

NBER WORKING PAPER SERIES

ON INFERRING DEMAND FOR HEALTH CARE IN THE PRESENCE OF ANCHORING,  
ACQUIESCENCE, AND SELECTION BIASES

Jay Bhattacharya  
Adam Isen

Working Paper 13865  
<http://www.nber.org/papers/w13865>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
March 2008

Bhattacharya acknowledges the National Institute on Aging for financial support. Isen acknowledges the Stanford Undergraduate Research Program, the Stanford Graduate Student Council, and the Associated Students of Stanford University for financial and other support for the project. We thank William Vogt for helpful discussions. All errors remain our own. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2008 by Jay Bhattacharya and Adam Isen. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

On Inferring Demand for Health Care in the Presence of Anchoring, Acquiescence, and Selection Biases

Jay Bhattacharya and Adam Isen  
NBER Working Paper No. 13865  
March 2008  
JEL No. C42,C81,C9,I1

**ABSTRACT**

In the contingent valuation literature, both anchoring and acquiescence biases pose problems when using an iterative bidding game to infer willingness to pay. Anchoring bias occurs when the willingness to pay estimate is sensitive to the initially presented starting value. Acquiescence bias occurs when survey respondents exhibit a tendency to answer 'yes' to questions, regardless of their true preferences. More generally, whenever a survey format is used and not all of those contacted participate, selection bias raises concerns about the representativeness of the sample.

In this paper, we estimate students' willingness to pay for student health care at Stanford University while accounting for all of these biases. As there is no cost sharing for students, we assess willingness to pay by having a random sample of students play an online iterative bidding game. Our main results are that (1) demand for student health care is elastic by conventional standards; (2) ignoring anchoring bias would lead to a substantially biased measure of the demand elasticity; (3) there is evidence for acquiescence bias in student answers to the opening question of the iterative bidding game and failure to address this leads to the biased conclusion that demand is inelastic; and (4) standard selection correction methods indicate no bias from selective non-response and newer bounding methods support this conclusion of elastic demand.

Jay Bhattacharya  
117 Encina Commons  
Center for Primary Care  
and Outcomes Research  
Stanford University  
Stanford, CA 94305-6019  
and NBER  
jay@stanford.edu

Adam Isen  
University of Pennsylvania  
Steinberg-Dietrich Hall #1400  
3620 Locust Walk  
Philadelphia, PA 19104-6372  
isen@wharton.upenn.edu

# 1 Introduction

In the contingent valuation literature, both anchoring and acquiescence biases pose problems when using an iterative bidding game to infer willingness to pay. Anchoring bias occurs when the willingness to pay estimate is sensitive to the initially presented starting value. Acquiescence bias occurs when survey respondents exhibit a tendency to answer ‘yes’ to questions, regardless of their true preferences. More generally, whenever a survey format is used and not all of those contacted participate, selection bias raises concerns about the representativeness of the sample. It is difficult to adjust simultaneously for all these sources of bias when analyzing survey data to infer willingness to pay from an iterative bidding game. In this paper, we demonstrate a simple methodology to do so.

On most college campuses across the country, student health centers provide care to students at no cost at the point of service. At Stanford University specifically, all students are automatically enrolled in a first tier health care plan that provides services through the student health center (Vaden Health Center) with no cost sharing by students.<sup>1</sup> There are approximately 25,000 visits to Vaden annually, and over 67 percent of the roughly 12,000 enrolled students make such visits each year. Optimal coinsurance, which balances the marginal benefit of risk pooling with the marginal cost of moral hazard (Pauly, 1968), depends in part upon the price elasticity of demand for health care by students as well as the magnitude of the demand. Thus, determining this elasticity and willingness to pay are important to determining whether the imposition of cost-sharing would improve social welfare.

The lack of cost sharing by students prevents us from measuring the demand elasticity using standard revealed preference methods. We thus turn to contingent valuation methods. We conduct a representative online survey of Stanford students in which students play an iterative bidding game designed to reveal their willingness to pay for visits to Vaden. In this game, respondents are asked whether they would pay a specified amount for the good or service at hand. If respondents answer in the affirmative, the amount is incrementally increased until their willingness to pay is reached. Similarly, if respondents answer in the negative, the amount is incrementally decreased until their maximum willing to pay is reached.

---

<sup>1</sup>First tier health care consists of care given from nurses and primary care physicians at the student health center. Referrals to specialists and the provision of prescription drugs are not considered part of first tier care.

Contingent valuation methods, such as the one we employ, are subject to well known biases. Two such biases are due to anchoring (when the starting bracket amount influences the willingness to pay estimate) and acquiescence (the tendency to answer ‘yes’ to a question when it is thought to be the socially desirable response).

If anchoring is present, the higher the starting bid, the higher the stated willingness to pay estimates provided by respondents. Many studies in the health economics literature have used contingent valuation methods to measure demand elasticities without accounting for anchoring bias. One exception is Smith (2006) who experimented with payment cards with various starting brackets and found that the willingness to pay estimate is sensitive to the starting bracket. Two studies have explicitly tested for anchoring bias in the context of an iterative bidding game, with one finding an effect and the other no effect (Stalhammar, 1996; OBrien and Viramontes, 1994), and since their main aim was to test for the bias, neither study corrects for the bias in its measurement of willingness to pay. In this paper, we adapt an econometric model developed by Herriges and Shogren (1996) in the environmental economics literature to correct our elasticity estimate for anchoring bias.

Classically, acquiescence bias occurs when there is a social pressure to agree to a particular survey question (for example, “Is racism bad?”). However, such tendencies may manifest themselves even in situations in which social pressure is doubtful (see Hurd and van Soest, 2003). Both Ryan et al. (2004) and Blumenschein et al. (2001) argue that acquiescence bias plays a large role in explaining high willingness to pay estimates in contingent valuation studies of the value of health care.

Last, selection bias might manifest itself when those who respond to the survey differ from the whole population in salient ways. In our data collection, we provided randomized participation incentives to survey participants. We take advantage of this randomization to adjust our estimates for non-response bias using both standard and newer methods from the econometrics literature.

