Estimating Willingness to Pay by Means of Adaptive, Sequential Conjoint Experiments

Gregory M. Duncan, *Department of Economics, University of California, Berkeley, and The Huron Consulting Group*
Joel Huber, *The Fuqua School of Business, Duke University*
W. Kip Viscusi, *Harvard University and Vanderbilt School of Law*

Abstract

We present a new method for determining willingness to pay by sequentially offering alternatives to respondents; the alternatives are based on previous responses and can bracket the respondent's willingness to pay. The process while intuitive and conceptually easy to describe is difficult to model. The sequential structure causes problems familiar in empirical auction literature. Our approach is entirely different. Earlier attempts to analyze these data used empirical Bayes method. Here we develop and use a formal discrete choice model of the sequential choices. We show how the choices can be described by some simple moment conditions involving complicated probabilities. We use simulated method of moments to extract estimates from the moment conditions and from the latter the willingness to pay for water quality improvements.
The method accurately determines the willingness to pay for any product and will have use in real-time computer analysis of buying choices; it may also have application in the empirical analysis of auctions.

Introduction:

The problem addressed is estimating a person's willingness to pay for improved water quality in his or her region. The experiment we designed is conceptually simple but one difficult to model. In many aspects it presents problems similar to those found in the auctions literature where one observes a sequence of bids and tries to infer the willingness to pay from the sequence[1].

The economic benefit of water quality improvements is society's willingness to pay for increases in water quality. Early measures of water quality were derived from travel cost values of recreational benefits. Subsequent benefit assessments, which remain in use in some policy applications, consist of analyzing the value of improvements in the water's ranking on a water quality ladder. This unidimensional water quality index assumes that there is a hierarchy of quality levels in terms of whether the water is drinkable, swimmable, fishable, or boatable. Thus, water that is drinkable also meets acceptability criteria for all lower ranked uses. Unfortunately, this hierarchical characterization is problematic, as these categories of uses do not reflect our current scientific understanding of the empirical ordering of water quality. That is, if one

1 See for example, Paarsch, Harry J. and Han Hong (2006)

examines the pattern of quality levels across states, there is almost no evidence of such a hierarchy. The focus of the survey results reported here is on people's willingness to pay for water that is rated "good" based on an overall index, developed by the U.S. Environmental Protection Agency (EPA), that initially merges benefits with respect to fishing, swimming, and the quality of the aquatic environment. An additional survey component makes it possible to separate the component values. The survey results reported here will focus on the overall water quality valuation component.

This paper expands and tests the methodology developed by Magat, Huber, Viscusi, and Bell (2000), where water quality values are derived from hypothetical market choices. These values are based on simple choices between regions that differ on water quality and cost of living. A series of such choices yield bounds on the value of water quality improvements for each individual. The method has the advantage of generating estimates of the private value of improvements in water quality from a simple understandable task. An important innovation in this paper is a method for handling and modeling indifference or inability to distinguish between options in a discrete choice framework.

This paper discusses econometric stability of these estimates as well as some reliability and sampling questions that arise in this use of iterative choice to assess private values. The study is based on over 1,000 new surveys implemented through web-based interviewing. Generally, we find that water quality valuations follow expected economic patterns: factors such as income, education, and visits to lakes or rivers are appropriately related to the value of water quality. Further, a scope test indicates greater valuations for larger changes in water quality gains, increasing confidence in the metric quality of the results. We assess the reliability of this approach by testing for the stability of the results given different econometric assumptions, with particular focus on those responses for which the dollar value of water quality could only be bounded on one side.

A second important improvement in this study is the use of a national web-based panel rather than the recruitment to regional central sites or mall intercepts used in the Magat *et al.* (2000) study. The use of respondent panels for policy has emerged as a response to increasing difficulty and expense attached to recruiting probability-based random samples. It is fundamentally an empirical question whether a panel-based sampling approach will produce acceptable results. We find that the demographic characteristics of the final sample closely correspond to that of the target universe of U.S. adults. Additionally, we show that that the results are not affected by factors that might distinguish between those who take the survey against those who do not. Finally, a sample selection procedure adjusts the water quality valuations for the probability that a panel member will not take or successfully complete the survey. These estimates differed little from the unadjusted means, providing assurance that they are relatively independent of possible panel selection biases.

Section 2 describes the overall study design, the survey methodology and the iterative choice method for generating values for improvements in water quality. Section 3 explores the logical adequacy of the results, including an exploration of consistency tests for the responses as well as the variation of the valuation responses conditioned on demographics. Section 4 provides tests of survey and sample validity. The survey was internet-based, using the Knowledge Networks panel. We examine the extent to which attrition bias from the panel and other aspects of this survey mode influence the water quality values. As indicated in the concluding Section 5, the results are quite robust and meet a wide variety of tests for rationality and consistency.

2. Study Design

The survey used a computer-based methodology and was administered to a national sample. The average respondent completed the survey in 25 minutes. The instrument initially acquainted the respondent with the meaning of regional differences in lake and river water rated of good quality and differences in annual cost of living. This introductory section establishes the cognitive groundwork for the respondents so that a choice between regions differing in these aspects can be reliably answered.

Introductory section in the survey

The key valuation task involves choices between regions differing in their levels of water quality and the annual cost of living. A critical part of the method involves introductory sections that encourage the respondent to think about these tradeoffs. This process begins with some very general questions to encourage the respondent to think about the value of freshwater bodies. It also elicits information on the frequency of visits to lakes and rivers as well as related activities, such as boating, fishing, or swimming.

The primary reason for asking about usage is to encourage respondents to think about why they might value differences in water quality. However, it may also be the case that respondents reporting greater usage of lakes and rivers have higher valuations of improvements in the quality of those water bodies.

Immediately following the introduction to water usage, the survey explains the meaning of cost of living and elicits the respondent's level of concern with an annual increase in cost of living of $200. Respondents then respond to a question that tests comprehension involving a simple choice between two regions, identical except that one is more expensive. The few respondents who chose the more expensive location are provided a brief educational module before being asked to proceed.

Next, respondents are introduced to the criteria that define what it means for water quality to be "good." Consistent with definitions used by EPA's National Water Quality Inventory, the survey provides the following definition:

> The government rates water quality as either
>
> * Good, or
> * Not Good.
>
> Water quality is Good if the water in a lake or river is safe for all uses.
> Water quality is Not Good if a lake or river is polluted or unsafe to use.
> More specifically, water quality is Good if the lake or river
>
> * Is a safe place to swim,
> * Fish in it are safe to eat, and
> * Supports many plants, fish, and other aquatic life.
>
> Water quality is Not Good if the lake or river
>
> * Is an unsafe place to swim due to pollution,
> * Has fish that are unsafe to eat, or
> * Supports only a small number of plants, fish, and other aquatic life.

