"If you have the flu symptoms, your asymptomatic spouse may better answer the willingness-to-pay question".

Evidence from a Double-Bounded Dichotomous Choice Model with Heterogeneous Anchoring

Michaël Schwarzinger[a,b,⋆]

Fabrice Carrat[b] Stéphane Luchini[c]

[a]Center for Health Policy, Freeman Spogli Institute for International Studies / Center for Primary Care & Outcomes Research, School of Medicine, Stanford University, CA 94305-6019, USA

[b]INSERM U707, Paris-Université Pierre et Marie Curie-Paris 6, UMR-S 707, Paris, 75012, France

[c]GREQAM UMR 6579-IDEP, CNRS, Centre de la Vieille Charité, Marseille, 13002, France

February 2009

⋆Corresponding author: Tel: +1.650.723.0984;
Fax: +1.650.723.1919;
E-mail address: schwarzi@stanford.edu
Address: CHP/PCOR, 117 Encina Commons,
Stanford, CA 94305-6019, USA

Abstract

The small sample size of contingent valuation (CV) surveys conducted in patients may have limited the use of the single-bounded (SB) dichotomous choice format which is recommended in environmental economics. In this paper, we explore two ways to increase the statistical efficiency of the SB format: 1) by the inclusion of proxies in addition to patients; 2) by the addition of a follow-up dichotomous question, *i.e.* the double-bounded (DB) dichotomous choice format. We found that patients (n=223) and spouses (n=64) answering on behalf of the patient had on average a similar willingness-to-pay for earlier alleviation of flu symptoms. However, a patient was significantly more likely to anchor his/her answer on the first bid as compared to a spouse. Finally, our original DB model with shift effect and heterogeneous anchoring reconciled the discrepancies found in willingness-to-pay statistics between SB and DB models in keeping with increased statistical efficiency.

Keywords: Contingent valuation; double-bounded dichotomous choice; Patient; proxy; anchoring; structural shift; influenza

JEL Classification: I1, C2, H4

1 Introduction

Contingent valuation (CV) is used to assess the monetary value of non-market goods or services. This involves constructing a hypothetical market for those goods or services and directly asking individuals the maximum amount they are "willing-to-pay" (WTP) to get them. The CV method has been principally used in environmental and transport economics (Bateman and Willis, 1999; Jones-Lee, 1989; Beattie et al., 1998; Carthy et al., 1998) accounting for 70% of the 2,071 studies referenced in the Environmental Valuation Reference Inventory.[1] Over the last decade, CV has become increasingly popular in health economics to value health improvement from a patient's perspective (Olsen and Smith, 2001).

Several health economists recently acknowledged that transferring standards of practice from environmental economics to health economics may not be that straightforward (Smith, 2000, 2003; Hanley et al., 2003). To begin, the construction and presentation of the CV market should take into account the specific features of health care services: 1) scenario descriptions should include important and relevant dimensions to the patients, and these should be elicited in pilot studies; 2) elements of risk or uncertainty should be provided according to the perspective of the surveyed population, i.e. either the uncertain benefits of a health intervention in patients or the risk of getting a disease in the general population; 3) the payment vehicle, e.g. out-of-pocket payments, as well as the time period over which WTP is assessed, should be in accordance with the country health care financing system (Smith, 2003). If these features of health care services are not properly defined, then the CV market may not be realistic enough and WTP answers may be altered by a strong "hypothetical bias" (Liljas and Blumenschein, 2000).

Next, WTP can be elicited in CV surveys through various questionnaire formats (Smith, 2000;

[1] The Environmental Valuation Reference Inventory can be found at `http://www.evri.ec.gc.ca/` – accessed on the 12th of July 2008.

Hanley et al., 2003; Frew et al., 2003). The single-bounded (SB) dichotomous choice format was introduced by Bishop and Heberlein (1979). The respondent is asked if he/she would be willing to pay or not a given amount (bid) for the provision of the non-market goods or services. The proposed bids are varied across respondents to trace out the WTP distribution and to estimate WTP statistics. The SB dichotomous choice format has been recommended in environmental economics (Arrow et al., 1993). It is consistent with referendum voting behavior which is an incentive compatible mechanism for provision of public goods. In addition, a "take-it-or-leave-it" question is easier for respondents to answer because it is a more market-like question as compared to an open-ended question. However, the SB dichotomous choice approach requires large sample sizes to reach a sufficient level of precision of WTP estimates (Hanemann and Kanninen, 1999). In environmental economics surveys, the answers of the general population are of interest and the survey budget is the sole constraint to the sample size. In health economics, the answers of patients are usually of interest (Olsen and Smith, 2001); therefore the type of respondent limits the sample size as compared to the general population. Accordingly, other questionnaire formats that require fewer respondents have been used, i.e. open-ended WTP questions and payment cards (Olsen and Smith, 2001).

If health economics CV surveys are to follow the SB dichotomous choice format recommended in environmental economics, then increasing statistical efficiency (lowering the standard error) of the mean WTP estimate may be achieved in several ways: 1) increasing the sample size by surveying other types of respondents than the patient; 2) increasing information about each respondent's WTP by adding a follow-up dichotomous choice question; and 3) optimizing the bid design.

The type of respondent is a crucial issue in health economics (Dolan, 1999). When a patient-based perspective is relevant for the CV survey, the inclusion of proxies as defined by relatives living with patients is an interesting way to increase the sample size. The question begged is: Will the proxies have the same WTP as the patients? The literature in health economics is scarce on the level of agreement between proxies and patients. In previous CV surveys, proxies were interviewed on behalf of the patient, either parents of sick children (Sorum, 1999; Liu et al., 2000; Sach et al., 2004) or spouses of cognitively impaired patients (Chiu et al., 1998; Mulvaney-Day, 2005). Similarly, proxies were mostly interviewed on health status of patients when the need for proxies was the most salient. The level of agreement between proxies and patients remains unclear from the few studies having compared both of their answers as illustrated by the following results: 1) agreement on functional status was better for observable behaviors such as "Activities of Daily Living" rather than for more subjective areas such as functional limitations, sensorial limitations or mental health (Todorov and Kirchner, 2000); 2) agreement was only moderate on health-related quality of life for patients with mild cognitive impairment (Sneeuw et al., 1997; Tamim et al., 2002); 3) agreement was good on the utility of hypothetical disease vignettes (Saigal et al., 1999; Lenert et al., 2000; Dawson et al., 2008); 4) agreement was poor on the utility of actual health states of 80-year patients who were willing to trade significantly less time for a healthy life than their proxies (Tsevat et al., 1998).

The statistical efficiency of health economics CV surveys may also be increased by using the double-bounded (DB) dichotomous choice method. Surprisingly, the DB dichotomous choice method has received little attention in health economics (Bala et al., 1998; Slothuus and Brooks, 2000; Clarke, 2000; Dalmau-Matarrodona, 2001; Kennedy, 2002; Watson and Ryan, 2007). The DB dichotomous choice method refines the information on each respondent's WTP by adding to the initial dichotomous choice a follow-up dichotomous question: a higher bid is proposed if the initial bid was accepted, and a lower bid otherwise (Hanemann, 1985). Hanemann et al. (1991) showed that the DB method was asymptotically more efficient than the SB method, whereas a third discrete question delivered relatively little efficiency gains compared to those achieved by the DB method (Scarpa and Bateman, 2000). However, discrepancies in WTP statistics were commonly observed between the SB and DB methods. Several psychological and behavioral phenomenons were proposed to explain inconsistency between SB and DB methods. Starting point bias is a framing effect where respondents anchor their stated WTP value to the initial bid that represents a reasonable value (Boyle et al., 1985; Herriges and Shogren, 1996; Boyle et al., 1997; Frykblom and Shogren, 2000; Whitehead, 2002). Assuming a constant anchoring effect among respondents, Herriges and Shogren (1996) showed that both mean and variance of the prior WTP could be biased. The follow-up question can also induce behaviors such as "yea-saying" or incentive incompatibility resulting in a "shift" in respondent's answers between the first and second dichotomous question. "Yea-saying" behavior occurs when a respondent is more likely to answer "yes" to the follow-up dichotomous choice question when he/she has previously answered "yes" to the first dichotomous choice question (see for instance Alberini et al. 1997). Incentive incompatibility occurs when a respondent who initially said "yes" is asked the same valuation question at a higher price, which would be equivalent to raising the price after the purchase has been made in a market setting, and the respondent could then respond with a "protest no" to the follow-up question even if his true WTP is higher than the increased price (Alberini et al., 1997; Whitehead, 2002; Carson and Groves, 2007).