## 2 Data

In early 2006, we conducted a survey of the Stanford University student body. Stanford provided us with a randomly selected list of two thousand

undergraduates and two thousand graduate students. We elicited student participation through email contact (up to three times if students did not respond initially). In order to increase the response rate, we stressed the support of the Stanford student government organization for this study and we provided a financial inducement. Four hundred randomly chosen students who filled out the survey were given a \$5 Amazon gift certificate. The rest of the students were offered no compensation for participating in the study. The typical time to complete the survey was under five minutes. Overall, 30.8% of students responded to the survey.

The survey consisted of a part that included the iterative bidding game and a part asking about demographic and health background information. The iterative bidding game was structured to allow for tests of anchoring and acquiescence biases. First, respondents were randomly assigned to one of four bidding games, each of which had a different starting bid (\$5, \$20, \$35, and \$50). This strategy is called random bracket entry. We picked these bid levels, based on pilot runs of the survey, to be within a typical student’s budget constraint.

Second, the questions used neutral wording that avoided any “yes” or “no” questions and instead asked what the respondent would have done if an  $x$  dollar co-payment was charged. Hurd and van Soest (2003), using national consumption and income surveys, find that acquiescence bias can be avoided by slightly reformulating the question to more neutral wording. To explicitly test if the neutral wording made a difference, a separate random sample of students was assigned to bidding games that contained purely ‘yes’ or ‘no’ questions.

The remaining questions from the survey were used to address three main concerns raised by cost sharing: how a co-payment would generally affect health outcomes and how a co-payment would specifically affect low-income students and those with poor health backgrounds. We included questions about the student’s reason for his or her last visit, how long the reason for the visit existed, and how the student ranked the urgency of his or her reason for visiting. To address questions regarding differential willingness to pay by low-income students and those in poor health, we included socio-economic status and health background questions in the survey.

Table 1 shows descriptive statistics about about the demographics of the study population of 883 students. Stanford students tend come from families with high levels of educational attainment. Family income levels (collected only for graduate students) are, not surprisingly for a collection of graduate

students, low. A substantial proportion of the student population is on financial aid.

Table 2 shows descriptive statistics about the health status of the study population and their use of student health services. Students typically report themselves to be in very good or excellent health. Few students have a chronic medical condition. When they do use student health services, they tend to arrive with non-urgent complaints.<sup>2</sup>

Average reported willingness to pay for a visit (taken as the midpoint of the final interval in the iterative bidding game played each respondent) increases sharply with the randomly assigned starting bracket. Ignoring anchoring would thus lead, almost certainly, to a biased estimate of willingness to pay. We also see evidence of acquiescence bias as given by a higher percent of students posed the non-neutral wording responding in the affirmative to the first bracket question.

### 3 Anchoring Bias

We adapt a Bayesian learning model, developed by Herriges and Shogren (1996) to measure willingness to pay in the environmental economics context, to our problem. The intuition behind the learning model is that people do not have a point willingness to pay for a given good before they are asked about it in the survey. Rather, each person has a distribution over their willingness to pay that represents their beliefs about it. These beliefs are based on their past experiences with the specific good and related goods. Once respondents are exposed to the first bracket, the information the respondents held prior to viewing the bracket is then updated by the information in the first bracket, which signals to them the value of the “correct” willingness to pay. As a result, rather than comparing their “true” willingness to pay to all the subsequent brackets, respondents compare their “true” willingness to pay with the new information provided by the first bracket to form the “revised” willingness to pay, which is then compared to the subsequent bracket.

More formally, let  $W$  denote the “true” estimate of each respondent’s willingness to pay. In playing the iterative bidding game, the respondents are asked whether would have been willing to pay  $b_1$  to visit the health

---

<sup>2</sup>We measured visit urgency on a scale from 1 to 5. One corresponds to a benign cold, three corresponds to infectious mononucleosis, and five corresponds to an emergency room visit.

center. The respondent’s answer then reveals whether their willingness to pay estimate lies above or below  $b_1$ , the entry point into the game. If the respondent answers no, he is asked if he would be willing to pay  $b_2 < b_1$ . Alternatively, if the respondent answers yes, he is asked if he would be willing to pay  $b_2 > b_1$ . This pattern of questioning is repeated until bracket  $b_i$  is reached such that exactly one of the following is true:

$$\text{If } W < b_1 \text{ then } b_{i-1} > W > b_i \text{ and } b_1 > b_2 > \dots > b_{i-1} > b_i \quad (1)$$

$$\text{If } W > b_1 \text{ then } b_{i-1} < W < b_i \text{ and } b_1 < b_2 < \dots < b_{i-1} < b_i \quad (2)$$

The end result is an interval between  $b_i$  and  $b_{i-1}$  in which each respondent’s willingness to pay lies. However, this approach ignores any anchoring effect that may be influencing respondents’ stated willingness to pay. Instead, the observed willingness to pay can be thought of as a function of one’s true willingness to pay and the first entry point,  $b_1$ .

Let  $W$  be the respondent’s “true” willingness to pay estimate; let  $\tilde{W}$  be the reported willingness to pay estimate that is formed after the respondent is shown  $b_1$ , and let  $\alpha \in [0, 1]$  be a constant that mediates the extent of anchoring bias. Herriges and Shogren (1996) propose the following arithmetic model for the anchoring bias:

$$\tilde{W} = (1 - \alpha)W + \alpha b_1 \quad (3)$$

Instead, we model the anchoring effect as a geometric average of  $b_1$  and  $W$ .<sup>3</sup>

$$\ln \tilde{W} = (1 - \alpha) \ln W + \alpha \ln b_1 \quad (4)$$

In this model, reported willingness to pay is a weighted geometric average of the true willingness to pay and the entry value. As  $\alpha$  increases, so does the extent of anchoring bias. At  $\alpha = 1$ , there is complete anchoring so that a respondent will report  $\tilde{W} = b_1$  regardless of his true willingness to pay. On the other hand, at  $\alpha = 0$ , there is no anchoring bias and  $\tilde{W} = W$ .