The survey then explicitly excludes drinking water from the valuation task.

Once familiar with the concepts of water quality and cost of living, these contexts are framed within context of a region, defined as "within a 2-hour drive or so of your home, in other words, within 100 miles." A 100 mile radius is appropriate because it reflects a reasonable 2-hour drive for the recreational use of bodies of water, and about 80 percent of all recreational visits for lakes, rivers, and streams are within such a radius.[7] This text explanation of region contrasts with the method reported in Magat et al. (2000) where respondents viewed pictorial representations of the region size. However, our pretest interviews indicated that the 100-mile region radius could be well understood when described through the text used.

After they learned about water quality and the region, respondents received a warm-up choice. In this case they were asked to choose between two regions that differed in the percentage of water bodies with quality rated good. Respondents who preferred the region with a lower percent of lakes and rivers rated good received a brief interactive tutorial on the meaning of the benefit measure and the error in their response.

Key Valuation Choice Task

Once respondents learn about water quality, cost of living and their application to a region, they are ready for the iterative choice questions. This key valuation task is designed to elicit the respondent's tradeoff between water quality and cost of living in choices between different regions. These regions are "the same in all other ways, including the number of lakes and rivers near your home." As a final warm-up question respondents are asked to make a choice where one alternative dominated another on both cost of living and water quality. That is, they choose between two regions, where one region had more quality lakes and rivers and lower cost of living. Respondents who erred received a remedial tutorial that reviewed the nature of the choice being made.

The critical choice questions take the form shown in Figure 1. It is noteworthy that the task itself is not complex, which past evidence suggests should enhance the validity of the survey approach. We will also present a series of rationality tests of the survey responses as validity checks of the methodology.

If a respondent was indifferent in the initial choice presented in Figure 1, then the iterative choice process is complete. A choice of either alternative led to successive choices that terminated either at indifference or choice of one of the regions with a specified cost and water quality. Figure 2 displays the logic of the iterative choice questions. The program iterates choices, each time degrading the desirable aspect of the last alternative chosen until the selection reverses. For example, a respondent preferring the lower cost region on the initial question in Figure 1 then considers the same pairwise choice, except the cost of living in that region is raised. Continued preference for the lower cost region leads to continued increases in the cost of living in the chosen region until the respondent faces a dominated choice in which the regions have the same cost of living but differ only in terms of water quality. Similarly, continued preference for the higher quality region leads to continued reductions in the water quality of the chosen region until the regions have the same water quality but differ only in cost of living.

The survey also ends with a number of additional sections, such as a brief series of demographic questions and whether the respondent had difficulty understanding any part of the survey. This process of elaborate training before the choice questions is one we have used a similar formulation in a wide variety of other environmental risk contexts. We have found that with sufficient grounding, the tradeoff against cost of living can be well understood. We deliberately framed the choice as one between regions similar to but abstracted from the region where the person now lives. This abstraction is one that we believe contributes to the stability, validity and actionability of the results. In terms of stability, not having to focus on a particular body of water conditioned on the location of one's home discourages inferences about one's particular circumstance that may or may not apply to a particular change in the percent of good water quality in a region. In terms of validity, the survey focuses on a free market choice that has minimal social consequences—whether one buys in region A or B primarily influences one's own utility. These market choices contrast with referenda where one's vote can affect the welfare of others, confounding the results with an array of conflicting forces including altruism, confidence in the efficacy government action, willingness to impose costs on others, and attitudes about taxation to fund such referenda. Finally, the results are actionable in helping to establish a general social metric for policy decisions across regions. The projected dollar value for changes in water quality can be related to general citizen characteristics such as age, income and education. These values can be applied using census data to evaluate a broad range of options that affect the quality of water.

Experimental conditions

In order to test the robustness of the results to different versions of the questionnaire, randomly identified groups received alternative versions. These tests permit an assessment of the effects of anchoring and the initial range of the alternatives in the initial trade off.

Our study tests for anchoring influence by manipulating the presence of an external norm for water quality. Approximately half the respondents received information that the national average of water quality was rated 65% Good, whereas the other respondents received no national information. Being told the US 65% value may increase the sensitivity to water quality, since there is now an anchor that helps respondents value of the water percent amounts provided.

Second, the value of a given change in percent good may itself be affected by the range of percent good and dollars in the initial choice. For example, if the first choice is between a gain of 20% good in return for $400 in cost of living (e.g., $20 for one percentage point), then respondents may reasonably use that information to assume that, say, $15 is a good price to pay for one percentage point gain. By contrast, if the initial choice pits a 20% gain against $200, ($10 per one percentage point), then the $15 seems relatively high. This inference is understandable if one takes the Gricean (1975) assumption that the initial choices provided in such questionnaires are reasonable. To test the impact of the initial range we altered the initial range in cost of living to be either $200, $300 or $500, and the range of the gain in percent good to be either 20, 30 or 40 percentage points. This test is whether the initial choice is appropriately sensitive to ranges, as required for appropriate sensitivity to scope.

3. Econometric Model[2]

We use a random utility approach. We assume each individual has a utility depending on the cost, *C*, and the quality, *Q*. Specifically, $U = C\beta_C + Q\beta_Q + \eta$. Following Cardell (1977), Hausman and Wise (1978) and Duncan (1980) we model taste variation with random coefficients. We allow the possibility of either indifference or an inability to discern differences through the use of an indifference threshold Δ. One alternative is

2The approach here is based on a model proposed and used by Duncan and Perry (1988) for pricing new communications services. Their model was applied to a sequential querying process similar to the one described here except only one attribute, the price, varied in response to a choice.

chosen over another only if the difference in utilities exceeds this threshold. At each stage conditional on reaching the stage, there are three possible outcomes.

If region 1 was preferred to region 2 at stage i then the event

A1: $U_{1i} - U_{2i} > \Delta$

occurred.

If region 2 was preferred then

A2: $U_{2i} - U_{1i} > \Delta$ occurred.

Otherwise there was indifference, and inability to distinguish, or no preference expressed and

A3: $-\Delta < U_{1i} - U_{2i} < \Delta$

occurred.

