C9

THE USE OF SIMULATED POLITICAL MARKETS

TO VALUE PUBLIC GOODS

Richard T. Carson
University of California, San Diego

W. Michael Hanemann
University of California, Berkeley

Robert Cameron Mitchell
Resources for the Future

February 1987

The problem of determining the demand for public goods has long intrigued economists. Since markets do not generally exist for the sale and purchase of public goods, researchers cannot directly observe market behavior with respect to these goods. To overcome this problem, two approaches, one based on indirect observed behavior and another on direct hypothetical behavior, have been used to infer the benefits of unmarketed goods. Following economists' well-known preference for observed behavioral indicators, the indirect approach estimates elasticities or consumer surplus/willingness-to-pay benefit measures from observing consumer actions such as recreational travel patterns (Clawson and Knetsch, 1966; Vaughan and Russell, 1982), and housing purchases/prices, (Harrison and Rubinfeld, 1978), and from observing political behavior such as the decisions of elected officials (Borcherding and Deacon, 1972; Bergstrom and Goodman, 1973), and bureaucrats (McFadden, 1976). Any of the indirect-observed behavior techniques suffer from a number of serious limitations. The four most important tend to be: (1) the necessarily arbitrary chain of assumptions required to translate the behavior into values for specific amenity improvements; (2) the inability to measure certain important types of benefits; (3) in the presence of uncertainty, the measurement of *ex post* expected welfare measures rather than the correct *ex ante* welfare measures; and (4) the restriction to situations for which relevant and observable behavioral data are available.

The direct hypothetical methods attempt to circumvent these problems by using surveys, that is they describe desired improvements to randomly selected respondents and ask them how much they are willing to pay for these amenities. In some cases (Courant, Gramlich, and Rubinfeld, 1980; Bergstrom, Rubinfeld, and Shapiro, 1982), the survey questions resemble, in their simplicity and lack of detail, those used in conventional public opinion/voter surveys. In the case of the more widely used contingent valuation (CV) method (e.g., Randall, Ives, and Eastman, 1974; Brookshire, *et al.*, 1982; Jones-Lee, Hammerton, and Philips, 1985), however, a scenario is developed which presents a well-defined market in which a good, whose characteristics are described in substantial detail, can be purchased.[1]

[1] The term contingent valuation is used to refer to the fact that the result obtained are "contingent" on the particular scenario presented to respondents in the survey exercise. See Mitchell and Carson (forthcoming, Appendix A) for an extensive list and brief description of over 100 CV surveys).

A decade of intensive research on the CV method has demonstrated its theoretical validity, its flexibility, and its ability to measure a broader range of values (Cummings *et al.*, 1986). It is now generally accepted that economists' early fears that CV findings would be affected by strategic behavior were largely unfounded, because respondents lack the necessary information and motivation to act in this manner given the hypothetical character of the valuation process.[2] The fact that CV studies are hypothetical, however, makes many economists skeptical of estimates obtained by the CV method. Their concerns appear to be two-fold. First, they doubt that people can meaningfully value an unfamiliar good, such as air visibility or national freshwater quality, during the short space of a personal interview, as consumers typically take days, weeks, or even months to make a major durable good purchase decision. Second, they find it difficult to believe that respondents will be motivated to make the effort necessary to arrive at accurate estimates even if it were possible for them to do so. Another concern, which is raised more by cognitive psychologists and sociologists than by economists, is that the respondents will be vulnerable to various suggestive features of the scenario posed (Kahneman, Slovic, and Tversky, 1982). Considerations such as these led Freeman to conditionally dismiss the usefulness of the CV method:

> Thus, in the absence of evidence showing that people can and do make accurate assessments of their preferences in these situations, the results of these surveys should not be given great weight by decision makers (Freeman, 1979; 104).