According to models (3) and (4), when the first bracket question is asked, a respondent will compare his “true” willingness to pay with  $b_1$ . In answering

---

<sup>3</sup>The main advantage of a geometric model over an arithmetic form is that a geometric model restricts  $W$  to nonnegative values whereas an algebraic model could lower the “true” willingness to pay estimate to a negative value, which makes little sense in our context.

this first question, the respondent will be biased toward  $b_1$  in his response. However, the reported answer will be on the same side as  $W$ , the “true” willingness to pay. Thus, the answer to the first question can be taken at face value. On the other hand, all subsequent questions cannot be taken at face value in that they are comparing the bracket amount  $b_i$  to the reported willingness to pay estimate  $\tilde{W}$ .

The survey data from the iterative bidding game come in the form of intervals around the reported willingness to pay:

$$\begin{aligned} b_{i-1} > \tilde{W} > b_i & \text{ if (1) holds, or} \\ b_{i-1} < \tilde{W} < b_i & \text{ if (2) holds.} \end{aligned}$$

However, we are interested in  $W$ , not  $\tilde{W}$ . Applying (4), we rewrite the above equations as follows:

$$\begin{aligned} b_{i-1} > \exp((1 - \alpha) \ln W + \alpha \ln b_1) > b_i & \text{ if (1) holds.} \\ b_{i-1} < \exp((1 - \alpha) \ln W + \alpha \ln b_1) < b_i & \text{ if (2) holds.} \end{aligned}$$

We use maximum likelihood methods to estimate the model. We assume that the “true” willingness to pay for health care follows a three parameter Gamma distribution. We assume that willingness to pay comes from an independent draw of a random variable,  $W$ , with the following cumulative distribution (which can be expressed in terms of the lower incomplete gamma function  $\gamma$ ):<sup>4</sup>

$$\begin{aligned} P(W < w) &= \frac{1}{\Gamma(k)} \int_0^{(w-g)/b} e^{-t} t^{k-1} dt \\ &= \frac{\gamma(k, (w-g)/b)}{\Gamma(k)} \text{ for } w \geq 0, w \geq g, \text{ and } k, b, g > 0. \end{aligned} \tag{5}$$

With (5) and the data, we can calculate the probability that we observe the sample that we actually observe. If (1) holds:

---

<sup>4</sup>Here  $k$  is the shape parameter,  $b$  is the scale parameter, and  $g$  is the location parameter



$$P\left(b_i < \tilde{W} < b_{i-1}\right) = \frac{1}{\Gamma(k)} \begin{pmatrix} \gamma(k, (\exp\left(\frac{\ln b_{i-1} - \alpha \ln b_1}{1-\alpha}\right) - g)/b) \\ -\gamma(k, (\exp\left(\frac{\ln b_i - \alpha \ln b_1}{1-\alpha}\right) - g)/b) \end{pmatrix}.$$

If (2) holds:

$$P\left(b_{i-1} < \tilde{W} < b_i\right) = \frac{1}{\Gamma(k)} \begin{pmatrix} \gamma(k, (\exp\left(\frac{\ln b_i - \alpha \ln b_1}{1-\alpha}\right) - g)/b) \\ -\gamma(k, (\exp\left(\frac{\ln b_{i-1} - \alpha \ln b_1}{1-\alpha}\right) - g)/b) \end{pmatrix}.$$

Let  $A_1$  represent the set of respondents for whom (1) holds (true willingness to pay is below  $b_1$ ) and let  $A_2$  represent the set of respondents for whom (2) holds. Then the log likelihood function is:

$$\ln L = \sum_{n \in A_1} \ln P\left(b_{n,i-1} > \tilde{W}_n > b_{n,i}\right) + \sum_{n \in A_2} \ln P\left(b_{n,i-1} < \tilde{W}_n < b_{n,i}\right) \quad (6)$$

Our goal is to estimate  $\alpha$ ,  $k$ ,  $b$ , and  $g$ . We are also interested in how some important characteristics of each respondent,  $X$ , affect willingness to pay as well and the anchoring effect. The  $X$  vector includes demographic characteristics (such as race, income, and graduate student status) as well as health status characteristics.<sup>5</sup> We parameterize the dependence of  $k$ ,  $b$  and  $\alpha$  on  $X$  as follows:  $k = \left(\frac{1}{\exp(X'\kappa)}\right)$ ,  $b = \left(\frac{1}{\exp(X'\beta)}\right)$ , and  $\alpha = \left(\frac{\exp(Z'\delta)}{1+\exp(Z'\delta)}\right)$ .<sup>6</sup> This parameterization restricts  $k > 0$ ,  $b > 0$ , and  $0 < \alpha < 1$ , as is required for the Gamma distribution.

As is typical for maximum likelihood methods, our estimates of the sensitivity of the willingness to pay to covariates and to anchoring bias will depend in part on our distributional assumptions. While we cannot directly test our assumption that the “true” willingness to pay distribution is a member of

---

<sup>5</sup>Whether a student is “Low income” is defined by whether the student’s annual household income is less than \$25,000 (for graduate students) or by whether the student is on financial aid (for undergraduates).