Rather than work directly with utilities, we model differences in utilities. Let C_{ji} and Q_{ji} are respectively the cost and quality of region j at stage i and let $DC_i = C_{1i} - C_{2i}$ and $DQ_i = Q_{1i} - Q_{2i}$ be differences in costs and qualities. The respective probabilities conditional on reaching stage i are given by

(1.1) $$\Pr\left[A1_i\right] = \Pr\left[DC_i\beta_C + DQ_i\beta_Q + \varepsilon_i > \Delta\right]$$

(1.2) $$\Pr\left[A2_i\right] = \Pr\left[DC_i\beta_C + DQ_i\beta_Q + \varepsilon_i < -\Delta\right]$$

(1.3) $$\Pr\left[A3_i\right] = \Pr\left[-\Delta \le DC_i\beta_C + DQ_i\beta_Q + \varepsilon_i \le \Delta\right]$$

Another parameterization that will make it easier to determine willingness to pay is

(1.4) $$\Pr\left[A1_i\right] = \Pr\left[-DC_i - \frac{\beta_Q}{\beta_C}DQ_i + \frac{\varepsilon_i}{-\beta_C} > \frac{\Delta}{-\beta_C}\right]$$

(1.5) $$\Pr\left[A2_i\right] = \Pr\left[-DC_i - \frac{\beta_Q}{\beta_C}DQ_i + \frac{\varepsilon_i}{-\beta_C} < -\frac{\Delta}{-\beta_C}\right]$$

(1.6) $$\Pr\left[A3_i\right] = \Pr\left[-\frac{\Delta}{-\beta_C} \le -DC_i - \frac{\beta_Q}{\beta_C}DQ_i + \frac{\varepsilon_i}{-\beta_C} \le \frac{\Delta}{-\beta_C}\right]$$

Note that since we expect the effect of cost to be negative we divide by $-\beta_C$ and do not reverse signs. The ratio $-\frac{\beta_Q}{\beta_C}$ can be interpreted as the willingness to pay.

(1.7) $$\Pr\left[A1_i\right]=\Pr\left[-DC_i-wtpDQ_i+\frac{\varepsilon_i}{-\beta_C}>\frac{\Delta}{-\beta_C}\right]$$

(1.8) $$\Pr\left[A2_i\right]=\Pr\left[-DC_i-\frac{\beta_Q}{\beta_C}DQ_i+\frac{\varepsilon_i}{-\beta_C}<-\frac{\Delta}{-\beta_C}\right]$$

(1.9) $$\Pr\left[A3_i\right]=\Pr\left[-\frac{\Delta}{-\beta_C}\le -DC_i-\frac{\beta_Q}{\beta_C}DQ_i+\frac{\varepsilon_i}{-\beta_C}\le\frac{\Delta}{-\beta_C}\right]$$

A simple but inappropriate assumption would be that the ε_i are independent. But clearly there is dependence between the events. One way to introduce that is to model the dependencies between the ε_i. This approach involves difficult integrals that require simulation. Simulation in and of itself is not serious drawback. The problem here is that the regions of integration are extremely complicated: intersections of cones. We may pursue that approach in a latter paper. For now we use a simpler approach that loses little generality.

We allow the triple $\left\{\beta_C,\beta_Q,\Delta\right\}$ to be a random vector, fixed and constant for each individual, but varying over individuals to pick up random taste variations. Also assume that conditional on $\left\{\beta_C,\beta_Q,\Delta\right\}$,the ε_i are iid logistic random variables. Let

$$\begin{aligned}\delta_{ij}&=1 \text{ if } i \text{ is chosen at stage } j\\ &=0 \text{ otherwise}\end{aligned}$$

then the likelihood of an outcome $\left\{\delta_{1,1},\cdots,\delta_{3,11}\right\}$ conditional on $\left\{\beta_C,\beta_Q,\Delta\right\}$ is

(1.10) $$\begin{aligned}L&=\prod_{j=1}^{11}P\left[A1_j\right]^{\delta_{1j}}P\left[A2_j\right]^{\delta_{2j}}P\left[A3_j\right]^{\delta_{3j}}\\ &=\prod_{j=1}^{11}P\left[DC_j\beta_C+DQ_j\beta_Q+\varepsilon_j>\Delta\right]^{\delta_{1j}}\\ &\qquad\cdot P\left[DC_j\beta_C+DQ_j\beta_Q+\varepsilon_j\le-\Delta\right]^{\delta_{2j}}\cdot P\left[-\Delta<DC_j\beta_C+DQ_j\beta_Q+\varepsilon_j\le\Delta\right]^{\delta_{3j}}\end{aligned}$$

There are also sign restrictions required by economics and to prevent zero probabilities during iterations. Specifically, $\beta_C\le 0,\ \beta_Q\ge 0,$ and $\Delta\ge 0.$ So $\beta_C=-\left|\alpha_C\right|$, $\beta_Q=\left|\alpha_Q\right|,\Delta=\left|\alpha_\Delta\right|$. In addition, we allow for the possibility that these coefficients depend on observable characteristics of the respondents specifically, where the x_i are the explanatory variables for the corresponding coefficient. Of course, these explanatory variables may be the same, overlap or be completely different, the estimation will tell. If we further assume that the α are jointly $N\left(Bx,\Omega\right)$, then the marginal distributions of $\left\{\beta_C,\beta_Q,\Delta\right\}$ are folded normals, which are proportional to non-central chi-squares. So for example, $\beta_C \sim -\sigma_C \cdot \chi^2_{(1)}\left(\zeta^2_C\right)$, the negative of a scaled non-central chi-square with scale $\sigma_C = Var\left(\alpha_C\right)$ and non-centrality parameter $\zeta^2_C = \left(\frac{x_C^{'}\mu_C}{\sigma}\right)^2$.

We recall for future reference that if $Y = \left|Z\right|$ and $Z \sim N\left(\xi,\sigma^2\right)$, then

(1.11)
$$E\left(Y\right) = \sigma\left(1+\left(\frac{\xi}{\sigma}\right)^2\right)$$
$$V\left(Y\right) = 2\sigma^2\left(1+2\left(\frac{\xi}{\sigma}\right)^2\right)$$

Thus the likelihood can be written

(1.12)
$$\begin{aligned} L = \int\Bigg\{\prod_{j=1}^{11}\Bigg(& P\left[DC_j\beta_C + DQ_j\beta_Q + \varepsilon_j > \Delta\right]^{\delta_{1j}} \\ & \cdot P\left[DC_j\beta_C + DQ_j\beta_Q + \varepsilon_j \le -\Delta\right]^{\delta_{2j}} \\ & \cdot P\left[-\Delta < DC_j\beta_C + DQ_j\beta_Q + \varepsilon_j \le \Delta\right]^{\delta_{3j}}\Bigg)\Bigg\}\phi\left(\alpha;\mu,\Omega\right)d\alpha \end{aligned}$$

Where the β are the explicit functions of the α and the *x* as above.

