mean WTP of DBDC is lower than that of the SDBC (0.15<0.68)[1]. Accordingly, the use of DBDC in the current study yields to more efficient WTP estimates than the SDBC. However, this efficiency gain comes at the cost of biasness since there is a downward mean shifting in WTP from the second responses ($\mu_{SBDC} > \mu_{DBDC}$).

[Insert Table 1 and Table 2 here]

To conduct the test of internal inconsistency, the parametric bootstrap technique was used. Table 3 shows that there is difference between the μ_{SBDC} and μ_{DBDC} (Δ=91.32). An interesting result provided by the p-value (0.83) confirms the non rejection of the null hypothesis of internal inconsistency, meaning that there is statistical evidence to support the internal inconsistency in response patterns hypothesis. Nevertheless, it is possible to use

[1] The DBDC yields four times efficiency gains as compared to the SDBC.

follow-up certainty questions (FCQ) to calibrate respondents' WTP and also reduce the discrepancy between the mean WTP calculated from the SBDC and DBDC and maintain the efficiency gain as well.

[Insert Table 3 here]

As can be seen in table 3, the use of double calibration reduces the discrepancy between the mean WTP from the SBDC and DBDC $\Delta = 36.75$. In fact, the use of calibration technique has reduced the internal inconsistency by 60 % (see table 3). The p-value (0.65) of the test of internal inconsistency constructed by the bootstrap technique indicates that there is failure to reject the null hypothesis of internal inconsistency. Thus, the null hypothesis of internal inconsistency cannot be rejected at 5% level. The failure to reject this hypothesis is due to a higher variance in the first bid calibrated. This result is similar to the findings of Bateman et al. (2008) though in different context.

In theory, calibrating the responses of respondents must not affect the efficiency gain of the DBDC over SBDC though the central tendency could be affected. As argued by Alberini et al., (2003), there is no reason to believe that allowing uncertain responses will affect the efficiency of welfare estimates. We further investigate the efficiency gain when calibration is applied. As can be seen in Tables 1 and 2, there is still gain efficiency of the DBDC over SBDC when the calibration technique is applied. For instance, in Tables 1 and 2, the ratio of the confidence interval to the mean WTP of DBDC is lesser than that of SBDC (0.41<1.33). In other words, confidence intervals around the mean WTP estimates of DBDC are still tighter than the one around the mean WTP estimates of SBDC. Thus, there is a consistency of efficiency gain in the study. Lastly, the mean WTP of SBDC and DBDC are both reduced when the calibration is applied. In fact, before the double calibration, the mean WTP of

SBDC and DBDC are respectively 1064.95 CFA francs and 973.63 CFA, while these means are 975.49 CFA francs, 938.74 CFA francs when the calibration is then applied. Accordingly, if policymakers are keen to know what the poor rural households are willing to pay for CBHI in Bandjoun, they may set the premium at 938.74 CFA francs/person/month (about 2 US dollars).

6. Discussion

As argued by Hanley et al. (2009), the respondents would prefer to state a range of values instead of a point estimate, because they are unsure about the value they place on the proposed goods or policy. This uncertainty could lead to an overestimation of the mean WTP and a behavioral inconsistency. Yet, the calibration technique could be used to mitigate this hypothetical bias and the anomalous behavior in response patterns.

Results of the empirical study suggest that the double calibration technique, which consists at recoding yes/no answers, reduces internal inconsistency. Thus, DBDC with calibration might not only be effective at mitigating hypothetical bias than SDBC, since the mean WTP is similar, but it provides higher statistical information. Accordingly, applying the calibration approach on DBDC may be preferred to applying calibration on SDBC. Consequently, from a methodological point of view, this study distinguishes from previous research by being the first to implement a calibration technique in DBDC in order to reduce internal inconsistency, hypothetical bias and maintain efficiency gains.

7. Conclusions

The use of contingent valuation (CV) method in the health sector is gaining popularity since policymakers may rely on the results of CV survey to improve the well-being of their populations. Over the past decades, there has been a shift from single bounded dichotomous choice (SBDC) to double-bounded dichotomous choice (DBDC) because of the statistical efficiency gains of the DBDC. Nevertheless, the use of DBDC has been criticized on the ground that responses from the initial bids are inconsistent with the responses to the second bid, with a downward mean shifting in the second responses. Empirical evidences from previous studies have confirmed this internal inconsistency in DBDC. This paper aims at using the certainty calibration to reduce this internal inconsistency in DBDC by focusing on the community-based health insurance (CBHI). To the best of our knowledge, this is the first paper which addresses the internal inconsistency in DBDC by using certainty calibration technique.

The results of the study confirm the internal inconsistency between the initial and the follow-up responses and the statistical efficiency gains. Indeed, there is substantial difference between the mean WTP of the SBDC and DBDC for CBHI, and a four times efficiency gains of the DBDC over the SDBC. The parametric bootstrap technique used confirms the statistical evidence of internal inconsistent in responses. Furthermore, the use of certainty calibration reduces this internal inconsistency patterns in responses by 60% and still maintains the efficiency gain of the DBDC over the SBDC. In other words, by calibrating the WTP of the respondents, the discrepancy between mean WTP of SBDC and DBDC is reduced and there is an efficiency gain in the use of DBDC over SBDC. This finding has two major implications. The first implication is the methodological one: the presence of inconsistency in DBDC can be mitigated by the certainty calibration technique and yield theoretically consistent preferences. The second implication is policy-relevant: policymakers may decide to set the premium of the rural poor household heads at 2 about US dollars.