<sup>6</sup>When allowing for  $\alpha$  to depend upon  $X$ , the function could not be maximized using standard hill climbing methods or through a grid search. Instead,  $Z$  contains a subset of the characteristics from  $X$  which might theoretically affect the level of anchoring or for which there is evidence of this when estimating the model separately for different subgroups.

the three-parameter gamma family, we can conduct some tests to examine whether this distributional assumption is plausible. In particular, we run Kolmogorov-Smirnov tests to see if the underlying probability distribution (the reported willingness to pay distribution  $\bar{W}$ ) differs from our hypothesized distributions. At the 5% level, we cannot reject the hypothesis that the observed willingness to pay was generated by the three parameter gamma distribution implied by our maximum likelihood estimates.<sup>7</sup>

Table 3 shows the effect of covariates on the willingness to pay distribution (that is, the parameters of the gamma distribution). Because covariates enter in different parameters of the distribution, we report in Table 4 the marginal effect of each covariate on willingness to pay,  $E(f(x = 1)) - E(f(x = 0))$ , and calculate variances using an asymptotic bootstrap.<sup>8</sup> We find that sicker students have a higher willingness to pay, while low-income individuals and those with chronic conditions have a lower willingness to pay. In dollar terms, students with more urgent symptoms are willing to pay about \$10 more on average than those with less urgent symptoms, while low-income students and those with chronic conditions have a willingness to pay of \$7 dollars and \$4 less, respectively.<sup>9</sup>

Table 5 shows the maximum likelihood estimates of the effect of anchoring. Aside from the fact that we can reject the hypothesis that  $\alpha = 0$  (no anchoring bias) at the 0.01 level for every subgroup, it is difficult to interpret the parameter estimates. We simulate some scenarios that make more clear what our estimates mean:

- Starting the bidding game at \$50 results in willingness to pay estimates over \$7 higher on average than when starting at \$5, an increase of about 60%.

---

<sup>7</sup>We also tried other functional form assumptions, including an exponential distribution, a weibull distribution, and a two-parameter Gamma distribution. Based upon the Kolmogorov-Smirnov statistic, the three parameter gamma provided the best fit among these, and it was the only distribution to not be rejected by the test.

<sup>8</sup>Under standard assumptions for maximum likelihood methods, our parameter estimates are asymptotically distributed joint normal. We run the asymptotic bootstrap by taking 1000 draws from the asymptotic joint distribution of the parameters. For each draw and for each parameter, we calculate  $E(f(x = 1)) - E(f(x = 0))$ .

<sup>9</sup>Even though these latter groups are not willing to pay as much for Vaden services, we find in results not reported here (see Isen, 2006) that the low-income and those with chronic conditions would not be deterred from seeking care for urgent symptoms with modest cost sharing levels.

- When pooling all the brackets in the sample (where some of the anchoring is randomly dispersed throughout the distribution), we still get a willingness to pay estimate that is upwardly biased by approximately 9%.
- The bias in the price elasticity implied by this estimate varies between 9% and 80% depending upon the starting bracket, and when pooling all the brackets, the bias is 27%

The anchoring effect varies for different subgroups of students. To simplify the interpretation of results, we report the marginal effect on  $\alpha$  of changing each covariate from zero to one. The only statistically significant difference is for white students; relative to non-whites, anchoring is lower for white students by 0.135 points. One possible explanation for this result is that, as the literature on racial disparities in health care has documented, whites tend to use more health care services in their lifetime; perhaps individuals with more experience with a good are less affected by anchoring.

## 4 Acquiescence Bias

To test for acquiescence bias, we compare the responses of students for whom the first bracket question in the iterative bidding game was posed with a ‘yes’ or ‘no’ answer against students for whom the first question was more neutrally worded. Recall that students were assigned the form of this first question randomly. We run a probit regression of whether a participant would have visited the health center (for his last visit) if he had had to pay a fee equivalent to his first bracket amount against an indicator for whether the participant was assigned the ‘yes’ or ‘no’ form of the first bracket question.

Table 6 shows the results for this test in column 1. The probability of answering in the affirmative to the first bracket when given a ‘yes’ or ‘no’ question increases by 11.7 percentage points, which makes these students 35% more likely to answer in the affirmative relative to the main sample of students. This is strong evidence of acquiescence bias. The other results are unsurprising—an increase in the (randomly assigned) fee students are asked if they would pay reduces the likelihood they answer in the affirmative, low income students are less willing to visit despite a fee, chronically ill students and those with urgent medical conditions are more willing.

We turn now to measuring the effect of anchoring bias on the measured willingness to pay. We do this by first regressing the “true” willingness to pay estimates (corrected for anchoring bias using the method described in Section 3) on whether respondents were exposed to the neutral version of the first bracket question in the iterative bidding game. We then take these fitted values to those exposed to the non-neutral wording to predict what their willingness to pay would have been and compare them to those exposed to the neutral wording.

The results are in the second column of Table 6. Non-neutral wording leads to a willingness to pay estimate that is upwardly biased by around 30% as well as an elasticity estimate that is downwardly biased by around 20%. It is clear from these results that posing the first bracket entry question in a non-neutral manner causes a substantive bias. In the rest of the paper, we thus analyze only the subgroup of respondents who were posed a neutral version of that first question.

## 5 Willingness to Pay and Estimated Demand Elasticities.

From Section 3, we obtain an estimate of each students’ true willingness to pay interval for health care that is purged of anchoring bias using (4). For this section, let  $s$  be a unique identifier for each student and let the pair  $(\tilde{b}_{i-1}^s, \tilde{b}_i^s)$  be the purged bounds on willingness to pay for student  $s$  from the iterative bidding game. We estimate  $W_s$  using the following equation, which is implied by our distributional assumption (5):

$$W_s = \min(\tilde{b}_{i-1}^s, \tilde{b}_i^s) + \int_{\min(\tilde{b}_{i-1}^s, \tilde{b}_i^s)}^{\max(\tilde{b}_{i-1}^s, \tilde{b}_i^s)} \frac{(x-g)^{k-1}}{\Gamma(k) b^k} \exp\left(\frac{-(x-g)}{b}\right) x dx \quad (7)$$

The demand curve for health care at Stanford is defined as a function,  $Q(P)$ , which represents the number of visits to Vaden over a year at price,  $P$ . The empirical analog of this object, say  $\hat{Q}(P)$ , in our context is the number of students in our sample, weighted by their number of visits in the previous year,  $v_s$ , whose willingness to pay exceeds  $P$ :

$$Q(P) = \sum_{s=1}^{n-1} v_s 1(W_s \geq P) \quad (8)$$

Figure 1 shows a non-parametric plot of  $\ln Q(P)$  against  $\ln P$ . It also shows the best fitting quadratic curve that approximates this plot. When price is set at the mean student willingness to pay, the elasticity of demand  $\frac{d \ln Q}{d \ln P}$  equals  $-0.7$ .