(1.13)
$$\begin{aligned} L = \int\Bigg\{\prod_{j=1}^{11}\Bigg(& 1 - F\left[\left|b_\Delta x_\Delta + \omega_\Delta\right| - DC_j\left|b_C x_C + \omega_C\right| - DQ_j\left|b_Q x_Q + \omega_Q\right|\right]^{\delta_{1j}} \\ & \cdot F\left[-\left|b_\Delta x_\Delta + \omega_\Delta\right| - DC_j\left|b_C x_C + \omega_C\right| - DQ_j\left|b_Q x_Q + \omega_Q\right|\right]^{\delta_{2j}} \\ & \cdot\Big(F\left[\left|b_\Delta x_\Delta + \omega_\Delta\right| - DC_j\left|b_C x_C + \omega_C\right| - DQ_j\left|b_Q x_Q + \omega_Q\right|\right] \\ & - F\left[-\left|b_\Delta x_\Delta + \omega_\Delta\right| - DC_j\left|b_C x_C + \omega_C\right| - DQ_j\left|b_Q x_Q + \omega_Q\right|\right]\Big)^{\delta_{3j}}\Bigg)\Bigg\}\phi\left(\omega;\Omega\right)d\omega \end{aligned}$$

We will use simulation to calculate the required integrals, by drawing ω from a multivariate $N\left(0,\Omega\right)$.

We noted that the assumption of independence was innocuous. This is because large values of the parameters turn the probabilities above into indicator functions for the events. To see this write (1.10) as

$$(1.14) \qquad \prod_{j=1}^{11} P\left[DC_j\beta_C + DQ_j\beta_Q - \Delta > -\varepsilon_j\right]^{\delta_{1j}}$$
$$\cdot P\left[DC_j\beta_C + DQ_j\beta_Q + \Delta \le -\varepsilon_j\right]^{\delta_{2j}}$$
$$\cdot P\left[-DC_j\beta_C - DQ_j\beta_Q - \Delta < \varepsilon_j \le -DC_j\beta_C - DQ_j\beta_Q + \Delta\right]^{\delta_{3j}}$$

through a reparameterization write this as

$$(1.15) \qquad \prod_{j=1}^{11} P\left[\tau\left(DC_j\beta'_C + DQ_j\beta'_Q - \Delta'\right) > -\varepsilon_j\right]^{\delta_{1j}}$$
$$\cdot P\left[\tau\left(DC_j\beta'_C + DQ_j\beta'_Q + \Delta'\right) \le -\varepsilon_j\right]^{\delta_{2j}}$$
$$P\left[\tau\left(-DC_j\beta'_C - DQ_j\beta'_Q - \Delta'\right) < \varepsilon_j \le \tau\left(-DC_j\beta'_C - DQ_j\beta'_Q + \Delta'\right)\right]^{\delta_{3j}}$$

or

$$(1.16) \qquad \prod_{j=1}^{11} P\left[DC_j\beta'_C + DQ_j\beta'_Q - \Delta' > \frac{-\varepsilon_j}{\tau}\right]^{\delta_{1j}}$$
$$\cdot P\left[DC_j\beta'_C + DQ_j\beta'_Q + \Delta' \le \frac{-\varepsilon_j}{\tau}\right]^{\delta_{2j}}$$
$$\cdot P\left[-DC_j\beta'_C - DQ_j\beta'_Q - \Delta' < \frac{\varepsilon_j}{\tau} \le -DC_j\beta'_C - DQ_j\beta'_Q + \Delta'\right]^{\delta_{3j}}$$

Now for any distribution function $F(x)$ with median at zero, we have

$\lim_{\tau\to\infty} F\left(\frac{x}{\tau}\right) = H(x)$, the Heaviside (or step)function. This is nothing more than the

indicator of the event. Thus, τ is large implies: $(\alpha \cdot \tau)$ is large, implies: the coefficients in the probabilities would be large. So in the τ limit, (1.16) is nothing more than the indicator of a sequence of events with random coefficients. Its integral with respect to the coefficients, again in the limit, is a probability without but without an independence assumption. Thus we can allow the data to tell us whether the conditional independence assumption is appropriate or not, large coefficients diminish the effect of the assumption.

4.Estimation

As mentioned, to calculate the probabilities, we must simulate. Simulated maximum-likelihood was attempted but did not converge, moreover, simulated maximum-likelihood (using a plug-in estimator of the log probability) requires huge numbers of simulants. As a consequence we took a Simulated Method of Moments (SMM) approach.

Any random variable can be decomposed into an expectation conditional on other variables, denoted z, plus an error with zero expectation conditional on z. Because of this the error is also uncorrelated with z and any function of the z. Let $z \in R^k$, let $g : R^k \to R^m \;\; m \le k$. Then we have the following moment conditions (see Hall(2005)

$$E\left(g(z)\left(\delta_{ij} - \Pr\left(\delta_{ij} = 1 \mid z\right)\right)\right) = 0, \quad j = 1,2,3; \; i = 1,\ldots,11; \tag{1.17}$$

and estimation can proceed by SMM (McFadden(1989). That is if

$$\phi(\theta)_t = g(z_t)\left(\delta_{ijt} - \Pr\left(\delta_{ijt} = 1 \mid z_t\right)\right) \tag{1.18}$$

where θ is the vector of all parameters, we find parameter estimates to minimize

$$\sum_{t=1}^{T} \phi_t(\theta)' \hat{W} \sum_{t=1}^{T} \phi_t(\theta) \tag{1.19}$$

where $\hat{W}$ is the optimal weighting matrix (see Cameron Trivedi (2005)).
The advantages of SMM are: first, while the probabilities must be simulated, the moments can be estimated in an unbiased fashion. Second, the number of simulants required is small, as suggested by Gourieroux and Monfort (2002) we use 10. Finally, with only a specification of the model, it can be directly estimated with existing software. We used SAS, Proc Model with the GMM and NDRAW options. The covariance matrix for the random coefficients Ω must be estimated as well. We parameterized this as $\Omega = H^T H$ where the H is lower triangular. For convergence we restricted the first row to a vector of ones and also $H_{ij} = 1$, for $i = 1$, $j = 1,2,3$ and $i = 2$, $j = 2$. We provide tests of the restrictions as well, although there are directions in which these tests must be powerless because some are needed for identification.

We use as instruments: the personal characteristics of the individuals, the cost and quality differences, and the interactions of these characteristics with the cost and quality differences. We did not use the same explanatory variables for the $\{\beta_C, \beta_Q, \Delta\}$ but tested down to find a parsimonious set. We also did not use, as moments conditions, any involving the impossible probabilities; only 22 of the events for which probabilities can be written can be positive. The results are presented in Table 1.

{Table 1 here}

It is difficult to interpret the coefficients if for no other reason than the appearance of the absolute value. Our interest is in the willingness to pay. We define it as the change in cost arising from a change in quality that would hold utility constant. That is,

(1.20) $$wtp_t = \frac{dC_t}{dQ_t} = -\frac{\partial V_t / \partial Q}{\partial V_t / \partial C} = -\frac{\beta_{Qt}}{\beta_{Ct}}.$$

We approach this in two ways. First we use these results to examine simulate the β using estimated coefficients, estimated variances, the explanatory variables and simulated errors and calculate the willingness to pay for each person in the sample. We then treat the simulated willingness to pay as a simple univariate random variable to be statistically analysed. We do this using Proc Univariate in SAS.