References

Alberini A, Boyle K, Welsh M (2003) Analysis of Contingent Valuation Data with Multiple Bids and Response Options Allowing Respondents to Express Uncertainty. 45:40-62

Arrow K, Solow PR, Leamer EE, Radner R, Shuman H (1993) Report of NOAA Panel on Contingent Valuation Method. Fed Regist 58 (10):4601-4614

Asenso-Okyere WK, Osei-Akoto I, Anum A, Appiah EN (1997) Willingness to pay for health insurance in a developing economy. A pilot study of the informal sector of Ghana using contingent valuation. Health Pol 42 (3):223-237

Ataguba J, Ichoku EH, Fonta W Estimating the willingness to pay for community healthcare insurance in rural Nigeria. In, Dakar, 2008. Poverty and Economic Policy,

Bateman I, Burgess GD, Hutchinson W. G., Matthews DI (2008) Learning design contingent valuation (LDCV): NOAA guidelines, preference learning and coherent arbitrariness. 55:127-141

Bateman IJ, Langford IH, Jones AP, Geoffrey N. K. (2001) Bound and Path effects in double and triple bounded dichotomous choice contingent valuation. 23:191-213

Bishop RC, Heberlein T (1979) Measuring the values of extra-market goods: Are indirect measures biased? . 61:926-930

Blumenschein K, Blomquist GC, Johannesson M, Horn N, Freeman P, (2008) Eliciting willingness to pay without bias: Evidence from a field experiment. The Econ J 118 (525):114-137

Burton AC, Carson KS, Chilton SM, Hutchinsonet WG (2003) An experimental investigation of explanation for inconsistencies in responses to second offers in double referenda. 46:472-489

Cameron TA, Quiggin J (1994) Estimation using contingent valuation data from a “dichotomous choice with follow-up” questionnaire. J Envrion Manag 27:218-234

Carson RT (2000) Contingent valuation: A user’s guide. Environ Sci Technol 34:1413-1418

Carson RT, Hanemann M, Mitchell RC (1986) Determining the Demand for Public Goods by Simulating Referendums at Different Tax Prices. Department of Economics, University of California, San Diego (1986),

Carson RT, Wilks L, Imber D (1994) Valuing the Preservation of Australia's Kakadu Conservation Zone. 46:727-749.

Champ PA, Bishop RC, Brown TC, McCollum DW (1997) Using donation mechanisms to value nonuse benefits from public goods. J Environ Econ Manag 33 (2):151-162

Champ PA, Moore R, Bishop RC (2009) A Comparison of Approaches to Mitigate Hypothetical Bias. Agric Resour Econ Rev 38 (2):166–180

Cooper J, Hanemann MW, Signorelli G (2002) One and One-half Bids for Contingent Valuation. 84:742-750.

Cummings R, Taylor LO (1999) Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method. 89 (3):649-665.

DeShazo JR (2002) Designing Transactions Without Framing Effects in Iterative Question Formats. 43:360–385

Dong H, Kouyate B, Cairns J, Mugisha F, Sauerborn R (2003) Willingness-to-pay for community-based insurance in Burkina Faso. Health Econ 12 (10):849-862

Dror M, Radermacher R, Koren R (2007) Willingness to pay for health insurance among rural and poor persons: Field evidence from seven micro health insurance units in India. Health Pol 82:12-27

Efron, Tibshirani (1993) An Introduction to the Bootstrap. Chapman and Hall, New-York

Ethier RG, Poe GL, Schulze WD, Clark JE (2000) A comparison of hypothetical phone and mail contingent valuation responses for green pricing electricity programs. Land Econ 76 (1):54-67

Flachaire E, Hollard G (2007) Starting Point Bias and Respondent Uncertainty in Dichotomous Choice Contingent Valuation Surveys. 29 (3):183-194.

Haab TC, McConnell KE (2002) Valuing Environmental and Natural Resources, The Econometrics of Non Market Valuation. Cheltenham U.K., Edward Elgar

Hanemann M, Loomis WJ, Kanninen B (1991) Statistical efficiency of double-bounded dichotomous choice contingent valuation. Am J Agric Econ 73 1255-1263.