While the famous RAND study found a price demand elasticity for all health care service in the general population to be about  $-0.2$ , the estimate for first tier care at Stanford is not very surprising. After all, the Stanford population is healthier than the general population, the student health center provides only first tier care, and all students possess other medical insurance which serves as a close substitute for the care available at Stanford student health. Demand elasticities at levels close to one such as this imply that cost sharing would increase welfare by stemming moral hazard and the overuse of medical care. Additionally, given the low average willingness to pay of \$15.28, the gains from risk pooling are limited.

## 6 Bias From Selective Non-Response

Although we contacted a random sample of Stanford students for the survey, response rates were low enough—30.8%—to raise concern about the representativeness of our final sample. We use two different methods to address the possibility that our willingness to pay estimates are biased by the high rate of non-response.

First, we estimate a Heckman selection model (Heckman, 1979). The Stanford Administration provided demographic information on all individuals generated in the random sample, whether or not the student replied to our survey. We used these demographic data to match respondents and non-respondents in the context of a Heckman selection model. In the first stage, we model whether a student responded to the survey, and in the second stage we model respondents' 'true' willingness to pay, adjusted for anchoring bias.

We use the \$5 gift certificate that we offered to a randomized subset of respondents as an instrumental variable. Though the Heckman selection model is identified by functional form assumptions without any required exclusion restrictions, having a plausible instrumental variable makes the estimates

more believable. We include an indicator for whether each student was offered a participation incentive in the first stage, but not in the second stage.

Table 7 shows the results from the Heckman selection model. The probability of responding is increased by 5.8% when a student is offered compensation, but the coefficient on the inverse Mills ratio in the second stage indicates that we cannot reject the hypothesis that there is no selection bias.

While these results are suggestive, we do not view them as definitive. The Heckman selection model requires a strong and untestable functional form assumption about the joint distribution of the errors in the selection and outcome equations—bivariate normality. We thus turn to bounding methods to see how sensitive our estimates are to the possibility of non-response bias. We estimate three different sets of bounds: (1) the (effectively) assumption-free bounds of Manski (1990); (2) the Manski instrumental variables (IV) bounds (Manski and Pepper, 2000); and (3) the structural bounds of (Philipson, 2001) which require both an instrumental variable and some structural assumptions about the effects of selection.<sup>10</sup> We use the randomized participation incentive as our instrumental variable for the Manski IV bounds and the structural bounds.

Let  $W_1$  be the (anchoring corrected) willingness to pay among people who responded to the survey and let  $W_0$  be the anchoring corrected willingness to pay among people who did not respond. Let  $D$  be an indicator of whether a student responded to the survey. We are interested in population mean willingness to pay,  $W = DW_1 + (1 - D)W_0$ :

$$\begin{aligned} E[W] &= E[DW_1] + E[(1 - D)W_0] \\ &= E[W_1|D = 1]P[D = 1] + E[W_0|D = 0]P[D = 0] \end{aligned} \tag{9}$$

The terms  $E[W_1|D = 1]$  and  $P[D = 1]$  are readily observable from the data we have, but we do not observe  $E[W_0|D = 0]$ , which is the mean willingness to pay among the non-responders. We can, however, impose some *a priori* bounds on it. Willingness to pay must not be negative:  $W_0 > 0$ . Replacing  $E[W_0|D = 0]$  with zero in (9) thus yields the following lower bound on  $E[W]$  that depends only on observed quantities:

$$E[W] \geq E[W_1|D = 1]P[D = 1] \tag{10}$$

---

<sup>10</sup>For related bounding approaches, see Shaikh and Vytlacil (2004); Bhattacharya et al. (2005).

Similarly, since our iterative bidding game has an upper end point,  $M = \$150$ , we have that  $W_0 < M$ . Applying this inequality yields an upper bound on  $E[W]$ :

$$E[W] \leq E[W_1|D = 1]P[D = 1] + MP[D = 0] \quad (11)$$

Together (10) and (11) constitute the Manski bounds without an instrumental variable. These can be sharpened if an instrumental variable, such as we have, is available.

Let  $\rho$  be an indicator for whether each person contacted was provided a participation incentive. For an instrumental variable, we require that  $P[D = 1|\rho = 1] > P[D = 1|\rho = 0]$  (which holds in our data) and that  $E[W_0|\rho] = E[W_0]$  and  $E[W_1|\rho] = E[W_1]$  (which we are willing to assume since  $\rho$  is randomly assigned). Clearly, the latter assumptions imply that  $E[W] = E[W|\rho]$  for all  $\rho$ . The Manski bounds above still hold conditional on  $\rho$  since  $W$  is still bounded below by zero and above by  $M$ :

$$\begin{aligned} E[W_1|D = 1, \rho]P[D = 1|\rho] &\leq E[W] \\ &\leq E[W_1|D = 1, \rho]P[D = 1|\rho] + MP[D = 0|\rho] \end{aligned}$$

Applying the above inequalities twice—once for  $\rho = 1$  and once for  $\rho = 0$ —yields the Manski IV bounds:

$$\begin{aligned} \max_{\rho=0,1} (E[W_1|D = 1, \rho]P[D = 1|\rho]) &\leq E[W] \leq \\ \min_{\rho=0,1} (E[W_1|D = 1, \rho]P[D = 1|\rho] + MP[D = 0|\rho]) & \end{aligned} \quad (12)$$

We can obtain tighter bounds with an additional structural assumption, similar to that imposed by Philipson (2001). In particular, we assume that students who responded to the survey about the health center are more likely to have a higher willingness to pay for services from the center. This is a plausible assumption since those who did not respond to the survey presumably care less about student health services than those who did respond.<sup>11</sup>

---

<sup>11</sup>In our initial contact with each student, we included statements from the various Stanford authorities emphasizing the importance of this research to improving student health services. Those who care about these services seem more likely to respond to such statements by responding to the survey.