{Table 2 here}

{Histogram here}

We find the willingness to pay distribution has a mean of about $9 dollars, and ranges between 0 and $4000. There are less than 10 with these super values. They most likely are an anomaly of the simulation. A small number of people <1% have a value > $100.

Second, we utilize the parameterization implicit in equations (1.7)-(1.9). For interpretability, we take the *wtp* parameter to be a linear combination of the respondent characteristics. This model therefore differs from the previous model in two ways, the aforementioned one, measuring wtp directly, and whether the error in the first model or this model is heteroscedastic. These results appear in Table 1A. It must be that one or the other model is heteroscedastic because the coefficient on ΔC, a function of respondent characteristics, divides the event defining the outcome. In each case we define the residual error as homoscedastic.

4. Evaluation of the Panel Sample

Sample Characteristics

The sample used for the study came from the Knowledge Networks (Menlo Park, CA, www.knowledgenetworks.com) panel. Researchers on environmental benefits valuations have increased their use of internet panels, so that the performance of this survey approach has broad implications beyond our particular study. The Knowledge Networks sample consists of a national sample of households recruited by random-digit dialing, who either have been provided internet access through their own computer or are given a WebTV console.

Sample Validity Tests

Because the survey was administered via the internet using an existing panel of respondents, we undertook a series of validity tests specifically determining whether their panel membership influenced the valuation results. To the best of our knowledge, these are the first such tests to have been undertaken for this sampling methodology. We tested the panel influences of four variables on the regression analysis of the determinants of the value of water quality benefits.

The first variable is whether the respondent stopped the survey and then continued the survey at a later time. Conceivably, such respondents might be less engaged in the survey task. However, there was no significant effect of this variable on benefit values.

The second variable of interest is the time the respondent has been a member of the Knowledge Network panel. Length of time in the panel may affect attentiveness to surveys and potentially could be correlated with other personal characteristics that influence water quality valuations. The estimates in Table 1 fail to indicate any significant influence of this variable either.

Third, the number of days the respondent took to complete the survey after being offered the opportunity to participate could reflect a lack of interest in the survey topic or in taking surveys generally. Nevertheless, there is no significant effect of this variable on benefit valuations in either of the equations estimated in Table 1

The final survey methodology variable tested is whether the respondent subsequently quit the panel either immediately after the invitation for this survey or at any later time until May 2004, when data for this variable were collected. Such respondents could be less interested in taking surveys and might have different valuations. However, this variable was also not statistically significant in the water quality valuation equations.

Overall, there is no indication that any of these key aspects of the panel methodology bias the survey responses. In addition to the general match of our respondents to the U.S. population, we also examined whether these four variables reflecting the methodology had any influence on the probability that the respondent failed to pass the consistency test with respect to the benefit valuations. There were no significant effects of any of the Knowledge Networks panel variables so that there is no evidence that national performance of the survey task is importantly influenced by any of these variables.

Selection Effects

Of 1,587 panel members invited to take the survey, 74% of respondents chose to participate. Of the 1,174 participants, three respondents did not complete the portion of the survey that elicits water quality value. Finally, 6% of participants completed the survey but were dropped because they chose the dominated alternative and continued with that choice even after being so informed. Therefore, 1,103 of 1,587 invitees consistently completed the water quality valuation portion of the survey. For the selection correction for bias, we used variables for which we had the values for non-respondents as well as survey respondents. This data is routinely collected by Knowledge Networks on its panel members. Thus, an additional advantage of such panels is that there is information available to analyze who chose not to take the survey after being offered the chance to do so.

To predict participation, we identified a number of variables that significantly affected survey completion. In particular, we found that being African American or Hispanic was negatively associated with completing the survey, as was household size. We also constructed two health-related stress dummy variables. The first stress variable was for individuals who reported that they had a high stress level. The high stress variable indicated respondents who reported more “stress, strain, or pressure” than usual “during the past few months.” The second stress variable was for people who failed to respond to the stress information question. Each of these variables was negatively related to the probability of taking the survey but not significantly related to the water quality valuation amount V, thus achieving the appropriate identification.

5. Conclusions

The survey results presented here passed a variety of consistency tests and rationality checks. These tests included dominance tests as part of the iterative choice process and external scope tests across respondents. In addition, the internet-based methodology itself was tested with respect to a variety of potential sources of bias, such as sample attrition, and these panel characteristics had no significant effect on the results.

It is appropriate to speculate on why these results are much stronger than those reported in Magat et al. (1988). The earlier study produced similar aggregate values, but the covariates with water quality value were largely insignificant, and a scope test was not even attempted. The Magat et al (1988) study had less than half the number of respondents, but the main differences are methodological. In the current study, greater effort was placed on preparing the respondent to make the trade-off between water quality and cost of living. Three warm up questions involving dominated choices provided easy ways to understand the choice task, and for the relatively low percent of respondents who 'failed' those questions, provided a way to communicate the importance of their answers.

Working with a panel had several advantages. First, since our survey design involved the use of a computer-based sample, the Knowledge Networks panel may have yielded a more representative sample of survey participants than other survey methods such as those used by Magat et al. (2000) in which a group of subjects contacted by phone came to a central location to take the survey. Second, respondents in the panel are accustomed to taking surveys, so they are not confused by the process. Third, and most important, because there are data on those who declined to take the survey, it is possible to estimate the impact of that self-selection on our results. In this case, that self-selection had minimal effect on our estimates. However, that result strictly applies only to our focal question about the value of water quality. The real value of panels is that they contain the information that permits an assessment of the impact of respondent selection mechanism that will certainly be an even greater problem in the future.

The practical benefit of these results is that they provide unit water quality benefit values that can be matched to existing EPA measures of water quality to provide an assessment of benefits of water quality programs. Good water quality has a median unit value of $2.31 per percentage point increase in water quality. This value is dependent on variables such as income, education, and personal use of lakes and rivers in the expected fashion. To value water quality improvements, one can use these values in conjunction with results that break down the benefits in terms of benefits for the components of water quality—fishing, swimming, and health of the aquatic environment— to gauge the economic benefit of an improvement project to the affected local population.

References:

Bajari, Patrick and Ali Hortaçsu. (2003). "The Winner's Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions,"*The RAND Journal of Economics*, Vol. 34, No. 2, pp. 329-355.

Bajari, Patrick and Ali Hortaçsu. (2003). "Economic Insights from Internet Auctions" *Journal of Economic Literature* Vol. XLII pp. 457–486

Berkman, Richard and W. Kip Viscusi. (1973). *Damming the West*. New York: Grossman Publishers.