Finally, dichotomous choice-related WTP statistics are analytically a function of both the set of bids proposed to respondents and the unknown true WTP distribution. Changing the set of bids may result in significantly different WTP statistics (Boyle et al., 1998; Cameron and Hupper, 1991; Cooper and Loomis, 1992; Kanninen, 1995a). While the mean WTP estimate converges asymptotically to the true parameter, standard deviation estimate still depends on the proposed bids (Abdelbasit and Plackett, 1983; Sitter, 1992; Hanemann and Kanninen, 1999). The set of bids is usually selected to cover WTP dispersion as found in a pilot study. Otherwise, optimal bid designs rely typically on the choice of one, two or three bids that depend on the underlying true WTP distribution and the number of observations (Abdelbasit and Plackett, 1983; Minkin, 1987; Ford et al., 1992; Nyquist, 1992; Alberini, 1995b). Paradoxically, an optimal bid design provides optimal efficiency when the underlying true distribution and parameters are known. Sequential bid design overcomes partially this limitation (Abdelbasit and Plackett, 1983; Minkin, 1987; Nyquist, 1992; Kanninen, 1995b). In sequential CV sub-experiments, the WTP distribution is estimated using all available observations so that a new "optimal" bid design is derived for use with the subsequent iteration of the survey. While successive updates improve the efficiency of the design for poor initial estimates, the sequential bid design is questionable in practice.

The previous literature review showed that each way to increase statistical efficiency of SB dichotomous choice data in health economics has the potential to bias WTP statistics. In this paper, we explore the statistical effects of increasing sample size by surveying patients as well as proxies and increasing information about each respondent's WTP by adding a follow-up dichotomous choice question. WTP for earlier alleviation of flu symptoms was elicited among adult patients or spouses by use of the DB dichotomous choice method. In each household surveyed, the respondent was either the patient or his/her spouse. We assumed that the SB dichotomous choice data of patients provided an unbiased WTP distribution of reference for earlier alleviation of flu symptoms. Three DB dichotomous choice models were compared: the "naïve" DB model; the DB model with shift effect and constant anchoring; and an original DB model with shift effect and heterogeneous anchoring, where anchoring depends on individual characteristics. In the four dichotomous choice models, a specific parameter accounted for potential discrepancies on mean WTP estimate between patients and spouses. Accordingly, we assessed the level of agreement on mean WTP between patients and spouses as well as potential bias on WTP statistics vs. statistical efficiency achieved by the follow-up dichotomous choice question.

The paper is organized as follows. Section 2 describes the DB dichotomous choice data used that relate to the WTP for earlier alleviation of flu symptoms; the econometric model; and the estimation procedure. Section 3 provides the comparative results between the four dichotomous choice models.

In Section 4, we discuss the results and the implications of the study.

2 Material and methods

2.1 Contingent valuation data

The CV survey was nested into a prospective household contact study on influenza burden of disease in France. Patients suffering from influenza-like illness were included when visiting their general practitioner during the influenza outbreak from January to March 2000. The general practitioner checked off flu symptoms (a flu symptom score was computed at entry with a score of 1 indicating the presence of all symptoms and a score of 0 indicating the absence of symptoms), and took nasal swabs for testing influenza A and B viruses. A household diary questionnaire had to be filled up over a 2-week period by the patient and each household member. The diary questionnaire included daily questions about flu symptoms and medical consumptions, as well as questions about individual past experience with flu (vaccination this year, influenza-like illness the previous year), the presence of chronic disease, risk factors for lung disease, and household socio-economic characteristics. More details on the prospective household contact study are provided elsewhere (Carrat et al., 2002).

On day three after inclusion of the patient by the general practitioner, a CV survey was conducted by telephone. Due to budget constraints, we interviewed the first 435 recruited households (out of 946 – 45%) (Carrat et al., 2001, 2002). Respondents were asked about their WTP for earlier alleviation of flu symptoms. In each household, the respondent was either the adult patient or his/her spouse depending on who was answering the phone. Consequently, two WTP scenarios were prepared to reflect both a patient-based perspective where the respondent is the patient and a proxy-based perspective where the spouse is responding on behalf of the patient. Whatever the WTP scenario, respondents were randomly asked about their WTP for the alleviation of flu symptoms by 1 day or 3 days (see Appendix A).

A pilot CV study was performed by mail in autumn 1999 to test WTP scenarios for the alleviation of flu symptoms by 1 day and to select the bids. After visiting their general practitioner, patients with flu symptoms were sent a WTP question with a single bid randomly selected from [25; 50; 100; 200; 400] French Francs (FF). Two hundred seven patients of mean (sd.) age of 29 (21) years including 54% female, returned the CV questionnaire. The return rate differed significantly according to the bid proposed (25FF: 23%; 50FF: 15%; 100FF: 21%; 200FF: 34%; 400FF: 6%). In the standard probit model, the mean WTP for the alleviation of flu symptoms by 1 day was 137FF (deciles: 20th: 26FF; 40th: 103FF; 60th: 170FF; 80th: 247FF; 100th: 1216FF). Despite a possible bias on WTP distribution, we decided to keep the same set of bids for the main telephone CV survey as it traced out well the WTP distribution.

The elicitation mechanism in the telephone survey was based on the double-bounded dichotomous choice format. In other words, respondents were asked for two successive bids. The initial bid proposed to the respondent was randomly selected from [25; 50; 100; 200; 400] FF or [50; 100; 200; 400; 800] FF for the alleviation of flu symptoms by 1 day or 3 days, respectively. The follow-up bid was set at twice the initial bid if the respondent accepted the initial bid, and at half the initial bid otherwise. Out of 435 household respondents, 287 (66.0%) WTP interviews were used for the analysis, 10 (2.3%) were not willing to participate in the CV survey, and 138 were excluded from the analysis because of missing information - 95 (21.8%) households did not send back the diary questionnaire including clinical and socio-economic variables, and 43 (9.9%) households did not complete the household income question. In the following, all bids, costs and WTP were converted to Euros (1 French Franc=0.152449€).

2.2 Econometric Model

The true WTP distribution for earlier alleviation of flu symptoms is unknown. For comparison purposes among types of respondent and dichotomous choice models, we assumed that the SB model of patients' data provided an unbiased estimator of WTP statistics for earlier alleviation of flu symptoms. In all dichotomous choice models, we modeled the prior WTP of the proxy as a combination of the expected patient's prior WTP and a parameter θ. This parameter accounts for potential discrepancies on mean WTP estimate between patients and proxies. Potential heteroscedasticity between patients and proxies was also taken into account by an error term specific to each perspective. We used standard interval-data modeling of double-bounded dichotomous choice data. Interval-data modeling assumes no changes in underlying WTP distribution between the first bid (the prior WTP according to the SB model) and the follow-up bid (the posterior WTP according to the DB model). Interval-data modeling allows specification of the psychological effect (starting point bias) and behavioral effect (shift effect) induced by the follow-up discrete question. In addition to the naïve DB model without psychological/behavioral effects (where prior=posterior WTP) and the DB model with shift effect and constant anchoring (where prior differs from posterior WTP), we developed an original DB model with shift effect and heterogeneous anchoring–where anchoring depends on individual characteristics, e.g., the type of respondent. Originally, the econometrics developed in this paper allowed: 1) to assess the level of agreement on mean WTP estimates between patients and spouses in all dichotomous choice models; 2) to assess potential bias on WTP statistics vs. statistical efficiency induced by the follow-up dichotomous choice question.

2.2.1 Prior WTP of patients and proxies

First, let us consider the patient-based perspective where the patient answers the WTP question. The prior WTP of the patient is defined as follows:

$$W_{i1}^{*} = X_i \beta + \epsilon_i, \qquad \epsilon_i \sim NID(0, \sigma^2) \tag{1}$$

where X_i is a set of explanatory variables which captures consumer preferences, and unknown parameters β and σ are respectively a vector and a scalar.

Second, let us consider the proxy-based perspective where the proxy answers the WTP question on behalf of the patient. The prior WTP of the proxy j answering on the behalf of the patient i is defined as follows:

$$W_{j1}^{*} = E(W_{i1}^{*}) + \theta + \nu_j, \qquad \nu_j \sim NID(0, \omega^2) \tag{2}$$

In our specification, the prior WTP of the proxy is a combination of the expected patient's prior WTP and a parameter θ. The parameter θ accounts for potential discrepancies between the prior WTP of the proxy and that of the patient. When θ is positive, the proxy overstates the patient's prior WTP and understates it when it is negative. If θ is not significantly different from zero, the prior WTP of the proxy does not differ from that of the patient. The error term ν_j is specific to the proxy-based perspective and it has a different variance ω^2 than the error term ϵ_i in the patient-based perspective.[2] We therefore take into account a potential heteroscedasticity induced by the two perspectives that could lead to misleading results otherwise (which is likely to occur in econometric models with discrete dependent variables, see Yatchew and Griliches 1984).

[2]We did not restrict the structure of the relationship between ω and σ so that ω can be either higher or lower than σ.

Equations (1) and (2) can be combined into one equation such that

$$W_{i1}^* = X_i\beta + \theta P_i + u_i, \qquad u_i \sim NID(0, \sigma_i^2 = \sigma^2(1 - P_i) + \omega^2 P_i) \tag{3}$$

where P_i is a dummy variable which is one when the proxy answers on behalf of the patient and zero otherwise. Equation (3) therefore reduces to (1) when $P_i = 0$ and to (2) when $P_i = 1$.