Our randomized incentive provides further evidence that this assumption is reasonable. Among those provided the incentive, average WTP is \$13.66 while for those not so provided, it is \$15.44. Since those provided the incentive were more likely to respond, these results support the idea that students who are willing to pay more for student health services are more likely to respond (notwithstanding our results from the Heckman selection model). Formally, this assumption amounts to the following:

$$E[W_0|D = 0, \rho = r] < E[W_1|D = 1, \rho = r] \text{ for } r = 0, 1.$$

Unlike the assumptions imposed for the Heckman selection model, these structural assumptions are not tight enough to ensure point identification. Imposing the above structural assumption on (9) yields the following bound on the population mean willingness to pay:

$$\begin{aligned} \max_{\rho=0,1} (E[W_1|D = 1, \rho]P[D = 1|\rho]) \leq E[W] \leq \\ \min_{\rho=0,1} (E[W_1|D = 1, \rho]) \end{aligned} \quad (13)$$

Finally, we calculate the elasticity of demand at the extreme point of each bound. We need one more piece of notation. Let  $N$  be the size of the sample contacted (and let  $n$  be the number of responders). We are interested in the population elasticity, which we get from (8):<sup>12</sup>

$$Q(P) = \sum_{s=1}^{N-1} 1(W_s \geq P) = \sum_{s=1}^{n-1} 1(W_s \geq P) + \sum_{s=1}^{N-n-1} 1(W_s \geq P) \quad (14)$$

We observe in the right hand equation the first term but not the second, which is the non-responder quantity at each given price. At the lower bound, we set the willingness to pay of each non-responders to the lowest price level, which yields the following elasticity at the lower bound:

$$Q(P) = \sum_{s=1}^{N-1} 1(W_s \geq P) \text{ where } W_s = \min P \text{ for } 1 < s < N - n - 1 \quad (15)$$

---

<sup>12</sup>For simplicity, we do not weight by the number of visits for each student in the previous year in this section. If one wishes to include the weight, one can use the total number of visits in the last year to determine the number of visits of non-responders and construct the bounds accordingly, which in our case, provides very similar results.



The elasticity at the upper bound follows in a similar way where we set the willingness to pay of non-responders to the highest price level:

$$Q(P) = \sum_{s=1}^{N-1} 1(W_s \geq P) \text{ where } W_s = \max P \text{ for } 1 < s < N - n - 1 \quad (16)$$

Together (15) and (16) constitute the elasticity corresponding to the Manski bounds without an instrumental variable (where the elasticity  $\frac{d \ln Q}{d \ln P}$  is evaluated at the mean willingness to pay of the quadratic demand approximation). Similar logic applies in constructing the elasticities corresponding to the narrower bounds.

The results, in Table 8, indicate a much tighter set of willingness to pay estimates as we move from the Manski bounds, to the Manski IV bounds, to the structural bounds. While the elasticity estimates at the extreme points of the Manski bounds and the Manski IV bounds are substantively wide, the elasticity estimates at the extreme points of the structural bounds are nearly identical to the elasticity estimate (uncorrected for selection bias) that we report in Section 5.

## 7 Conclusion

While contingent valuation games are an important method used to calculate demand when prices are difficult or impossible to observe, in practice the method can be subject to a number of problematic biases, including anchoring and acquiescence biases. In a survey context, selective non-response by those who care least about the good being valued may also lead to incorrect inferences. In this paper, we demonstrate ways to adjust inferences from an iterative bidding game that simultaneously accounts for all three of these sources of bias.

In our example of calculating the demand curve for student health at Stanford University, it is useful to see how large the bias in the willingness to pay estimate and the elasticity estimate is induced by each type of bias separately. We show in Table 9 the relative effects of anchoring, acquiescence, and selection biases on willingness to pay and elasticity:

$$\left| \frac{WTP_{biased} - WTP_{corrected}}{WTP_{corrected}} \right| \text{ and } \left| \frac{Elasticity_{biased} - Elasticity_{corrected}}{Elasticity_{corrected}} \right|$$

In our context, it appears that anchoring and acquiescence are of similar magnitude in biasing the results whereas selection appears to do little in biasing the elasticity (although may bias willingness to pay estimates).

Anchoring bias can be overcome by using random bracket entry and creating a model to assess and control for the effect. Some studies have only used random bracket entry (for example Asgary et al., 2004, and the Health and Retirement Survey), but the results from pooling all of the brackets indicate that it is important to also estimate an anchoring model. Acquiescence bias can be significantly reduced by avoiding “yes” or “no” questions.

Lastly, even though we find that selection bias is of little concern for our elasticity estimate, it is useful to induce a higher response rate among a random subsample, through randomized payment incentives or by expending more effort to elicit responses. These strategies make possible robust inference in the face of selective non-response. In particular, they provide both an unassailable exclusion restriction for the first stage of a selection model, and they permit the construction of tighter bounds on the bias from nonresponse. When all of these strategies are combined, they yield more plausible estimates of demand from iterative bidding games in the contingent valuation tradition.