In this paper, we will argue that the appropriate market model for CV surveys to emulate, in many cases, particularly those involving pure public goods, is a political market rather than a private goods market, as has been typical of most CV surveys to date. We propose a simple and relatively inexpensive variant of the CV method using survey questions to simulate referenda at different tax prices, which can provide a complete description of the public's demand curve for the good in question. In doing so, we draw upon three disparate literatures: voting results in public finance, CV surveys in environmental economics, and stimulus response estimation in experimental biology. By combining the double sampling scheme proposed by Neyman (1938)

---

[2] See comments by Arrow, Rosen, and others in Cummings *et al.*, (1986), as well as Mitchell and Carson (forthcoming).

with a very weak postulate from reveal preference theory, we show how the relevant demand parameters can be estimated in a highly efficient manner. Finally in an empirical application, we take up Freeman's challenge and show that by using this method we are able to predict the actual outcome of the 1984 California Ballot Proposition 25 (Clean Water Bond Act).

I. The Political Market Versus the Private Goods Market Model

It has usually been argued that the market which a CV survey should attempt to emulate is the private goods consumer market.[3] In its pure form, such a model embodies the notion of a consumer with well developed tastes whose purchase decisions are based on a full understanding of the available alternatives based on prolonged experience in the market. An alternative framework, which has received recent support from some CV researchers, is provided by the political market model.[4] There is a considerable body of theory developed by economists and political scientists on the properties of these markets.[5] The form most relevant to the CV method is the referendum, where the voter is faced with a one-time (or at best a very infrequent) choice on a predetermined policy package to which he must vote yes or no.[6] Here the behavior the researcher desires to be able to predict using a CV study is how voters would actually vote if a proposition to provide the specified amenity was actually placed on the ballot. The voting decision suggest a more complex, and some (e.g. Morgan, 1978) would say a more realistic, model of decision making than the one implied by the idealized private goods market model.[7] This model assumes that people make choices which are influenced by multiple motives, by contextual factors, and by less than perfect information.

---

[3] See in particular the reference operating conditions proposed by Cummings *et al.* (1986).

[4] See Cummings, Cox, and Freeman (1984), Randall *et al.* (1985), and Cummings *et al.* (1986). An early advocate of the referendum model for CV surveys was Ridker (1967).

[5] See for example Bowen (1943), Downs (1957), Black (1958), Buchanan and Tullock (1962), Mueller (1979), and Enelow and Hinich (1984).

[6] In virtually all states voters may vote on binding propositions placed on the ballot by the state legislature (and in some instances by citizen petition). Some states provide for advisory referendums as well. See Magleby (1984) for an overview of the actual referendum process and research on it.

[7] Marketing researchers (Bettman, 1979) have long recognized that many purchases made in the private market place are infrequently made and that the information people gather prior to purchase decisions differs greatly depending on the purchase situation, the type of good, and their past experience.

Several aspects of a referendum make it appropriate as a model for CV surveys. First, referenda are actually used as a mechanism to enable citizens to make binding decisions about the provision of public goods such as a new school building or an increase in the size of the local police force to be finance by a bond issue or higher taxes. Second, the voter's decision in a referendum has clear economic implications since the voter will have to bear his share of any cost implied by the proposition if it passes. Third, it is a plausible choice framework for respondents. They are likely to be familiar with its method of operation and its role in the political system. Finally, the referendum model lends itself well to the survey setting. A ballot, after all, is similar to a series of dichotomous choice questions, and political polling to predict election outcomes is a well recognized feature of public life.

II. Estimating Demand Curves from Political Market Data

There are several ways to differentiate the various methods for measuring the demand for specific public goods which use political market data. Among them are: (1) those based on the actions of voters versus those which are based on the actions of their elected or appointed representatives; (2) those based on the median voter model versus those based on more general (political) equilibrium concepts; (3) those which use micro as opposed to those using aggregate data; (4) those based on observed behavior as opposed to those based on stated preferences; and (5) those which explicitly assume a particular utility function and those which do not. These should be kept in mind when considering the discussion below.