2.2.2 Dichotomous choice modeling

Let us now focus on the econometrics of the elicitation procedure. In the single-bounded dichotomous choice model (SB model), the respondent's answer to the first bid b_{i1} is observed, but not W_{i1}^*. The answer of respondent i is then defined as follows:

$$W_{i1} = 1 \quad \text{if } W_{i1}^* \geq b_{i1} \quad \text{, and } W_{i1} = 0 \quad \text{otherwise.} \tag{4}$$

Assuming that the error term is identically and independently distributed according to a normal distribution with mean 0 and standard deviation σ_i, W_{i1}^* is estimated by means of the standard probit model with W_{i1} as the dependent variable.

In the double-bounded dichotomous choice model, a follow-up bid is proposed to the respondent. When respondent i answers *yes* to the first bid, b_{i1}, the second bid b_{i2} is higher, and lower otherwise. The "naïve" DB model assumes that WTP is independent of both first bid (b_{i1}) and follow-up bid (b_{i2}). W_{i1}^* is estimated by means of standard multinomial model.

However, the follow-up discrete answer may be biased. First, the first bid may induce a starting point bias as respondents may anchor their WTP on this bid which may be seen as providing information about the "correct" price. Thus, respondents combine their prior WTP, W_{i1}^*, with the first bid b_{i1} to form a revised or posterior WTP, W_{i2}^*. To account for this phenomenon, Herriges and Shogren (1996) proposed a DB model with constant anchoring in which the WTP W_{i2}^* is defined as a convex combination of the prior WTP W_{i1}^* and the first bid b_{i1} such that:

$$\begin{cases} W_{i1}^* = & X_i\beta + \theta P_i + u_i \\ \\ W_{i2}^* = & (1-\gamma)W_{i1}^* + \gamma b_{i1} \end{cases} \tag{5}$$

where γ lies between 0 and 1.

Secondly, the follow-up bid may induce behaviors such as yea-saying or incentive incompatibility. Yea-saying and incentive incompatibility are usually treated as a "shift" in respondents' answers between the first and second dichotomous question (Alberini et al., 1997; McLeod and Bergland, 1999; Whitehead, 2002). Alberini et al. proposed a DB model with a shift parameter considering that the posterior WTP is based on the true prior WTP with a shift parameter (Alberini et al., 1997). Although considered separately in previous papers, Whitehead (2002) proposed an econometric model which takes into account both constant anchoring and shift induced by the follow-up question (DB model with shift effect and constant anchoring). Formally, the econometric model is defined as follows:

$$\begin{cases} W_{i1}^* = & X_i\beta + \theta P_i + u_i \\ \\ W_{i2}^* = & (1-\gamma)W_{i1}^* + \gamma b_{i1} + \delta \end{cases} \tag{6}$$

where γ lies between 0 and 1, and δ is the "structural shift" parameter. The structural shift parameter could be positive due to yea-saying behavior (Kanninen, 1995a) or negative due to incentive incompatibility (Alberini et al., 1997).

In the previous DB model, the anchoring bias parameter γ was the same for all respondents. However, in the presence of individual heterogeneity, results based on standard regression can be seriously misleading if this heterogeneity is not taken into account (see Aprahamian et al. 2007 and Aprahamian et al. 2008). In this paper, we propose a DB model with shift effect and heterogeneous anchoring where the anchoring parameter is variable depending on individual characteristics. The model is now defined as:

$$\begin{cases} W_{i1}^{*} = & X_i\beta + \theta P_i + u_i \\ \\ W_{i2}^{*} = & [1-\gamma(Z_i;\eta)]W_{i1}^{*} + \gamma(Z_i;\eta)b_{i1} + \delta \end{cases} \tag{7}$$

where $\gamma(.)$ is a function of explanatory variables Z_i and a parameter vector η. As in previous DB model with constant anchoring, $\gamma(.)$ has to lie between 0 and 1. When one only includes the constant term in the explanatory variables Z_i, the heterogeneous anchoring model 7 reduces to equation 5.

This model is of particular interest as it allows testing for different anchoring patterns depending on the perspective of the respondent, either patient-based or proxy-based. This is done by considering the dummy variable P_i as an explanatory variable included in the function $\gamma(.)$.

2.3 Estimation procedure

Both the prior WTP W_{i1}^{*} and the posterior WTP W_{i2}^{*} are not directly asked for. To estimate our econometric model, we have to define the probabilities associated to each pair of responses (*no, no*; *no, yes*; *yes, no*; *yes, yes*). The probability of a *yes* to the first bid depends on whether WTP is greater than the bid b_{i1} and is then defined as:

$$P(\text{yes}) = P(W_{i1} > b_{i1}) = \Phi\left[\frac{b_{i1} - X_i\beta - \theta P_i}{\sigma_i}\right] \tag{8}$$

where $\Phi(.)$ is the normal c.d.f. since error terms are assumed to be normally distributed. Using equation 7, the probability of a *yes* to the first bid and a *no* to the second bid is given by:

$$P_i^{yn} = P(W_{i2} > b_{i2}) - P(\text{yes}) \tag{9}$$

$$= \Phi\left[\frac{b_{i2} - \gamma(Z_i;\eta)b_{i1} - \delta}{(1-\gamma(Z_i;\eta))\sigma_i} - \frac{X_i\beta + \theta P_i}{\sigma_i}\right] - \Phi\left[\frac{b_{i1} - X_i\beta - \theta P_i}{\sigma_i}\right] \tag{10}$$

where the function $\gamma(Z_i;\eta)$ is such that

$$\gamma(Z_i;\eta) = \gamma_{Pa}(1-P_i) + \gamma_{Pr}P_i + M_i\lambda \tag{11}$$

where γ_{Pa} is the anchoring parameter when the patient answers, γ_{Pr} is the anchoring parameter when the proxy answers on behalf of the patient and λ a vector of parameters associated with the variables M_i. The variables M_i are additional variables that may induce heterogeneity in anchoring. It is possible to allow for some more flexibility such that:

$$\gamma(Z_i;\eta) = \gamma_{Pa}(1-P_i) + \gamma_{Pr}P_i + M_i\lambda_{Pa}(1-P_i) + M_i\lambda_{Pr}P_i \tag{12}$$

In this specification, the effect of additional variables M_i on anchoring can differ depending on the perspective of the respondent.[3]

[3]Note that it is useful to force γ to lie between 0 and 1 as commonly done with the double exponential function: $\gamma(Z_i;\eta) = \exp(-\exp(Z_i\eta))$.

The probabilities associated with other possible pair of response are calculated in the same way:

$$P_i^{nn} = \Phi\left[\frac{b_{i2} - \gamma(Z_i;\eta)b_{i1} - \delta}{(1-\gamma(Z_i;\eta))\sigma_i} - \frac{X\beta + \theta P_i}{\sigma_i}\right] \tag{13}$$

$$P_i^{ny} = \Phi\left[\frac{b_{i1} - X_i\beta - \theta P_i}{\sigma_i}\right] - \Phi\left[\frac{b_{i2} - \gamma(Z_i;\eta)b_{i1} - \delta}{(1-\gamma(Z_i;\lambda))\sigma_i} - \frac{X\beta + \theta P_i}{\sigma_i}\right] \tag{14}$$

$$P_i^{yy} = 1 - \Phi\left[\frac{b_{i2} - \gamma(Z_i;\eta)b_{i1} - \delta}{(1-\gamma(Z_i;\eta))\sigma_i} - \frac{X\beta + \theta P_i}{\sigma_i}\right] \tag{15}$$

Accordingly, the following log-likelihood function is defined as:

$$\ell(y;\beta,\theta,\delta,\eta,\sigma,\omega) = \sum_{i\in N}\left[I_i^{yy}\log(P_i^{yy}) + I_i^{yn}\log(P_i^{yn}) + I_i^{ny}\log(P_i^{ny}) + I_i^{nn}\log(P_i^{nn})\right] \tag{16}$$

where I_i^{yy}, I_i^{yn}, I_i^{ny} and I_i^{nn} are dummy variables with, for instance, $I_i^{yy} = 1$ if respondent answered *yes* to the first bid and *yes* to the second bid, and 0 otherwise. The log-likelihood function is estimated by maximum likelihood.

We first tested for the significance of γ_{Pa} and γ_{Pr} in the $\gamma(.)$ function. We then tested additional variables M_i that may significantly induce heterogeneous anchoring. All explanatory variables recorded in the survey were tested one by one by means of standard likelihood ratio (LR) tests according to specification (11). We then introduced simultaneously all significant M_i variables (p-value less than 0.10) in the $\gamma(.)$ function by using specification (12). Hence, the remaining variables were included in the final DB model with shift effect and heterogeneous anchoring. Finally, we assessed the relevance of the DB model with shift effect and heterogeneous anchoring by testing the null hypothesis of no anchoring and no shift effect ("naïve" DB model) and the null hypothesis of constant anchoring and structural shift (DB model with shift effect and constant anchoring). This was done by means of LR tests.