Figure 1: Non-Parametric Demand Function

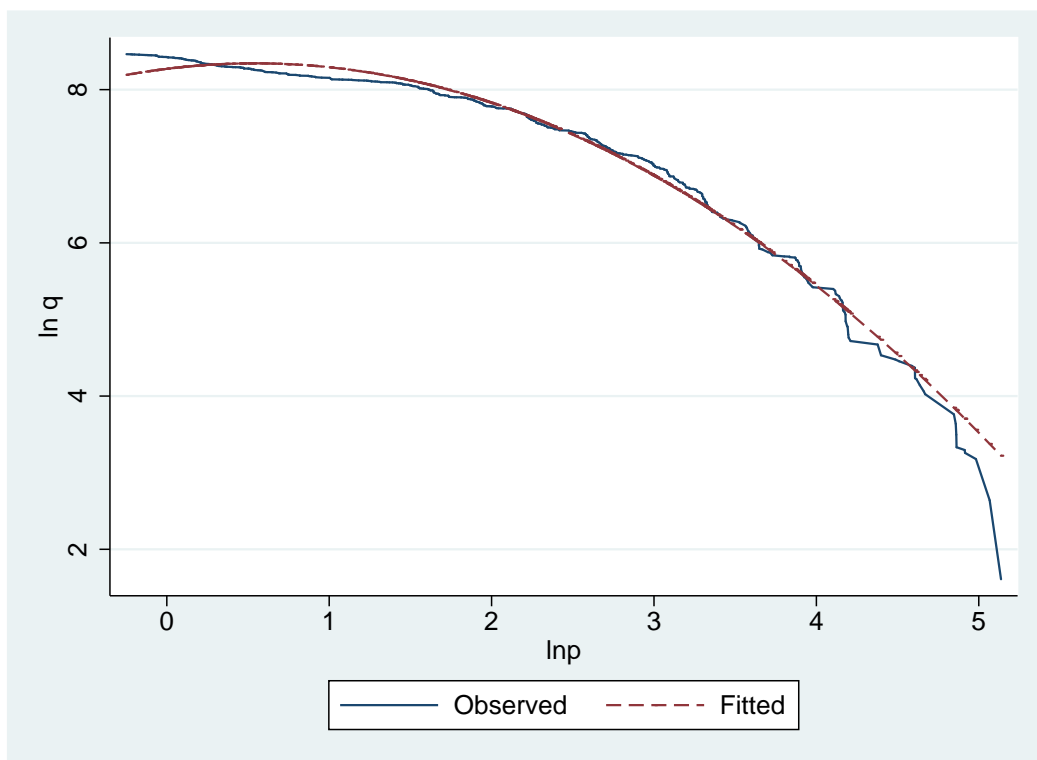


Table 1: Demographic and Health Descriptive Statistics

Category	Variable	All	Undergrad.	Grad.
Sex	Male	47%	38%	53%
Race/Ethnicity (People may report multiple races)	White	60%	58%	62%
	Black	6%	10%	3%
	Asian	29%	26%	31%
	Latino	10%	12%	9%
On Financial Aid	Yes	46%	46%	-
	No	54%	54%	-
Parent Education	High School	7%	7%	-
	Some College	6%	6%	-
	College	16%	16%	-
	Graduate School	71%	71%	-
Household Income (Grad. students only)	<\$25K	37%	-	37%
	\$25K-\$35K	37%	-	37%
	\$35K-\$55K	10%	-	10%
	>\$55K	17%	-	17%
Health Insurance	Through Stanford?	58%	29%	62%
Health Status	Excellent	29%	29%	-
	Very Good	53%	53%	-
	Good	13%	13%	-
	Fair	4%	4%	-
	Poor	1%	1%	-
	Med. Condition	16%	17%	15%
Sample Size		883	382	501

Table 2: Descriptive Statistics About Health Care

Category	Variable	All	Undergrad.	Grad.
2005 Visits	None	25%	11%	12%
	1 visit	18%	24%	19%
	2-3 visits	36%	43%	41%
	4-5 visits	12%	11%	16%
	6+ visits	10%	11%	11%
Urgency of Last Visit 1 = not urgent 5 = most urgent	1	39%	35%	43%
	2	32%	34%	30%
	3	20%	26%	20%
	4	5%	6%	5%
	5	3%	5%	2%
Reason for Last Visit (May report multiple reasons)	Digestive	8%	8%	8%
	Dermatological	13%	12%	14%
	Ear/Nose/Throat	33%	37%	30%
	Migraine/Headache	5%	6%	4%
	Gynecological	23%	22%	23%
	Mental Health	11%	13%	10%
	Chest	4%	3%	3%
	Lethargy	4%	6%	3%
	Cold/Fever/Flu	24%	31%	20%
Willingness-to-pay by starting bracket	All brackets	\$16.83	\$14.27	\$18.79
	\$5 bracket	\$13.20	\$8.51	\$16.81
	\$20 bracket	\$15.27	\$11.94	\$17.88
	\$35 bracket	\$17.96	\$16.80	\$18.80
	\$50 bracket	\$20.79	\$19.76	\$21.59
Answer in the affirmative to first bracket question	Neutral wording	27.41%	21.99%	31.54%
	Non-neutral wording	38.10%	36.07%	40.00%

“Digestive” indicates vomiting, diarrhea, abdominal pain, or blood in stool.

“Ear/Nose/Throat” indicates allergies, cough, throat, eye problem, ear or hearing problem, sinus problem, and difficulty breathing.

“Mental Health” indicates anxiety, depression, or other psychological problems.

“Chest” indicates chest pain and rapid or irregular heartbeat.

Table 3: Willingness to Pay Estimates

Variable	Shape Param. $k$	Scale Param. $b$	Location Param. $g$
More than 1 visit	-0.200 (0.171)	0.057 (0.203)	-
Chronic Condition	0.071 (0.229)	-0.365 (0.249)	-
Student health insurance	-0.010 (0.127)	0.216 (0.157)	-
Low Income	-0.008 (0.113)	-0.459** (0.145)	-
Nonurgent	0.120 (0.113)	-0.668** (0.155)	-
Graduate Student	0.529** (0.183)	-0.156 (0.220)	-
White	0.293 (0.223)	0.004 (0.255)	-
Constant	-0.887** (0.236)	3.762** (0.286)	0.989 (0.409)
N	883		

Standard errors are included in parantheses. Log-likelihood = -1737.48

\*\* significant at 1%

\* significant at 5%

Table 4: Effect of coefficients on WTP

Variable	$F(x = 1) - F(x = 0)$
More than 1 visit	-2.21** (0.050)
Chronic Condition	-4.25** (0.051)
Student health insurance	3.16** (0.050)
Low Income	-6.89** (0.045)
Nonurgent	-9.67** (0.069)
White	5.22** (0.046)
Graduate Student	4.40** (0.054)

Standard errors are included in parantheses and constructed by an asymptotic bootstrap