Berrens, Robert P., Alok K. Bohara, Hank C. Jenkins-Smith, Carol L. Silva, and David L. Weimer. (2004). "Information and Effort in Contingent Valuation Surveys: Application to Global Climate Change Using National Internet Samples," *Journal of Environmental Economics and Management*, 47(2), 331-363.

Bishop, Richard C. and Michael P. Welsh. (1992). "Existence Values in Benefit-Cost Analysis and Damage Assessment," *Land Economics*, 68(4), 405-417.

Bishop, Richard C. and Thomas A. Heberlein. (1990). "The Contingent Valuation Method." In Rebecca L. Johnson and Gary V. Johnson (eds.), *Economic Valuation of Natural Resources: Issues, Theory, and Applications.* Social Behavior and Natural Resources Series. Boulder: Westview Press.

Cameron, Colin and Pravin Trivedi. (2005) *Microeconometrics: Methods and Applications*, Cambridge University Press

Cameron, Trudy, and Jeffrey Englin. (1997). "Respondent Experience and Contingent Valuation of Environmental Goods," *Journal of Environmental Economics and Management*, 33(3), 296-313.

Carson, Richard T. and Mitchell, Robert Cameron. (1993). "The Value of Clean Water: The Public's Willingness to Pay for Boatable, Fishable, and Swimmable Quality Water," *Water Resources Research*, 29(7), 2445-2454.

Cardell, Scott. (1977). "A Logit Specification with a Random Distribution of Preferences," mimeo.

Chapman R.G. and Richard Staelin. (1982). Exploiting Rank Ordered Choice Set Data Within the Stochastic Utility Model. *Journal of Marketing Research* 19,3,288-301.

DeShazo, J.R., and German Fermo. (2002). "Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency," *Journal of Environmental Economics and Management*, 44(1), 123-143.

DeShazo, J.R., and Trudy Cameron. (2004). "Moving Beyond the Value of a Statistical Life: A New Theoretical and Empirical Analysis." Paper presented at 2004 AEA meetings.

Duncan, Gregory M. (1980). "Formulation and Statistical Analysis of the Mixed Continuous/Discrete Dependent Variable Problem in Classical Production Theory", *Econometrica* 48:839-852.

Duncan G.M. and D. Perry. (1989). "Sequential Querying: A New Approach to Estimating Willingness to Pay for New Services," *GTE Laboratories Technical Report 89-06-009.* Waltham, MA

Duncan, Gregory M. (1992). “A Common Probit/Random Utility Framework For Conjoint Experiments: Formulation, Efficient Estimation, and Inference,” *GTE Laboratories Technical Report 92-11-382.* Waltham, MA
Duncan G.M. and D. Perry. (1995). “Models for Completely, Partially and Contingently Ranked Data,” *GTE Laboratories Technical Report 95-09-001.* Waltham, MA
Fischhoff, Baruch and Lita Furby. (1988). “Measuring Values: A Conceptual Framework for Interpreting Transactions with Special Reference to Contingent Valuation of Visibility,” *Journal of Risk and Uncertainty*, 1(2), 147-184.
Fligner, Michael A. and Joseph Verducci (Eds.) (1992) *Probability Models and Statistical Analyses for Ranking Data*. Springer-Verlag Lecture Notes in Statistics, Volume 80.
Gourieroux, Christian and Alain Monfort. (2002) *Simulation Based Econometric Methods,* Oxford University Press.
Grice, H.P. (1975) “Logic and Conversation” In PCole and J.P. Forgas (eds.) *Syntax and Semantics: 3. Speech Acts (*p. 41-58*),* New York, Academic Press.
Hall, Alastair (2005) *Generalized Method of Moments*, Oxford University Press, USA
Hausman J.A. and Paul Ruud, (1987) “Specifying and Testing Econometric Models for Rank-Ordered Data,” *Journal of Econometrics* 34, 83-104.
Hausman J.A. and David A. Wise (1978) “A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences,” *Econometrica* 46:403-426.
Huber, Joel, W. Kip Viscusi, and Jason B. Bell. (2004) “Coping with the Contingency of Valuation: Range and Anchoring Effects in Choice Valuation Experiments,” Working Paper, Fuqua School of Business, Duke University.
Luce, R. Duncan, (1959) *Individual Choice Behavior*, New York, Wiley.
Kahneman, Daniel and Amos Tveski (1979) “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica, 47*, 263-291.
Krupnick, Alan, Anna Alberini, Maureen Cropper, Nathalie Simon, Bernie O’Brien, Ron Goeree, and Martin Heintzelman. (2002). “Age, Health and Willingness to Pay for Mortality Risk Reductions: A Contingent Valuation Survey of Ontario Residents,” *Journal of Risk and Uncertainty*, 24(2), 161-186.
Magat, Wesley A., Joel Huber, W. Kip Viscusi, and Jason Bell. (2000). “An Iterative Choice Approach to Valuing Clean Lakes, Rivers and Streams.” *Journal of Risk and Uncertainty*, 21:1, 7-43.
Magat, Wesley A., W. Kip Viscusi, and Joel Huber. (1988). “Paired Comparison and Contingent Valuation Approaches to Morbidity Risk Valuation,” *Journal of Environmental Economics and Management*, 15, 395-411.
McFadden, D. (1989). “A Method of Simulated Moments for Estimation of the Multinomial Probit Without Numerical Integration,” *Econometrica* 57:995-1026.
Mitchell, Robert Cameron and Richard T. Carson. (1989). *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington: Resources for the Future.
Paarsch, Harry J. and Han Hong. (2006). *An Introduction to the Structural Econometrics of Auction Data.* Cambridge: The MIT Press.
Schkade, David A. and John W. Payne. (1993). "Where Do the Numbers Come from? How People Respond to Contingent Valuation Questions." In Jerry A. Hausman

(ed.), *Contingent Valuation: A Critical Assessment*. New York: North-Holland, 271-293.
Smith, V. Kerry, and William H. Desvousges. (1986). *Measuring Water Quality Benefits*. Boston: Kluwer Academic Publishers.
Smith, V. Kerry, and Laura L. Osborne. (1996). "Do Contingent Valuation Estimates Pass a 'Scope' Test? A Meta-Analysis," *Journal of Environmental Economics and Management*, 31(3), 287-301.
United States Environmental Protection Agency. (1994). *National Water Quality Inventory, 1992 Report to Congress*, EPA 841-R-94-001.
Viscusi, W. Kip, Joel Huber, and Jason Bell. (2004) "The Value Of Regional Water Quality Improvements", Harvard Law School Discussion Paper No. 477