3 Empirical results

Out of the 287 adult patients with flu symptoms included in the analysis, there were 51.9% male respondents and the mean (sd.) age was 41 (16) years (Appendix B). They had a mean flu symptom score at entry of 0.41 (0.16) with a higher score indicating higher severity, and 62.0% of cases were tested positive for influenza A or B virus, 14.3% had an influenza-like illness the previous year, and 10.8% had flu vaccination during the year of the study. Households had on average 3.1 (1.2) members, and a monthly income under €1,524 in 31.0% and above €3,049 in 21.3% of the cases. At the time of the CV survey, another member had flu symptoms in 24.4% of households in addition to the patient recruited by the general practitioner. Respondents provided their WTP for earlier alleviation of flu symptoms for themselves when they were the patient (77.7%) or for their symptomatic spouse (22.3%) when they were not the patient.

The discrete choice CV data is given in Tables 1.a and 1.b according to the size of clinical benefits (alleviation of flu symptoms by 1 or 3 days respectively) and the respondents' perspective. The columns separate respondents according to which initial bid was proposed and the rows provide the bid acceptance rate assuming that a respondent who accepts a bid accepts all lower bids. All columns show a decreasing probability of bid acceptance from lower to higher follow-up bids. In most instances when the respondent was the patient, the rows show a positive relationship between increasing initial bids and the probability of follow-up bid acceptance. For example, the probability of acceptance of the bid €61 for the alleviation of flu symptoms by 3 days increased from 16% when the initial bid was €30.5 to 68% with an initial bid of €122 ($p<0.001$, Fisher's exact test). Accordingly, the presence of a starting point bias was likely when the respondent was the patient. When the respondent was the spouse answering on behalf of the patient, the trend was less salient, although sample size was smaller.

Table 2 presents the results of the four econometric models we used to describe the CV dichotomous choice data: the SB model (only the answer to the initial bid is considered), the “naïve” DB model (both answers to the initial and follow-up bids are considered), the DB model taking into account the starting point bias with constant anchoring (a unique parameter γ) and our original DB model with heterogeneous anchoring that depends on respondent’s characteristics. A structural shift parameter was included in the two later models to control for possible behaviors associated with the follow-up question (Whitehead, 2002).

A large set of explanatory variables was considered to avoid a potential misspecification bias. All CV dichotomous choice models showed construct validity: WTP increased significantly with the size of the clinical benefit from 1 to 3 days of alleviation of flu symptoms (INT3_DAYS significant in all four models) and WTP increased with income (although, not significantly). Among other explanatory variables of WTP, having had influenza-like illness the previous year significantly decreased WTP in the DB model with heterogeneous anchoring ($p = .080$). Other explanatory variables were not significant.

The level of agreement between patients and spouses for the WTP for earlier alleviation of flu symptoms was assessed by means of a parameter θ (INT_PROXY) accounting for discrepancies between the expectation of the patient’s prior WTP and that of the proxy. In all four econometric models, θ was not significantly different from zero ($p = .246$ in the SB model; $p = .648$ in the “naïve” DB model; $p = .591$ in the DB model with constant anchoring; $p = .786$ in the DB model with heterogeneous anchoring).

The “naïve” DB model provided efficiency gains when compared to the SB model, i.e. the standard error of mean WTP estimate decreased from 1.23 to 0.79 (Table 2). However, the “naïve” DB model significantly biased the mean WTP downward from €58.8 (95% CI [56.4 to 61.3] in the SB model) to €44.7 (95% CI [43.12 to 46.3] in the “naïve” DB model). In the DB model with shift effect and constant anchoring, we found a significant constant anchoring effect $\gamma = .421$ ($p < .001$), but no significant structural shift ($\delta = -.79$; $p = .391$). The anchoring effect estimate seemed reasonable and was within the range of previous estimates in CV environmental surveys (from $\gamma = .22$ to $\gamma = .38$) (Herriges and Shogren, 1996; Green et al., 1998; Holmes and Boyle, 2005). However, efficiency gains of the DB model with constant anchoring had to be balanced by the remaining bias downward of the mean WTP estimate (€48.1, 95% CI [46.4 to 49.8])

In the DB model with shift effect and heterogeneous anchoring, the structural shift parameter was also negative ($\delta = -.67$) and not significant ($p = .483$). Explanatory variables of the heterogeneous anchoring were selected in a three-stage procedure. We first considered whether anchoring was different across respondents' perspectives. The anchoring effect was significant when the respondent was the patient, but it was not the case when the respondent was the spouse answering on behalf of the patient ($p < 0.001$ and $p = 0.601$, respectively).

We then added explanatory variables one by one to test for further heterogeneity in anchoring of patients' answers (tests presented in Appendix C). The only significant explanatory variables were the occurrence of flu symptoms in another household member at the time of the CV survey (INT_OTHERFLU; $p = .012$) and the flu symptom score (FLU_SCORE; $p = .058$). Finally, we tested a heterogeneous anchoring model with 4 parameters whether INT_OTHERFLU and FLU_SCORE had a differential effect on the anchoring depending on the respondents' perspectives. Only one parameter was significant, i.e. INT_OTHERFLU on the anchoring effect of the patient-based perspective ($p = .029$).[4] Accordingly, the final DB model with heterogeneous anchoring presented in Table 2 included an anchoring parameter when the respondent was the patient depending on the occurrence of new cases of influenza-like illness in the household at time of interview ($\gamma = .563$ for INT_ONESELF only and decreased by 0.230 for INT_ONESELF $\times$ INT_OTHERFLU), and no anchoring parameter for proxies.

[4]INT_OTHERFLU was not significant in the anchoring effect of the proxy perspective ($p = .901$). FLU_SCORE was not significant in both cases ($p = .138$ and $p = .475$)

Among the three DB econometric models, the DB model with heterogeneous anchoring presented the best fit for our CV data. The DB model with heterogeneous anchoring strongly rejected the null hypothesis of no anchoring and no shift effect ($p < .001$) and the null hypothesis of constant anchoring and shift effect ($p = .004$). As compared to the SB model, the DB model with heterogeneous anchoring provided efficiency gains (the standard error of mean WTP estimate decreased from 1.23 to 0.93) and the less biased mean WTP estimate among DB models (50.5 €, 95% CI [48.8 to 52.6] in the DB model with heterogeneous anchoring as compared to €58.8, 95% CI [56.4 to 61.3] in the SB model). In addition, the standard deviation estimates obtained with the DB model with heterogeneous anchoring model were in accordance with those found in the SB model. In the SB model, the standard deviation of WTP of proxies ω (25.6, se. 7.8) was smaller than that of patients σ (63.8, se. 13.2), with no significant difference ($p = .235$). The DB model with heterogeneous anchoring showed a similar pattern: $\omega = 42.0$ (se. 5.7) and $\sigma = 55.1$ (se. 7.2) with no significant difference ($p = .127$). Opposite results were found in other DB models where the standard deviation of the WTP of patients was smaller than that of proxies ($\omega > \sigma$). This phenomenon was significantly marked in the DB model with constant anchoring ($p = .147$ in the "naïve" DB model and $p = .008$ in the DB model with constant anchoring).

Figure 1 illustrates the WTP for the earlier alleviation of flu symptoms by 1 day according to the level of anchoring of the respondent using the DB model with heterogeneous anchoring. In this example, the patient had the average characteristics provided in Appendix A. The spouse had no anchoring on the first bid and therefore provided a constant WTP of €50.5 whatever the initial bid. On the contrary, the patient anchored on the first bid. Accordingly, his WTP increased with the initial bid proposed, although to a lesser extent in case of another household member having the flu at the time of the CV survey. If the actual price of the drug related to the earlier alleviation of flu symptoms by 1 day had been proposed as a first bid to those later respondents (about €30)

(Schwarzinger et al., 2003) their stated WTP would have been €39.4 and €44.3 respectively. If the initial bid had been 3-times higher (€90), their stated WTP would have been €73.2 and €64.3, respectively.

4 Discussion

The context of health care services may partly limit the transfer of the standards of practice of contingent valuation (CV) surveys from environmental economics to health economics (Smith, 2000, 2003; Hanley et al., 2003). In particular, the relatively small sample size of CV surveys conducted in patients as compared to the general population in environmental economics may have limited the use of the recommended single-bounded (SB) dichotomous choice format (Arrow et al., 1993). Two ways to increase the statistical efficiency of health economics CV surveys based on the SB dichotomous choice format were explored in this paper: 1) the inclusion of proxies in addition to patients; 2) the addition of a follow-up discrete question as commonly used in environmental economics, *i.e.* the double-bounded (DB) dichotomous choice format.

In this paper, we analyzed DB dichotomous choice data for earlier alleviation of flu symptoms where patients or proxies were interviewed. We estimated a DB dichotomous choice model that accounted for respondent's perspective and heterogeneous anchoring. Our results are three-fold. First, patients and proxies had on average a similar WTP. Second, our model showed that anchoring is a heterogeneous phenomenon. In particular, while a patient was likely to anchor his/her answer on the initial bid, a spouse answering on behalf of the patient was not. Finally, our original DB model with heterogeneous anchoring had the best statistics properties as compared to other DB dichotomous choice models.