\*\* significant at 1%

\* significant at 5%

Table 5: Anchoring Results

$Z$	Estimates	$\alpha(Z = 1) - \alpha(Z = 0)$
More than 1 visit	0.217 (0.385)	0.034
Chronic Condition	0.377 (0.434)	0.058
Graduate student	-.232 (0.431)	-0.036
White	-0.863* (0.326)	-0.135*
Constant	-0.994* (0.310)	-
Sample Anchoring Mean ( $\bar{\alpha}$ )	0.203	
N	883	

Standard errors are included in parentheses

\*\* significant at 1%

\* significant at 5%



Table 6: Tests for Acquiescence Bias

Variable	Effect on First Bracket Question Marginal Effect	Effect on WTP Estimate
Non-neutral question	0.117*** (0.0473)	4.45** (2.00)
First bracket amount	-0.0111*** (0.000891)	0.06 (0.04)
Undergraduate student	-0.106*** (0.0284)	-5.11*** (1.33)
Low Income	-0.0890*** (0.0291)	-5.30*** ( 1.36)
Two or more Visits	0.00778 (0.0316)	-1.15 (1.43)
White	0.0878*** (0.0289)	4.95*** (1.37)
Chronic condition	0.0867** (0.0354)	-4.10* (1.81)
Nonurgent visit	-0.141*** (0.0340)	-7.93*** (1.45)
Constant	-	22.07*** (2.31)
obs. P	0.287	
pred. P	0.245	
N	1,009	1,009

All estimates are reported as changes in the probability of being willing to come for visit despite a fee equal to b1 being charged. Standard errors are included in parentheses.

\*\*\* significant at 1%

\*\* significant at 5%

\* significant at 10%

Table 7: Heckman Selection Correction Model

Variable	Responded to Survey	WTP
Inverse Mills Ratio		-3.09 (10.74)
Offered compensation	0.058*** (0.016)	
Student Health Insurance	-0.017 (0.018)	4.09*** (1.59)
Graduate Student	0.088*** (0.031)	-2.69 (2.86)
Professional Student	-0.069** (0.027)	8.44*** (2.69)
PhD Student	-0.080*** (0.026)	3.40 (2.55)
BA Student	-0.090*** (0.024)	0.92 (3.07)
BS Student	-0.102*** (0.025)	-1.46 (3.54)
Masters Student	-0.162*** (0.024)	3.70 (5.10)
Asian	0.008 (0.025)	2.13 (2.00)
Black	-0.025 (0.034)	-1.29 (3.12)
White	0.022 (0.020)	6.11*** (1.64)
Latino	0.003 (0.031)	2.86 (2.56)
N	3,557	1,057

We report marginal effect estimates rather than coefficients throughout.

Standard errors are included in parentheses.

\*\*\* significant at 1%

\*\* significant at 5%

\* significant at 10%

Table 8: Non-Response Bias–Bounding Estimates

	Bounds		
	Manski	Manski IV	Structural
WTP Lower Bound	5.69	5.89	5.89
WTP Upper Bound	103.77	97.30	13.66
Elasticity Lower Bound	-.06	-.08	-.72
Elasticity Upper Bound	-.76	-.72	-.72

Table 9: Magnitude of Biases

		WTP Bias	Elasticity Bias
Anchoring	Pooled brackets	7.7%	27.4%
	\$5 bracket	20.3%	9.3%
	\$50 bracket	27.2%	80.1%
Acquiescence	Non-neutral wording	29.2%	19.3%
Selection	Heckman correction	0%	0%
	Structural bounding	11.7 - 159.0%	3.3-3.7%

## References

- Asgary, A., Willis, K., Taghvaei, A. A., and Rafeian, M. (2004). Estimating rural households willingness to pay for health insurance,. *European Journal of Health Economics*, 5(3):209–215.
- Bhattacharya, J., Shaikh, A., and Vytlačil, E. (2005). Treatment effect bounds: An application to swan ganz catheterization. Working Paper 11263, National Bureau of Economic Research.
- Blumenschein, K., Johannesson, M., Yokoyama, K., and Freeman, P. (2001). Hypothetical versus real willingness to pay in the health care sector: Results from a field experiment. *Journal of Health Economics*, 20(3):441–457.
- Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*, 47:153–161.
- Herriges, J. and Shogren, J. (1996). Starting point bias in dichotomous choice valuation with follow-up questioning. *Journal of Environmental Economics and Management*, 30:112–131.
- Hurd, M. and van Soest, A. (2003). ”a test for anchoring and yea-saying in experimental consumption data”. Working paper, RAND.
- Isen, A. (2006). First tier care at stanford university. Senior Honors Thesis, Department of Economics.
- Manski, C. (1990). Nonparametric bounds on treatment effects. *American Economic Review, Papers and Proceedings*, 80:319–323.
- Manski, C. and Pepper, J. (2000). Monotone instrumental variables: With an application to the returns to schooling. *Econometrica*, 68:997–1010.
- O'Brien and Viramontes (1994). Willingness to pay: a valid and reliable measure of health state preference? *Medical Decision Making*, 14(3):289–297.
- Pauly, M. (1968). The economics of moral hazard. *American Economic Review*, 58:531–537.
- Philipson, T. (2001). Data markets, missing data, and incentive pay. *Econometrica*, 69(4):1099–1111.

- Ryan, M., Scott, D., and Donaldson, C. (2004). Valuing health care using willingness to pay: A comparison of the payment card and dichotomous choice methods. *Journal of Health Economics*, 23(2):237–258.
- Shaikh, A. and Vytlačil, E. (2004). Limited dependent variable models and bounds on treatment effects: A nonparametric analysis. mimeo, University of Chicago and Columbia University.
- Smith, R. (2006). Its not just what you do, its the way that you do it: the effect of different payment card formats and survey administration on willingness to pay for health gain. *Health Economics*, 15:281–293.
- Stalhammar, N.-O. (1996). An empirical note on willingness to pay and starting-point bias. *Medical Decision Making*, 16:242–247.