We begin by considering the long line of work, typified by Borcherding and Deacon (1972) and Bergstrom and Goodman (1973), which attempts to ascertain public demand by looking at the actions of elected officials, specifically expenditures on different public goods. These studies typically assume that the actions of elected officials represent the desire of the median voter in their respective political districts.[8] A demand curve can be traced out by comparing the provision of public goods in different districts with the characteristics of the median voter in those

[8] The rationale behind such models is often the Tiebout (1956) "voting-with-your-feet" proposition that consumers tend to move to political jurisdictions which provide the desired mix of public goods.

districts. This framework has a number of problems. Two of the most severe are the implicit assumptions: (1) that the elected officials know and carry out the wishes of the median voter; and (2) that there is no vote trading or logrolling between different issues.

These two problems can be overcome by using the results of votes on actual referenda. Here, the voter can usually be assumed to be voting on a single issue and to be voting his preferences on that issue. This approach has been taken by Birdsall (1965) and Deacon and Shapiro (1975), among others. Most authors following this approach have used fairly aggregate data (e.g., vote returns for cities or counties) and have taken the percentage yes vote to be associated with the median (or mean) voter in the political jurisdiction.[9] A demand curve similar to that in the earlier case can be traced out by regressing the percent in favor in each jurisdiction against the characteristics of the median voter in that jurisdiction.

While this approach has several attractive features, the Hotelling-Bowen-Black-Downs median voter model has come under increasing criticism (Romer and Rosenthall, 1979) for being too simplistic and often indeterminant.[10] Many more sophisticated models of voting equilibrium have recently been put forth in the economics and political science literature (Enelow and Hinich, 1984). Another fundamental problem is that the researcher is limited to examining those public goods for which actual referenda have been held and to the quantities of those public goods embodied in those referenda. This type of analysis is most likely to be fruitful when there are a large number of different referenda for different quantities of the same public goods held in a large number of political jurisdictions. Consequently, it is not surprising that referenda on local expenditures have been those most examined by economists.[11]

The median voter model (or some variant of it) is necessary as long as aggregate data is because there must be some way to tie the observed voting outcomes to observable demographic

[9] The use of aggregate voting data in conjunction with the median voter assumption becomes less troublesome as the units of observation become smaller and more homogeneous, as is likely to be the case when precinct or town meeting data is used (Noam, 1982).

[10] In most of the empirical formulations of the median voter model, it is not possible to distinguish the median voter from any number of other approval quantiles as being the decisive one. Also if the distribution of the demographic variables is assumed to be symmetric, it will not generally be possible to distinguish the mean from the median voter. Further, there is no way to justify the median voter as having the median of all possibly relevant characteristics. Groves and Todo-Rovira (1986) have recently proposed an alternative empirical formulation incorporating distribution parameters which avoids some of these problems.

[11] A table summarizing the results of different studies is given in Bergstrom, Rubinfeld, and Shapiro (1982).

characteristics. In order to overcome this limitation, and for other reasons, such as the stronger foundation of economic theory when dealing with individual preferences, economists have been trying to obtain more micro data on which to perform their analyses. Since individual voters in referenda are not identifiable, researchers have turned to surveys.[12] The survey approach carries with it a change from revealing preferences through binding actions to revealing preferences in response to a survey questions, and the accompanying threat of strategic behavior against which Samuelson (1954) warned.[13]

The use of surveys by economists to determine the demand for public good has taken two distinct directions. The first is the public finance tradition where in the provision of local education is a major area of interest and voting behavior the model of choice for measuring public approval. The second is the environmental economics tradition where the valuation of the benefits of non-marketed goods such as outdoor recreational amenities is a major area of concern and the construction of an artificial private goods market in which to sell the amenity is the model of choice for measuring benefits.