In all dichotomous choice models, patients and proxies had on average a similar WTP for earlier alleviation of flu symptoms. Two arguments from the health state valuation literature support this original finding. Firstly, the knowledge of the health state plays a key role on health state valuation. Both patient and proxy have an intimate knowledge about the subject's health state, whether the general public is usually in need for a description of the health state (Brazier et al., 2005). In case of influenza-like illness, the knowledge of the health state is even more realistic for proxies as they probably suffered the condition in the past or even may be infected during current epidemics, e.g., 24% of households had at least one secondary case of flu at time of interview. Secondly, patients provide on average a better health state valuation than other types of respondents due to adaptation, coping and adjustment of the patient to a chronic condition (Cassileth et al., 1984; Damschroder et al., 2005). In case of influenza-like illness, i.e. an acute condition lasting about a week, these processes are unlikely to alter valuations between patients and proxies. Presumably, health benefits presented in the contingent valuation were valued by patients or proxies in reference to a disease of a similar perception. The economic interpretation of this finding is two-fold, though it is not possible to discriminate one interpretation from the other: either the proxy knows the preferences of his/her spouse and does not distort them during the interview, or the proxy has the same preferences as his/her spouse.[5]

Anchoring is a well-established phenomenon that has been recorded in many studies on healthy subjects (Tversky and Kahneman, 1974; Chapman and Johnson, 1994; Mussweiler, 2003). Previous DB models assumed anchoring as a constant effect across the surveyed population. As shown by this study in health economics and recent studies in environmental economics, anchoring may depend on individual characteristics (Flachaire et al., 2007; Aprahamian et al., 2007). In other words, framing effects may not be of the same magnitude for all respondents. To that extent, the DB model with heterogeneous anchoring provides an original means to measure the reliability of stated preferences

[5]Note that this interpretation is reminiscent of the distinction between *pure altruism* and *pure paternalism* (see Jones-Lee 1992b and Jones-Lee 1992a).

by the level of anchoring of respondents.

Although patients and spouses had a similar WTP for earlier alleviation of flu symptoms, the DB model with heterogeneous anchoring showed that patients had a significant level of anchoring in their answers as compared to spouses. In addition, the level of anchoring of patients decreased from γ=0.563 to 0.333 when another household member had flu symptoms at time of the interview. A possible explanation could involve notions of cognitive ability and sense of responsibility of the patient. On the one hand, former clinical experiments in adult volunteers showed that influenza viruses impaired cognitive performance to a level comparable to the deterioration seen in alcohol consumption (Smith et al., 1993). Accordingly, patients suffering from influenza-like illness may be cognitively impaired and more prone than asymptomatic proxies to framing effects in subjective valuations. On the other hand, the level of anchoring of patients decreased when another household member had flu. This trend may become critical when a dependent child becomes also ill as the adult patient's responsibility and reliability towards himself/herself and the child is more acute. It may lessen the impact of framing effects on valuation, but it does not cancel it.

Respondents in health economics may have salient characteristics, i.e. patient vs. asymptomatic proxy or disease-related characteristics such as cognitive impairment that could explain heterogeneous anchoring. It contrasts with environmental CV surveys on healthy subjects recording anchoring as a general phenomenon. However, recent environmental CV surveys have shown that some respondents were not likely to anchor their answer on the inititial bid. Flachaire et al. showed that respondents who had a more constructed opinion on the French natural park of Camargue were not subject to anchoring when asked for their WTP for a preservation program (Flachaire et al., 2007). Aprahamian et al. showed that 25% of respondents did not anchor when asked for their WTP for an air quality improvement (Aprahamian et al., 2007). The potential role on anchoring of explanatory variables related to the disease and recorded in health economics CV surveys should be subjected to further investigation.

Finally the DB model with heterogeneous anchoring was the best model among DB models to fit dichotomous choice WTP data for earlier alleviation of flu symptoms. It reconciled statistical discrepancies found previously between SB and DB models, i.e. the less biased mean estimate and similar standard deviation estimates. Similar results were found recently in a CV survey in environmental economics (Flachaire et al., 2007). Accordingly, the DB model with heterogeneous anchoring provides a means to both control for the anchoring effect on WTP estimation and optimize small sample sizes usually found in health economics CV surveys.

Several findings validated the DB model with heterogeneous anchoring presented in this paper. Firstly, WTP increased with the level of household income, in keeping with economic theory. Secondly, WTP passed a between-subject "scope test" since WTP increased significantly with the size of clinical benefit, i.e. from 1 to 3 days of earlier alleviation of flu symptoms (Hanley et al., 2003). Other explanatory variables could be explained intuitively: a positive relationship between WTP and reassurance (smoker at risk of complications, antibiotic consumption); a negative relationship between WTP and mistrust in clinical benefits of the hypothetical drug (influenza-like illness the previous year, occurrence of another case at home).

This study has some limitations. Firstly, the sample may not be representative of the French patients suffering from influenza-like illness. One-third of respondents in the CV survey were excluded because they did not send back the household diary questionnaire containing explanatory variables. However, there was no difference between excluded and included households according to the clinical variables recorded initially by the general practitioner. Secondly, the presentation of the CV market may not be realistic enough. Health benefits of the new treatment were limited to the description of an earlier alleviation of flu symptoms in accordance with the primary health outcome measured in related clinical trials (Cooper et al., 2003). However, elements of uncertainty such as the probability of success of the new treatment were not presented. Other features of the scenario descriptions such as the out-of-pocket payment were consistent with the lack of reimbursement of the drugs alleviating flu symptoms by 1 day and launched since then in the French health care system, i.e. zanamivir (RELENZA ®) in 1999 and oseltamivir (TAMIFLU ®) in 2002 (Schwarzinger et al., 2003). Further studies should validate the extent to which stated WTP from DB models relies to actual choices as it has been done with other WTP formats used in health economics (Onwujekwe et al., 2001; Blumenschein et al., 2001; Foreit and Foreit, 2003).

Finally, we used standard interval-data modeling which relies on the assumption that the underlying WTP distribution is identical between the first bid (SB model) and the follow-up bid (naïve DB model). Alternatively, the bivariate probit model relaxes this assumption by allowing for the possibility of different means, dispersions, and non-perfect correlation across the two answers (Cameron and Quiggin, 1994). Several arguments support our modeling approach. Theoretically, Monte Carlo simulations showed that interval-data mean WTP estimates are robust from departures from a perfect correlation, and they are often superior to the bivariate probit model in terms of the mean square error of estimates (Alberini, 1995a; Haab, 1998). Other simulations showed that bivariate probit model mean WTP estimates are unbiased, but does not always perform well in terms of statistical efficiency as compared to the SB model (Kang, 2006). Actual CV data sets from environmental economics (Kang, 2006) and health economics (Clarke, 2000; Kennedy, 2002) were inconclusive on the best modeling approach to fit double-bounded dichotomous choice data and relied eventually on the single-bounded model. In the present paper, we relied on interval-data modeling instead of the bivariate probit model to disentangle purposely the psychological effect (starting point bias) and behavioral effect (shift effect) induced by the follow-up discrete question.

5 Conclusion

The double-bounded (DB) dichotomous choice model has been less frequently used in contingent valuation surveys performed in health economics, relative to environmental economics. Recent studies in health economics using DB data did rely preferably on a single-bounded (SB) model due to “a starting point bias” (Clarke, 2000; Kennedy, 2002). This paper proposed a DB model with heterogeneous anchoring that reconciled the discrepancies that were commonly found between the DB and SB related-WTP statistics. After correction for heterogeneous anchoring, the DB method still refines the information about a respondent’s WTP, thereby allowing for smaller sample sizes than CV surveys relying on SB data. Furthermore, the model provides a means to test for the reliability of respondents’ answers in dichotomous choice CV surveys.

Previous CV surveys were mainly restricted to one type of target-population only. That is, either patients were invited to disclose their own evaluation or the general population was asked to reveal the value they attached to health programs. The novelty of our study is that it included, in addition to the patient, a second target group, namely the respective spouse of the patient. This evaluation from different perspectives brings to light some interesting results. In particular, it shows that patients and proxies provided similar WTP for an earlier alleviation of flu symptoms, but the patient-self evaluation tends to be less reliable. This result suggests some immediate practical consequences for economic evaluation in health care. Health evaluation of relatives, partners and other people who are in close contact with the patient should receive more attention. This is an interesting and new aspect that needs to be further explored in future research. Previous CV surveys only relied on proxies when the need for proxies was the most salient, e.g. children or cognitively impaired patients. As a result of this study, patients’ answers to WTP questions should be treated with more caution. However, patients still have more information about themselves than their relatives, thus their evaluation reveals some necessary information that should not be totally disregarded.

6 Acknowledgement

We would like to thank the patients, their families, and the 161 general practitioners of the French "Sentinelles" system who participated to the epidemiological study; Camille Sahler, MSc, Isabelle Goderel, BS, Audrey Lavenu, MSc, and Gregory Pannetier, BS, for technical assistance in data collection. We thank Frédéric Aprahamian, Susan Cleary, Cam Donaldson, Jean-Paul Moatti, Jason Shogren, Miriam Teschl, the Editor and two anonymous referees for helpful comments and remarks. We also thank Dana Conley for her help with editing. The CV study was funded by GlaxoSmithKline, Marly-le-Roi, France. The sponsors of the study had no further role in the study design, data collection, data analysis, data interpretation, writing, or the decision to submit the report for publication. The views expressed in this paper are those of the authors and usual caveat applies.