Hanley N, Kristrom B, Shogren JF (2009) Coherent Arbitrariness: On Value Uncertainty for Environmental Goods. 85 (1):41-50

Herriges JA, Shogren JF (1996) Starting Point Bias in dichotomous Choice Valuation with Follow-up Questioning. J Environ Econ Manag 30 (1):112-131

Jeanty PW (2007) Constructing Krinskyand Robb Confidence Intervals for Mean and Median Willingness to Pay (WTP) Using Stata. 6th North American Stata Users'Group Meeting, August13-14, 2007 Boston, MA.,

Kanninen BJ (1995) Bias in Discrete Response contingent valuation. 28 (1):144-125

Krinsky I, Robb AL (1986) On Approximating the Statistical Properties of Elasticities. 68:715-719

Loomis J, Ekstrand E (1998) Alternative Approaches for Incorporating Respondent Uncertainty When Estimating Willingness to Pay: The Case of the Mexican Spotted Owl. Ecol Econ 27:29–41.

Loureiro ML, Loomis JB, Vázquez MX (2009) Economic Valuation of Environmental Damages due to the Prestige Oil Spill in Spain. 44:537–553. doi:DOI 10.1007/s10640-009-9300-x

Mahieu PA, Rulleau B (2011) Does The Position of Income Questions influence Willingness-to-Pay in Contingent Valuation Surveys?

McFadden D (1994) Contingent valuation and social choice. 76:689-708

Mitchell RC, Carson RT (1989) Using surveys to value public goods: the contingent valuation method. Resource for the Future, Washington

Park T, Loomis JB, Creel M (1991) Confidence Interval for Evaluating Benefits Estimates from Dichotomous Choice Contingent Valuation Studies. 67:64-73

Poe GL, Clark JE, Rondeau D, Schulze W. D. (2002) Provision Point Mechanisms and Field Validity Tests of Contingent Valuation. Environ Resour Econ 23:105-131

Samnaliev M, Stevens T, More T (2003) A Comparison of Cheap Talk and Alternative Certainty Calibration Techniques in Contingent Valuation. University of Massachusetts Amherst, Department of Resource Economics, Working Paper No.11.

Tabor SR (2005) Community-Based Health Insurance and Social Protection Policy. Social Protection Discussion Paper Series. World Bank, Washington

Whittington D (2002) Improving the performance of contingent valuation studies in developing countries. Environ Resour Econ 22:323-367

Figure 1: The aggregate demand curve for CBHI without calibration

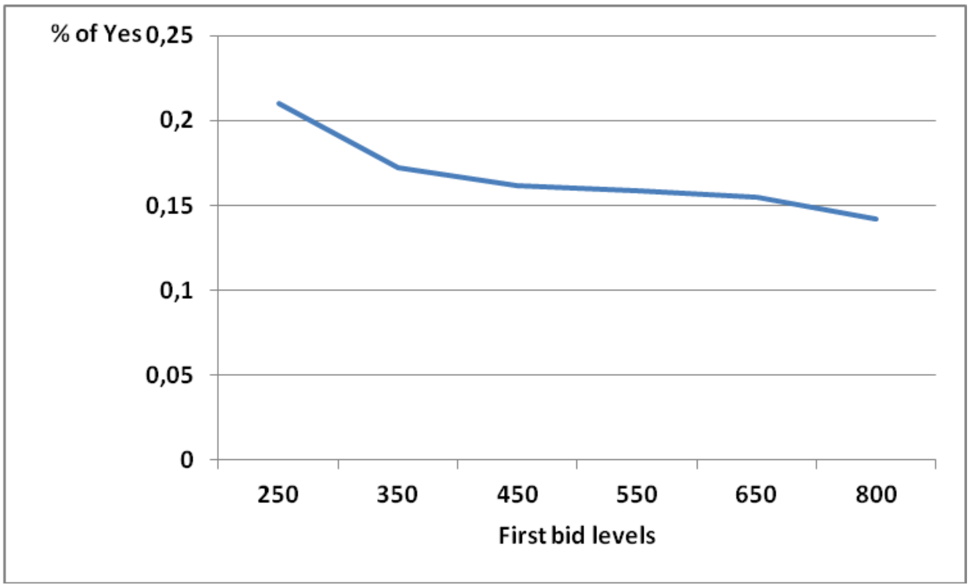

Table 1: Mean willingness to pay for SBDC

Statistic	*Value without calibration*	*Value with calibration*
μ_{SBDC}	1064.95 (140.74)	975.49 (179.41)
Krinsky-Robb[1]	[875.43 1598.91]	[755.43 2056.21]
CI/Mean	0.68	1.33

Notes: [1] confidence interval of the mean WTP obtained by Monte Carlo simulations on 50, 000 draws. Standard errors are in brackets.

Table 2: Mean willingness to pay for DBDC

Statistic	*Value without calibration*	*Value with calibration*
μ_{DBDC}	973.63 (38.03)	938.74 (86.86)
Krinsky-Robb[1]	[901.49 1052.29]	[797.48 1181.51]
CI/Mean	0.15	0.41

Table 3: Difference between mean WTP for SBDC and DBDC

Statistic	*Value without calibration*	*Value with calibration*
μ_{SBDC}	1064.95 (140.74)	975.49 (179.41)
μ_{DBDC}	973.63 (38.03)	938.74 (86.86)
$\Delta = \mu_{SBDC} - \mu_{DBDC}$	91.32 (1010.68)	36.75 (10431.83)
P-value[2]	0.83	0.65

Notes: [2] this is the achieved significance level.