The public finance approach leads naturally to discrete choice questions in a quasi-referendum framework. The questions either used a simple yes/no format or allowed some type of indifference in the form of yes-indifferent-no or more-the-same-less of the good in question.[14]

---

[12] In some cases, the public good in question is less than perfectly public so that its consumption leaves some visible signs through the purchase of some privately marketed good. As examples, consider the recreator at a public fishing lake who must pay for transportation to the lake or a house with good visibility due to air quality which sells for more than a similar house with bad visibility. Techniques for measuring the demand for the public good in these circumstances (hedonic pricing and the travel cost/household production method) are discussed by Freeman (1979).

[13] Economists who have used survey data on stated preferences have almost always been apologetic about using them for this reason. Theoretical developments beginning with a line of arguments put forth by Groves (1973) have shown that strategic behavior is not inevitable. While the mechanisms proposed in the incentive compatibility literature are difficult if not impossible to implement in a standard survey context, this literature does suggest a number of easily adapted features which tend to reduce the incentives for strategic behavior. The particular response format we will use, a discrete choice take-it-or-leave-it, has been shown to be incentive compatible (Zeckhauser, 1973; Hoehn and Randall, forthcoming) as long as the tax price and the quantity offered of the public good are taken as exogenous by the respondent. A respondent can do no better than to say yes if his willingness to pay is greater than the tax price asked, and to say no if it is less than the tax price asked, as long as the government forces all agents with the same identifiable characteristics (e.g., income) to pay the same tax price if the referenda passes irrespective of their earlier response/vote. The plurality decision criteria implied here does not, however, share the near Pareto optimality property possessed by the class of Groves/Clarke mechanisms. Recent theoretical arguments and experiment work related to strategic behavior are reviewed by Mitchell and Carson (forthcoming). It is interesting to note that political scientists, psychologists, and sociologists have taken the opposite view -- that people will generally try to tell the truth in response to survey questions (although they can fail to understand the questions asked of them in almost an infinite number of ways) -- and have used survey data on preferences extensively in their work.

[14] See Peterson (1975), Rubinfeld (1977), Gibson (1980), Gramlich and Rubinfeld (1982), Bergstrom, Rubinfeld, and Shapiro (1982), and Lankford (1985).

Often these questions were posed in the explicit context of voting on a referendum. In the analysis, an explicit utility function is postulated and used with one of the quantal choice statistical techniques popularized in economics by Daniel McFadden (1973).[15] As is typical with the estimation of statistically derived demand curves, specification issues abound. Perhaps, the most troublesome of these is deciding what tax price is perceived by consumers for the public good.[16] The emphasis in the public finance literature has been on estimating income and tax price elasticities.

Researchers undertaking a benefit analysis for the provision of environmental quality and recreational amenities did not originally think in terms of referenda. Rather they thought in terms of valuation questions such as: how far people would travel?, what maximum permit or entrance fee they would be willing to pay?, or how much extra a house with a better view was worth? This naturally led to a direct elicitation using CV of how much greater the costs of an environmental amenity would have to be before the survey respondent would prefer to stop enjoying the benefits of it.[17] The emphasis in the environmental economics literature has been on estimating Hicksian or Marshallian consumer surplus measures rather than income and price elasticities.

II. Simulating Referenda[18]

Economists are interested in the demand curve for a public good as a function of a price, a relationship which is often summarized in terms of elasticity (or consumer surplus) estimates. One natural elasticity measure in the case of a referenda describes how the percent of voters wil-

[15] See Amemiya (1981) for a recent survey of the applicable statistical techniques, and Hanemann (1984) for a discussion of the economic theory underlying these models.

[16] Researchers using this approach have tended to assume that voters either knew the tax price of the referendum or that knowledge of what the voter believed his current taxes to be was an appropriate surrogate.

[17] See for example Knetsch and Davis (1966), Randall, Ives, and Eastman (1974), Greenley, Walsh, and Young (1981), Brookshire *et al.*, (1982), and Desvousges, Smith, and McGivney (1983).