References

Abdelbasit, K. and R. Plackett (1983). Experimental design for binary data. *Journal of the American Statistical Association 78*(381), 90–98.

Alberini, A. (1995a). Efficiency vs bias of willingness to pay estimates: Bivariate and interval models. *Journal of Environmental Economics and Management 29*, 169–180.

Alberini, A. (1995b). Optimal designs for discrete choice contingent valuation surveys: Single-bound, double-bound, and bivariate models. *Journal of Environmental Economics and Management 28*(3), 287–306.

Alberini, A., B. Kanninen, and R. Carson (1997). Modeling response incentive effects in dichotomous choice contingent valuation data. *Land Econonomics 73*, 309–24.

Aprahamian, F., O. Chanel, and S. Luchini (2007). Modeling starting point bias and unobserved heterogeneity in contingent valuation surveys: an application to air pollution. *Amer J Agr Econ 89*(2), 533–47.

Aprahamian, F., O. Chanel, and S. Luchini (2008). Heterogeneous anchoring and the shift effect in iterative valuation questions. *Resource and Energy Economics 30*, 12–20.

Arrow, K., R. Solow, P. R. Portney, E. E. Leamer, R. Radner, and H. Schuman (1993). Report of the NOAA panel on contingent valuation. *Technical Report 58*(10), p1601–1614.

Bala, M., L. Wood, G. Zarkin, E. Norton, A. Gafni, and B. O'Brien (1998). Valuing outcomes in health care: a comparison of willingness to pay and quality-adjusted life-years. *J Clin Epidemiol 51*, 667–76.

Bateman, I. and K. Willis (Eds.) (1999). *Valuing environmental preferences: theory and practice* *of the contingent valuation method in the US, EU, and developing countries*. New York: Oxford University Press.

Beattie, J., J. Covey, P. Dolan, L. Hopkins, M. Jones-Lee, G. Loomes, N. Pidgeon, A. Robinson, and A. Spencer (1998). On the contingent valuation of safety and the safety of contingent valuation: Part 1- caveat investigator. *J Risk Uncertain 17*, 5–25.

Bishop, R. and T. Heberlein (1979). Measuring values of extra-market goods: are indirect measures biased? *Amer. J. Agr. Econ. 61*, 926–930.

Blumenschein, K., M. Johannesson, K. Yokoyama, and P. Freeman (2001). Hypothetical versus real willingness to pay in the health care sector: results from a field experiment. *J Health Econ 20*, 441–57.

Boyle, K., R. Bishop, and M. Welsh (1985). Starting point bias in contingent valuation bidding games. *Land Economics 61*, 188–194.

Boyle, K., F. Johnson, and D. McCollum (1997). Anchoring and adjustment in single-bounded, contingent-valuation questions. *Amer J Agr Econ 79*, 1495–1500.

Boyle, K., H. MacDonald, H. Cheng, and D. McCollum (1998). Bid design and yea saying in single-bounded, dichotomous-choice questions. *Land Economics 74*(1), 49–64.

Brazier, J., R. Akehurst, A. Brennan, P. Dolan, K. Claxton, C. McCabe, M. Sculpher, and A. Tsuchyia (2005). Should patients have a greater role in valuing health states? *Appl Health Econ Health Policy 4*(4), 201–8.

Cameron, T. and D. Hupper (1991). Referendum contingent valuation estimates: Sensitivity to the assignment of offered values. *Journal of the American Statistical Association 84*(416), 910–918.

Cameron, T. and J. Quiggin (1994). Estimation using contingent valuation data from "dichotomous choice with follow-up" questionnaire. *Journal of Environmental Economics and Management 27*, 218–234.

Carrat, F., C. Sahler, C. Le Galès, M. Schwarzinger, M. Nicolas, and M. Bungener (2001). How much patients are willing to pay for an earlier alleviation of influenza ? In A. Osterhaus, N. Cox, and A. Hampson (Eds.), *Options for the control of influenza IV*, pp. 67–72. Elsevier Science.

Carrat, F., C. Sahler, S. Rogez, M. Leruez-Ville, F. Freymuth, C. Le Gales, M. Bungener, B. Housset, M. Nicolas, and C. Rouzioux (2002). Influenza burden of illness: estimates from a national prospective survey of household contacts in france. *Arch Intern Med 162*, 1842–8.

Carson, R. and T. Groves (2007). Incentive and informational properties of preference questions. *Environmental and Resource Economics 37*(1), 181–210.

Carthy, T., S. Chilton, J. Covey, L. Hopkins, M. Jones-Lee, G. Loomes, N. Pidgeon, and A. Spencer (1998). On the contingent valuation of safety and the safety of contingent valuation: Part 2- the cv/sg "chained" approach. *J Risk Uncertain 17*, 187–213.

Cassileth, B., E. Lusk, T. Strouse, D. Miller, L. Brown, P. Cross, and A. Tenaglia (1984). Psychosocial status in chronic illness. a comparative analysis of six diagnostic groups. *N England J Med 311*(8), 506–511.

Chapman, G. and E. Johnson (1994). The limits of anchoring. *J Behavioral Decision Making 7*, 232–42.

Chiu, L., K. Tang, Y. Liu, W. Shyu, and T. Chang (1998). Willingness of families caring for victims of dementia to pay for nursing home care: results of a pilot study in taiwan. *J Manag Med 12*(6), 349–60.

Clarke, P. (2000). Valuing the benefits of mobile mammographic screening units using the contingent valuation method. *Appl Econ 32*, 1647–55.

Cooper, J. and J. Loomis (1992). Sensitivity of willingness-to-pay estimates to bid design in dichotomous choice contingent valuation models. *Land Economics 68*(2), 211–24.

Cooper, N., A. Sutton, K. Abrams, A. Wailoo, D. Turner, and K. Nicholson (2003). Effectiveness of neuraminidase inhibitors in treatment and prevention of influenza a and b: systematic review and meta-analyses of randomised controlled trials. *BMJ 326*(7401), 1235–41.

Dalmau-Matarrodona, E. (2001). Alternative approaches to obtain optimal bid values in contingent valuation studies and to model protest zeros. estimating the determinants of individuals' willingness to pay for home care services in day case surgery. *Health Econ 10*(2), 101–18.

Damschroder, L., B. Zikmund-Fisher, and P. Ubel (2005). The impact of considering adaptation in health state valuation. *Social Sciences & Medecine 61*(2), 267–277.

Dawson, N., M. Singer, L. Lenert, M. Patterson, S. Sami, I. Gonsenhouser, H. Lindstrom, K. Smyth, M. Barber, and P. Whitehouse (2008). Health state valuation in mild to moderate cognitive impairment: feasibility of computer-based, direct patient utility assessment. *Med Decis Making 28*(2), 220–32.

Dolan, P. (1999). Whose preferences count? *Med Decis Making 19*(4), 482–6.

Flachaire, E., G. Hollard, and S. Luchini (2007). Heterogeneous anchoring in dichotomous choice valuation framework. *Louvain Economic Review 73*(4), 369–384.

Ford, I., B. Torsney, and C. Wu (1992). The use of a canonical form in the construction of locally optimal designs for non-linear problems. *Journal of the Royal Statistical Society 54*(2), 569–583. Series B (Methodological).

Foreit, J. and K. Foreit (2003). The reliability and validity of willingness to pay surveys for reproductive health pricing decisions in developing countries. *Health Policy 63*, 37–47.

Frew, E., D. Whynes, and J. Wolstenholme (2003). Eliciting willingness to pay: comparing closed-ended with open-ended and payment scale formats. *Med Decis Making 23*, 150–9.

Frykblom, P. and J. Shogren (2000). An experimental testing of anchoring effects in discrete choice questions. *Environmental and Resource Economics 16*, 329–41.

Green, D., K. E. Jacowitz, D. Kahneman, and D. McFadden (1998). Referendum contingent valuation, anchoring, and willingness to pay for public goods. *Resource and Energy Economics 20*, 85–116.

Haab, T. (1998). Estimation using contingent valuation data from a "dichotomous choice with follow-up" questionnaire: A comment. *Journal of environmental economics and management 35*, 190–94.

Hanemann, M., J. Loomis, and B. Kanninen (1991). Statistical efficiency of double-bounded dichotomous choice contingent valuation. *Amer. J. Agr. Econ. 73*(4), 1255–63.

Hanemann, M. W. (1985). Some issues in continuous and discrete response contingent valuation studies. *Northeast Journal of Agricultural Economics 14*, 5–13.

Hanemann, M. W. and B. J. Kanninen (1999). *Statistical Analysis of Discrete Response CV Data, in « Valuing Environmental Preferences - Theory and Practice of the Contingent Valuation Method in the US, EC and Developing Countries »*. Oxford: Ian Bateman and Ken G. Willis.

Hanley, N., M. Ryan, and R. Wright (2003). Estimating the monetary value of health care: lessons from environmental economics. *Health Econ 12*, 3–16.

Herriges, J. A. and J. F. Shogren (1996). Starting point bias in dichotomous choice valuation with follow-up questioning. *J Environ Economics and Management 30*, 112–131.