Figure 1
Sample Private Water Quality Benefit Question

<table>
<tr><td colspan="4">We would like to ask you some more questions like these. However, in these questions, one region will have a lower annual cost of living and the other will have higher water quality. Remember that the national average for water quality is 65% Good.</td></tr>
<tr><td colspan="2">Region 1</td><td colspan="2">Region 2</td></tr>
<tr><td>Increase in
Annual Cost
Of Living</td><td colspan="2">$100
More
Expensive</td><td>$300
More
Expensive</td></tr>
<tr><td>Percent of Lake
Acres and River
Miles With Good
Water Quality</td><td colspan="2">40%
Good
Water
Quality</td><td>60%
Good
Water
Quality</td></tr>
<tr><td>Which Region
Would you Prefer?</td><td>Region 1
*</td><td>Region 2
*</td><td>No Preference
*</td></tr>
</table>

Table 1A
SGMM Results

Parameter	Estimate	Std Err	t Value	Pr > |t|
bc_1	0.137129	0.000109	1252.37	<.0001
bc_BLACK	0.01303	0.00002	656.34	<.0001
bc_NOSMOKE	0.004648	0.000011	435.14	<.0001
bc_FEDPCT	0.000099	3.88E-07	253.9	<.0001
bc_FEMALE	0.009462	0.000013	708.51	<.0001
bc_FILTER	-0.00323	0.000013	-240.99	<.0001
bc_FULLTIME	0.014889	0.000011	1378.77	<.0001
bc_GREEN	0.011533	0.000014	851.82	<.0001
bc_HHSIZE	-0.0039	4.31E-06	-904.42	<.0001
bc_LOGINC	-0.00683	9.70E-06	-704.07	<.0001
bc_YEAREDU	-0.00524	3.14E-06	-1667.9	<.0001
bq_1	0.122306	0.00062	197.17	<.0001
bq_BLACK	0.036929	0.000189	195.12	<.0001
bq_NOSMOKE	0.022158	0.000144	153.86	<.0001
bq_FEDPCT	-0.00025	2.74E-06	-90.6	<.0001
bq_FEMALE	0.014737	0.000144	102.17	<.0001
bq_FILTER	-0.02361	0.000137	-171.97	<.0001
bq_FULLTIME	-0.02197	0.000164	-134.3	<.0001
bq_GREEN	-0.51631	0.000205	-2520.8	<.0001
bq_HHSIZE	-0.0063	0.000041	-152.51	<.0001
bq_LOGINC	-0.00229	0.000062	-36.94	<.0001
bq_YEAREDU	0.005721	0.00003	188.75	<.0001
bd_1	0.257356	0.00183	140.26	<.0001
bd_BLACK	0.067043	0.000432	155.16	<.0001
bd_NOSMOKE	0.024502	0.000312	78.66	<.0001
bd_FEDPCT	-0.00039	5.69E-06	-67.95	<.0001
bd_FEMALE	-0.03351	0.00021	-159.72	<.0001
bd_FILTER	-0.01199	0.000181	-66.23	<.0001
bd_FULLTIME	-0.03564	0.000276	-129.34	<.0001
bd_GREEN	0.189742	0.000196	968.75	<.0001
bd_HHSIZE	-0.00548	0.000119	-45.97	<.0001
bd_LOGINC	0.001971	0.000167	11.8	<.0001
bd_YEAREDU	0.000509	0.000055	9.21	<.0001
hh4	-0.12966	0.000131	-989.89	<.0001
hh5	1.143571	0.00039	2932.29	<.0001
hh6	0.242955	0.000211	1153.87	<.0001

```
Nobs 4037   Objective        0.2529
            Objective*N        1021
```

Table 1A
SGMM Results

Parameter	Estimate	Approx Std Err	t Value	Approx Pr > |t|
bq_1	-13.3606	0.0205	-651.69	<.0001
bq_BLACK	-2.41908	0.00202	-1196.7	<.0001
bq_NOSMOKE	-7.00021	0.00354	-1977.2	<.0001
bq_FEDPCT	0.043751	0.00008	524.55	<.0001
bq_FEMALE	-3.2033	0.00295	-1085.2	<.0001
bq_FILTER	3.989688	0.00311	1281.2	<.0001
bq_FULLTIME	6.850559	0.00188	3644.88	<.0001
bq_GREEN	-1.63981	0.00516	-317.62	<.0001
bq_HHSIZE	0.317785	0.00111	287.42	<.0001
bq_LOGINC	2.151586	0.00168	1279.48	<.0001
bq_YEAREDU	-0.73422	0.00049	-1521.4	<.0001
bd_1	-0.50266	0.4043	-1.24	0.2139
bd_BLACK	0.473811	0.0749	6.33	<.0001
bd_NOSMOKE	-0.19635	0.0771	-2.55	0.0109
bd_FEDPCT	0.001535	0.00222	0.69	0.4897
bd_FEMALE	-0.26659	0.067	-3.98	<.0001
bd_FILTER	0.348152	0.093	3.74	0.0002
bd_FULLTIME	-0.23224	0.0518	-4.48	<.0001
bd_GREEN	0.694199	0.1337	5.19	<.0001
bd_HHSIZE	-0.30829	0.0268	11.51	<.0001
bd_LOGINC	0.35894	0.0427	8.41	<.0001
bd_YEAREDU	0.218195	0.0127	7.22	<.0001
hh1	14.68735	0.0054	2722.13	<.0001
hh2	18.239	0.00758	2404.64	<.0001

Nobs	3258	Objective	0.2396
		Objective*N	780.6945

Table 2
WTP Analysis

N	3258	Sum Weights	3258
Mean	9.26362307	Sum Observations	30180.884
Std Deviation	91.8929843	Variance	8444.32056
Skewness	44.0920666	Kurtosis	2235.20578
Uncorrected SS	27782736.4	Corrected SS	27503152.1
Coeff Variation	991.97672	Std Error Mean	1.60992934