[18] We note here the work which has been the most influential in developing this synthesis. From the public finance literature, Bergstrom, Rubinfeld, and Shapiro (1982) for their use of survey data as a surrogate for observing individual votes in a referendum, for their use of more than one simple yes/no question, and for their ability to estimate fairly precisely an own price elasticity. From the contingent valuation literature, the work of Cummings (Cummings, Cox, and Freeman, 1984; Cummings *et al.*, 1986) who advocated moving toward some type of voting mechanism to make the willingness to pay act more concrete. From the biological statistics literature the work of Cox (1970) and Finney (1971; 1978) on bioassays. We also note a particular intellectual debt to Bishop and Heberlein (1979; 1980) who proposed the rudiments of what we are doing here, in a different context and for a different purpose.

ling to vote for a given proposition changes with a percent change in tax price. Letting V_p stand for the percent of the voters who are willing to vote for the referendum at tax price, T_p, we can define the tax price elasticity of voter approval as,

$$\varepsilon_v = \frac{\partial \ln(V_p)}{\partial \ln(T_p)}, \tag{1}$$

where ln represents the natural logarithm.[19] This definition is somewhat unusual, because it is defined over the population of voters rather than a single consumer.[20]

In what follows, we propose a simple and effective method of directly estimating the demand for a public good which combines the public finance tradition of examining referendum results, the contingent valuation tradition of asking people directly about their willingness to pay, and the experimental biological tradition of estimating dose response relationships.

The Response to a Stimulus: Referenda Voting and Bioassays

Even though the term elasticity is infrequently used in the bioassay literature, the researcher conducting a bioassay is attempting to measure something directly analogous to a demand curve, but posed instead in terms of the voting population and a tax price.[21] In a classical bioassay, a large number of test specimens are randomly assigned to a small number of groups. Each of these groups is then exposed to different levels of the same stimulus (usually a poison or hazardous condition). The number of animals in each group responding to the stimulus (usually by dying) is counted, and the percentage responding calculated. These percentages are then plotted against the level of the stimulus and a curve statistically fitted to these points. The location estimate of primary interest is usually called the LD50 for the lethal dose where it is estimated exactly 50% of the specimens would die.[22]

---

[19] We could have replaced T_p with a continuous function that describes different tax prices for different members of society rather than the flat tax price which is assumed for convenience here.

[20] Almost all of the definitions in the literature explicitly incorporate voting are defined for an individual representative voter or the median voter. See Peterson (1975), for example.

[21] Finney (1978 and earlier editions) is a fairly complete source for information on bioassays. The focus of Cox (1970) may be more familiar to economists working in discrete choice situations. See also Finney (1971) and Mead and Curnow (1983).

[22] The term ED50, for median effective dose, is now frequently used to denote the location where 50% of the specimens would respond because experiments in which death is not the expected response are now common.

To draw out the analogy, the tax price is the stimulus and the percent willing to vote in favor is the response. The LD50 point is the minimum response necessary for the referendum to pass, that is, it is the highest tax price at which the median voter would vote to pass the referendum. In order to estimate the demand function, all that would theoretically have to be done is expose some different groups of voters to randomly assigned different tax prices, observe the percent of each group favorable toward voting yes on the referendum at their assigned tax price, and record these percentages and their accompanying tax prices.

In actuality, it is not quite this easy and there are some differences between conducting a bioassay and simulating referenda at different tax prices. One must go out and somehow survey representative samples of the voters at different tax prices in order to obtain the desired estimate, since they are not available in our "lab" for random assignment. The differences between a controlled experiment and random sampling are, however, not as great as one might expect.[23] We will show below how it is possible to combine the two. Another difference, which is more challenging, is that the random elements in a sample survey are considered to occur through the selection of particular individuals rather than through random responses to the stimulus by those selected individuals. This difference is important and we will return to it after discussing the basic sampling theory concepts.