Holmes, T. and K. Boyle (2005). Dynamic learning and context-dependence in sequential, attribute-based, stated-preference valuation questions. *Land Economics 81*, 114–126.

Jones-Lee, M. (1989). *The economics of safety and physical risk.* Oxford: Basil Blackwell.

Jones-Lee, M. (1992a). Altruism and the value of other people's safety. *Journal of Risk and Uncertainty 4*, 213–219.

Jones-Lee, M. (1992b). Paternalistic altruism and the value of statistical life. *The Economic Journal 102*, 80–90.

Kang, H. (2006). *Essays on methodologies in contingent valuation and the sustainable management of common pool resources.* Ph. D. thesis, Ohio State University. Timothy Haab advisor.

Kanninen, B. (1995a). Bias in discrete response contingent valuation. *J Environ Economics Management 28*, 114–25.

Kanninen, B. (1995b). Design of sequential experiments for contingent valuation studies. *Journal of Environmental Economics and Management 25*, S–1–11.

Kennedy, C. (2002). Revealed preference valuation compared to contingent valuation: radon-induced lung cancer prevention. *Health Econ 11*, 585–98.

Lenert, L., J. Ziegler, T. Lee, R. Sommi, and R. Mahmoud (2000). Differences in health values among patients, family members, and providers for outcomes in schizophrenia. *Med Care 38*(10), 1011–21.

Liljas, B. and K. Blumenschein (2000). On hypothetical bias and calibration in cost-benefit studies. *health Policy 52*, 53–70.

Liu, J., J. Hammitt, J. Wang, and J. Liu (2000). Mother's willingness to pay for her own and her child's health: a contingent valuation study in taiwan. *Health Economics 9*(4), 319–26.

McLeod, D. and O. Bergland (1999). Willingness-to-pay estimates using the double-bounded dichotomous-choice contingent valuation format: a test for validity and precision in a bayesian framework. *Land Economics 75*, 115–125.

Minkin, S. (1987). Optimal designs for binary data. *Journal of the American Statistical Association 82*(400), 1098–1103.

Mulvaney-Day, N. (2005). Using willingness to pay to measure family members' preferences in mental health. *J Ment Health Policy Econ 8*, 71–81.

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review 110*, 472–89.

Nyquist, H. (1992). Optimal designs of discrete response experiments in contingent valuation studies. *The Review of Economics and Statistics 74*(3), 559–563.

Olsen, J. and R. Smith (2001). Theory versus practice: a review of 'willingness-to-pay' in health and health care. *Health Econ 10*, 39–52.

Onwujekwe, O., R. Chima, E. Shu, D. Nwagbo, and P. Okonkwo (2001). Hypothetical and actual willingness to pay for insecticide-treated nets in five Nigerian communities. *Trop Med Int Health 6*, 545–53.

Sach, T., D. Whynes, C. O'Neill, G. O'Donoghue, and S. Archbold (2004). Willingness-to-pay for pediatric cochlear implantation. *Int J Pediatr Otorhinolaryngol 68*(1), 91–9.

Saigal, S., B. Stoskopf, D. Feeny, W. Furlong, E. Burrows, P. Rosenbaum, and L. Hoult (1999).

Differences in preferences for neonatal outcomes among health care professionals, parents, and adolescents. *JAMA 281*(21), 1991–7.

Scarpa, R. and I. Bateman (2000). Efficiency gains afforded by improved bid design versus follow-up valuation questions in discrete-choice cv studies. *Land Economics 76*, 299–311.

Schwarzinger, M., K. Lacombe, and F. Carrat (2003). Review of economic evaluations of neuraminidase inhibitors to control influenza. *Expert Rev Pharmacoeconomics Res 3*, 89–101.

Sitter, R. (1992). Robust designs for binary data. *Biometrics 48*(4), 1145–1155.

Slothuus, U. and R. Brooks (2000). Willingness to pay in arthritis: a danish contribution. *Rheumatology 39*, 791–9.

Smith, A., M. Thomas, P. Brockman, J. Kent, and K. Nicholson (1993). Effects of influenza b virus infection on human performance. *BMJ 306*, 760–1.

Smith, R. (2000). The discrete-choice willingness-to-pay question format in health economics: should we adopt environmental guidelines? *Med Decis Making 20*, 194–206.

Smith, R. (2003). Construction of the contingent valuation market in health care: a critical assessment. *Health Econ 12*, 609–28.

Sneeuw, K., N. Aaronson, D. Osoba, M. Muller, M.-A. Hsu, W. Yung, M. Brada, and E. Newlands (1997). The use of significant others as proxy raters of the quality of life of patients with brain cancer. *Med Care 35*(5), 490–506.

Sorum, P. (1999). Measuring patient preferences by willingness to pay to avoid: the case of acute otitis media. *Med Decis Making 19*(1), 27–37.

Tamim, H., J. McCusker, and N. Dendukuri (2002). Proxy reporting of quality of life using the eq-5d. *Med Care 40*(12), 1186–95.

Todorov, A. and C. Kirchner (2000). Bias in proxies' reports of disability: data from the national health interview survey on disability. *Am J Public Health 90*, 1248–1253.

Tsevat, J., N. Dawson, A. Wu, J. Lynn, J. Soukup, E. Cook, H. Vidaillet, and R. Phillips (1998). Health values of hospitalized patients 80 years or older. help investigators. hospitalized elderly longitudinal project. *JAMA 279*(5), 371–5.

Tversky, A. and D. Kahneman (1974). Judgments under uncertainty: heuristics and biases. *Science 185*, 1124–31.

Watson, V. and M. Ryan (2007). Exploring preference anomalies in double bounded contingent valuation. *J Health Econ 26*(3), 463–482.

Whitehead, J. (2002). Incentive incompatibility and starting-point bias in iterative valuation questions. *Land Economics 78*, 285–297.

Yatchew, A. and Z. Griliches (1984). Specification error in probit models. *Review of Economics and Statistics 66*, 134–39.

A Description of the willingness-to-pay scenario with a patient-based perspective

We would like to measure the importance for you of earlier alleviation of influenza-like symptoms. Please, assume that a new treatment could alleviate influenza-like symptoms by 1 (3) day(s). The symptoms will last 4 to 5 days (2 to 3 days) with the new treatment instead of 5 to 6 days without the new treatment. The new treatment should be taken in addition to usual flu treatments and will not be reimbursed by health insurances, either public or private. In your opinion, would you be willing to pay out-of-your pocket 100 FF[6] *to alleviate your*[7] *symptoms by 1 (3) day(s) if you had flu again?*

- If the respondent answered *Yes*, then a follow-up question was proposed: *In your opinion, would you be willing to pay 200 FF*[8] *to alleviate your symptoms by 1 (3) day(s) if you had flu again?*
- If the respondent answered *No*, then a follow-up question was proposed: *In your opinion, would you be willing to pay 50 FF*[5] *to alleviate your symptoms by 1 (3) day(s) if you had flu again?*
- If the respondent would not like to answer, then a debriefing question was proposed.

[6] *The initial bid was randomly selected in [25; 50; 100; 200; 400] FF or [50; 100; 200; 400; 800] FF for the alleviation of flu symptoms by one day or three days, respectively (1 Euro= 6.56 French Francs).*

[7] *In the proxy-based perspective, the question was framed as follows: 'In your opinion, would you be willing to pay out-of-your pocket 100 FF* to alleviate by 1 (3) day(s) the symptoms of your partner if she(he) had flu again?'*

[8] *The follow-up bid was doubled if respondent answered Yes to the initial bid, and halved otherwise.*

B Characteristics of patients, households, and willingness-to-pay survey. Except where stated otherwise, values are expressed in percentage of subjects ($n = 287$)

Name of variables	Description of variables	% or Mean (sd.)
Patient characteristics		
MALE	Gender male	51.9
AGE	Age, mean (sd.) years	40.9 (15.5)
FLU_SCORE	Clinical score from 0=no symptoms to 1=severe symptoms, mean (sd.) score	0.41 (0.16)
FLU_DIAG	Influenza infection confirmed by virological tests	62.0
ATB	Consumed antibiotic drugs	48.8
FLU_MEDCOSTS	Medical costs of flu, mean (sd.) (€)	44.6 (28.4)
FLU_BEFORE	Had ILI the previous year	14.3
FLU_VACC	Vaccinated this year	10.8
CHRONICDIS	Has a chronic disease	22.3
SMOKER	Current smoker	26.1
Household characteristics		
HOUSE_SIZE	Number of household members	3.1 (1.2)
HOUSE_INC1	Household income $<$ €1,524	31.0
HOUSE_INC2	Household income: €1,524 to €3,049	47.7
HOUSE_INC3	Household income $>$ €3,049	21.3
Willingness-to-pay survey characteristics		
INT_OTHERFLU	Another household member has ILI at time of interview	24.4
INT_PATIENT	Respondent is the patient	77.7
INT_PROXY	Respondent is the spouse of the patient	22.3
INT_3DAYS	Willingness-to-pay for earlier alleviation of flu symptoms by 3 days	47.4

ILI: Influenza-like illness
Survey was conducted in 1999 and FF were converted to euros (1 FF= €0.152449)