Basic Statistical Measures

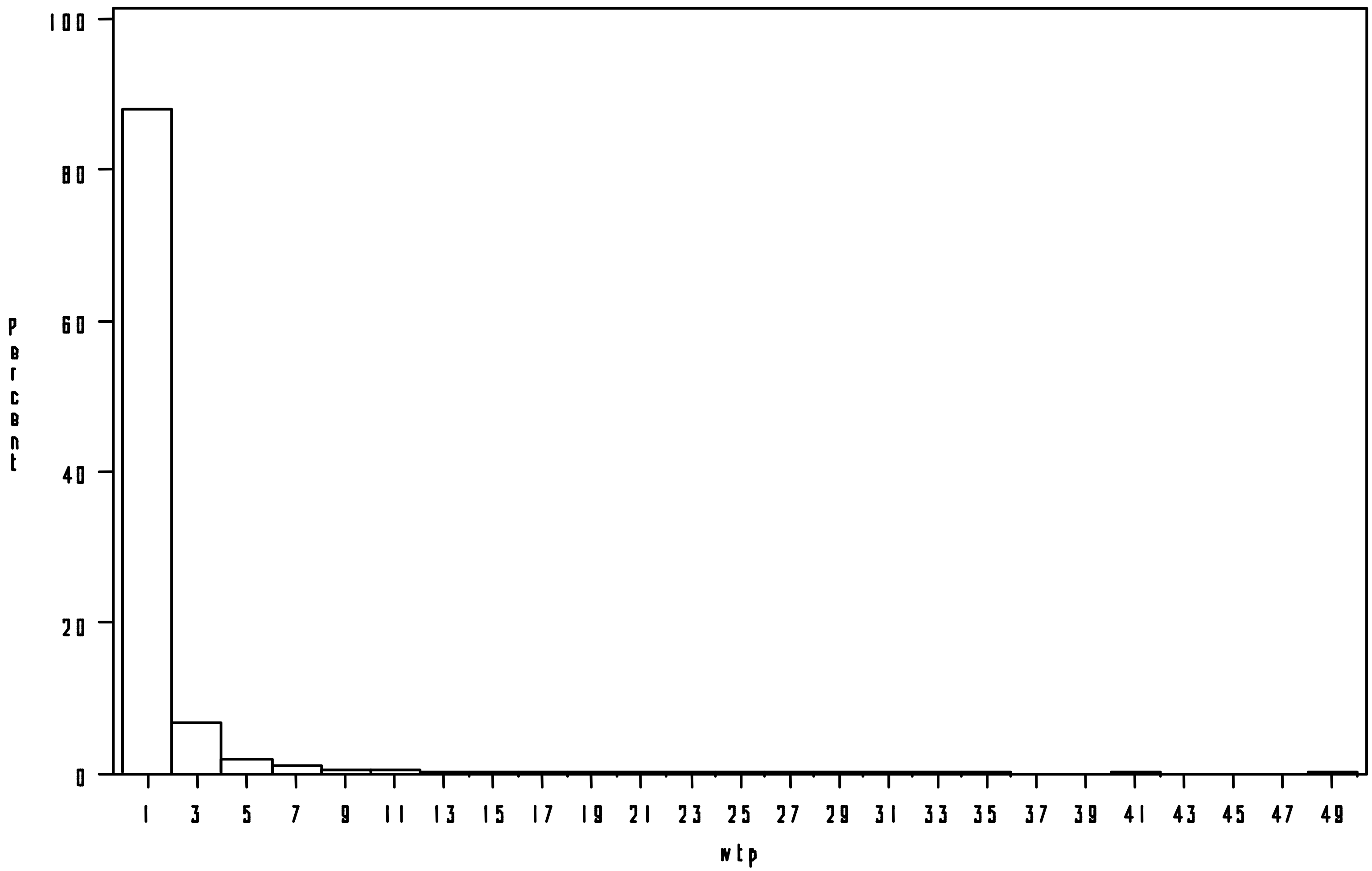
Location		Variability	
Mean	9.263623	Std Deviation	91.89298
Median	2.307592	Variance	8444
		Range	4776
		Interquartile Range	3.34578

Quantiles

Quantile	Estimate
100% Max	4.77608E+03
99%	1.13723E+02
95%	2.10014E+01
90%	1.08433E+01
75% Q3	4.50685E+00
50% Median	2.30759E+00
25% Q1	1.16106E+00
10%	4.64418E-01
5%	2.56118E-01
1%	4.76428E-02
0% Min	2.91176E-03

Extreme Observations

-------Lowest-------		------Highest-----	
Value	Obs	Value	Obs
0.00291176	3212	648.122	1420
0.00467253	7	698.802	1737
0.00636442	1306	741.465	2826
0.00762623	1235	1127.060	2107
0.00961946	1915	4776.083	560



100
80
60
40
20
0
percent
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49
wtp

WTP Regression Analysis

Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	2.57260	0.07345	35.02	<.0001
FEMALE_VISITED	0.23871	0.09112	2.62	0.0088
tld_stq_lakeper	-0.00018862	0.0000869	-2.17	0.0300
lakeper_LOGINC	0.00418	0.00209	2.00	0.0455
HHSIZE_VISITED	0.05426	0.02713	2.00	0.0455
WEST	1.26131	0.71137	1.77	0.0763
VISITED_LOGINC	-0.01730	0.00999	-1.73	0.0833
HHSIZE_FEMALE	-0.04255	0.02496	-1.70	0.0883
n_east_lakeper	-0.01444	0.00900	-1.60	0.1087
west_HHSIZE	-0.06998	0.04481	-1.56	0.1184
LAKEPER	-0.03265	0.02257	-1.45	0.1481
west_LOGINC	-0.08789	0.06647	-1.32	0.1862
south_tld_stq	0.00156	0.00137	1.13	0.2566
tld_stq_GREEN	-0.00347	0.00311	-1.11	0.2654
n_east_tld_stq	0.00279	0.00268	1.04	0.2980
west_lakeper	-0.01552	0.01673	-0.93	0.3537
lakeper_FEMALE	-0.00441	0.00476	-0.93	0.3540
N_EAST	0.11994	0.13941	0.86	0.3897
tld_stq_HHSIZE	0.00033149	0.00038574	0.86	0.3902
lakeper_GREEN	-0.00606	0.00843	-0.72	0.4721

Variable List

AGE	age
BLACK	Race, Black
BOTTLED	Drink Bottled Water most often
DIDSMOKE	Former Cigarette Smoker
FEDPCT	Percent of State Area Federal Land
FEMALE	Gender, Female
FILTER	Drink Filtered Water most often
FULLTIME	Employed Full Time
GREEN	Member of at least one listed organization
HHSIZE	Size of household
HICOST	Starting difference in cost of living between regions
HILEVEL	Starting difference in water quality between regions
HISPANIC	Ethnicity, Hispanic
LAKEPER	Lake Acers per Square Mile of State Area
LOGINC	Natural Log of income
MEANM	Mean elevation of State
NOSMOKE	Never Smoked Cigarettes
NOTWHITE	Race, not white
N_EAST	Region, Northeast
OTHER	Race, Non-Black, Non-White
RETIRED	Retired
RIVPER	River miles per square mile of state area
SOUTH	Region, South
STQLAK	State Lake Quality
TAPWATER	Drink tap water most often
TENURE	Panel tenure in months
TOLD	Told national water quality average
URBAN	Percent of State that is urban
URBANID	Self-Identify as living in urban setting (vs suburban or rural)

VISITED	Visited lake or river in last 12 months
WEST	Region, West
WORK	Currently employed
YEAREDU	Years of education
YESSMOKE	Current Cigarette Smoker