C Heterogeneous anchoring LR tests

Variable	LR test	p-value
FEMALE	1.072	0.300
AGE	0.945	0.330
FLU_SCORE	3.581	0.058
FLU_DIAG	1.497	0.221
ATB	0.002	0.960
FLU_MEDCOSTS	0.094	0.758
FLU_BEFORE	1.888	0.169
FLU_VACC	0.001	0.968
CHRONICDIS	0.079	0.778
SMOKER	0.018	0.891
HOUSE_SIZE	1.942	0.163
HOUSE_INC2	0.373	0.540
HOUSE_INC3	0.008	0.927
INT_OTHERFLU	6.179	0.012
INT_3DAYS	0.757	0.383

Tables

Table 1.a: Bid acceptance rate for the earlier alleviation of flu symptoms by one day ($n = 151$)

	Respondent is the patient ($n = 113$)						Respondent answering on behalf of the patient ($n = 38$)					
Follow-up bid	Initial bid						Initial bid					
	3.8 ($n = 23$)	7.6 ($n = 24$)	15.2 ($n = 24$)	30.5 ($n = 21$)	61.0 ($n = 21$)	Fisher p-value	3.8 ($n = 5$)	7.6 ($n = 7$)	15.2 ($n = 6$)	30.5 ($n = 13$)	61.0 ($n = 7$)	Fisher p-value
1.9	91%						100%					
3.8	91%	92%				ns	100%	86%				ns
7.6	78%	88%	79%			ns	100%	86%	100%			ns
15.2		63%	58%	76%		ns		71%	100%	69%		ns
30.5			38%	38%	71%	<.05			33%	46%	71%	ns
61.0				0%	38%	<.01				23%	29%	ns
122.0					10%						14%	

Note: Bids were presented in FF at time of the survey and are converted to Euros in the Table (1FF=€0.152449)

Table 1.b: Bid acceptance rate for the earlier alleviation of flu symptoms by three day ($n = 136$)

	Respondent is the patient ($n = 110$)						Respondent answering on behalf of the patient ($n = 26$)					
Follow-up bid	Initial bid						Initial bid					
	7.6 ($n = 26$)	15.2 ($n = 18$)	30.5 ($n = 25$)	61.0 ($n = 16$)	122.0 ($n = 25$)	Fisher p-value	7.6 ($n = 4$)	15.2 ($n = 5$)	30.5 ($n = 6$)	61.0 ($n = 5$)	122.0 ($n = 6$)	Fisher p-value
3.8	100%						75%					
7.6	88%	94%				ns	75%	100%				ns
15.2	65%	72%	92%			<.1	75%	100%	67%			ns
30.5		56%	76%	81%		ns		60%	50%	80%		ns
61.0			16%	50%	68%	<.001			17%	80%	17%	<.1
122.0				6%	24%	ns				60%	0%	<.1
243.9					0%						0%	

Note: Bids were presented in FF at time of the survey and are converted to Euros in the Table (1FF=€0.152449)

Table 2: Econometric results ($n = 287$)

Variables	Single-Bounded	Double-bounded		
	reference	naïve	shift &Constant Anchoring	shift &Heterogeneous Anchoring
	Parameter (p-value)	Parameter (p-value)	Parameter (p-value)	Parameter (p-value)
CONSTANT	-7.986	16.626	9.421	9.473
	(0.767)	(0.328)	(0.685)	(0.678)
FEMALE	5.917	1.063	1.445	3.534
	(0.520)	(0.846)	(0.845)	(0.656)
AGE	0.424	0.301	0.496	0.408
	(0.223)	(0.184)	(0.115)	(0.176)
FLU_SCORE	-4.995	0.097	-4.812	-8.185
	(0.864)	(0.995)	(0.836)	(0.748)
FLU_DIAG	10.088	0.141	2.728	6.577
	(0.238)	(0.979)	(0.718)	(0.443)
ATB	-10.400	-1.341	-2.228	-6.600
	(0.233)	(0.806)	(0.762)	(0.415)
FLU_BEFORE	-8.408	-12.397	-15.467	-21.611*
	(0.468)	(0.101)	(0.134)	(0.080)
FLU_VAC	-2.092	8.575	7.251	3.008
	(0.899)	(0.355)	(0.569)	(0.825)
CHRONICDIS	-13.294	-11.376	-13.706	-11.254
	(0.264)	(0.109)	(0.152)	(0.294)
SMOKER	17.649*	10.774*	13.698	9.547
	(0.099)	(0.082)	(0.103)	(0.310)
HOUSE_SIZE	7.612*	1.579	3.022	4.915
	(0.089)	(0.536)	(0.385)	(0.234)
HOUSE_INC2	15.199	2.325	3.669	4.463
	(0.157)	(0.711)	(0.666)	(0.648)
HOUSE_INC3	19.244	12.756	14.765	10.562
	(0.143)	(0.114)	(0.175)	(0.365)
HOUSE_MECOSTS	0.065	-0.009	0.025	0.074
	(0.715)	(0.920)	(0.838)	(0.576)
INT_3DAYS	17.248**	18.060***	13.407*	12.968*
	(0.047)	(0.000)	(0.071)	(0.098)
INT_OTHERFLU	3.080	-2.514	-8.120	-6.234
	(0.767)	(0.680)	(0.334)	(0.504)
STRUCTURAL SHIFT				
Shift parameter δ	-	-	-0.791	-0.673
	-	-	(0.391)	(0.484)
ANCHORING				
Constant Anchoring γ	-	-	0.421***	-
			(<0.001)	
Heterogeneous anchoring				
INT_PATIENT	-	-	-	0.563***
	-	-	-	(<0.001)
INT_OTHERFLU × INT_PATIENT	-	-	-	-.230***
	-	-	-	(<0.029)
PROXY-PERSPECTIVE				
INT_PROXY (θ)	-11.741	3.569	6.583	-2.454
	(0.246)	(0.648)	(0.591)	(0.786)
PRIOR PREDICTED WTP ($\widehat{W}_{i1}$)				
Mean, €	58.85	44.71	48.09	50.51
(sd.)	(1.23)	(0.79)	(0.89)	(0.93)
Dispersion, €				
Patient (σ)	63.81	35.55	47.34	55.14
(sd.)	(13.19)	(2.62)	(5.67)	(7.24)
Proxy (ω)	25.578	45.807	76.086	41.962
(sd.)	(7.769)	(7.430)	(15.869)	(5.768)
Log-likelihood	-148.127	-363.588	-348.613	-342.059

if *** if p-value<0.01, **if p-value<0.05, *if p-value<0.1

Figure

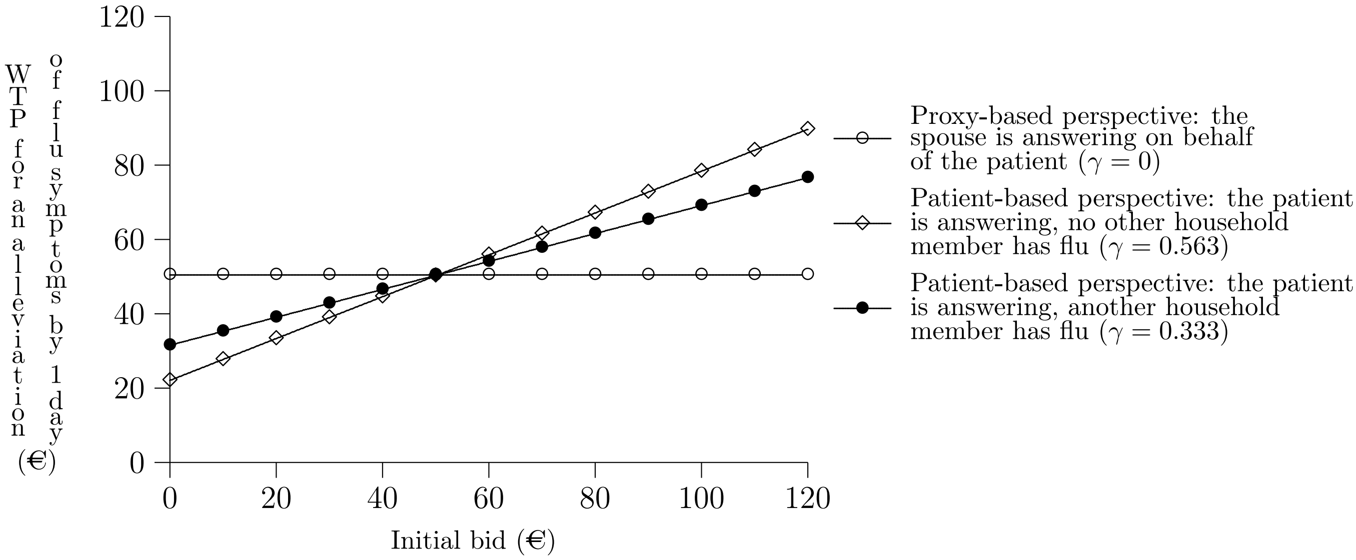

Figure 1: WTP for the alleviation of flu symptoms by 1 day according to the level of anchoring of the respondent and the initial bid proposed in the double bounded model with heterogeneous anchoring.