Sampling: Simple, Stratified, and Double

Because the human population is less homogeneous in its response patterns than are fruit flies and pure strains of white lab rats, large gains in efficiency may be obtained by changing from simple random sampling to stratified random sampling.[24] Consider a variable Y which takes on two values, say the presence (Y=1) or absence (Y=0) of a particular characteristic. We are interested in the estimated percentage, P, of the population with this characteristic.[25]

---

[23] Fienberg and Tanur (1985) develop the parallels between experimental design and sampling techniques.

[24] Under simple random sampling each individual in a population has an equal chance of being chosen to be interviewed for the survey. Cochran (1977) is a good source for information on different sampling schemes.

[25] We will follow the convention that population values will be denoted by capital letters while sample values and statistics will be denoted by lower case letters.

If we take a simple random sample of size n, P is estimated by

$$\bar{p} = n^{-1} \sum_{i=1}^{n} y_i \tag{2}$$

If an indicator variable I=1,2 is available and $VAR(Y|I=1) \neq VAR(Y|I=2)$, it is possible to define two potentially useful strata based on this indicator variable. Given a maximum available sample size of n and equal costs for sampling units when I=1 and I=2, the optimal (Neyman) allocation of n is proportional to the number of units in each of the i_{th} strata (N_1, N_2) and the variance of each strata, so that

$$n_i = \frac{N_i \sqrt{p_i q_i}}{N_1 \sqrt{p_1 q_1} + N_2 \sqrt{p_2 q_2}}, \tag{3}$$

where $q_i = 1 - p_i$. The stratified sample estimate of P is

$$\bar{p}_{st} = \frac{N_1 \bar{p}_1}{N} + \frac{N_2 \bar{p}_2}{N}, \tag{4}$$

where $\bar{p}_1$ and $\bar{p}_2$ are calculated within strata in the same manner as the estimator for simple random sampling in EQ. (1). The gain in efficiency of $\bar{p}_{st}$ over $\bar{p}$ is greatest if all the units for which Y=0 are in one strata and all the units for which Y=1 are in the other strata. One attempts to choose a variable for stratification that best accomplishes this.[26]

The variable on which the survey stratification is carried out must be visible before the units are chosen. If no suitable variable for stratification is available before undertaking the survey, the double sampling technique proposed by Neyman (1938) offers a useful alternative.[27] Double sampling is based on an extremely simple idea. If a good indicator variable is not available, conduct a large survey using a simple random sample, obtain the indicator variable desired, and then reinterview some percentage of the original sample after stratifying on the indicator variable. In many cases, double sampling will be more efficient than a single large simple random sample survey costing the same amount. Double sampling is likely to be a profitable stra-

[26] We assume here that N, N_1, and N_2 are known.

[27] Cochran (1977) provides a discussion of double sampling and relates it to other sampling techniques. Double sampling is used frequently in medical surveys when it is inexpensive to identify a group potentially having desired characteristics, but expensive to test for them. See Deming (1977) for an example.

tegy when the gains from stratification using the indicator variable obtained in the first stage sampling are very large. The percentage of respondents to be reinterviewed in the second phase is a function of the variances in the new strata and the cost of a second stage interview relative to first stage interviews. Double sampling also allows one to obtain unbiased estimates of population strata sizes N_1 and N_2 when these are unknown, as is often the case.

Efficient Referenda Simulation

In this section, we propose a more efficient method of simulating referenda. First, one draws a simple random sample of voters and asks each of them whether or not they would vote in favor of a specified referendum, if its tax price to the voter was x dollars. Having obtained a yes or no response to the x dollar question, we can use this variable as an indicator for stratification. If an individual said yes to x dollars, then the probability that the same individual in the same interview would say yes to any other amount *less than* x dollars can safely be assumed to be one as long as the good is normal. Similarly, the probability that an individual who said no to x dollars would also say no to any amount *greater than* x dollars is also one. These two statements utilize the weakest postulate of revealed preference theory -- that if a good is not preferred at a specified price, it will also not be preferred at any higher price, and that if it is preferred at the specified price, it will also be preferred at any lower price.[28]

It should be clear that the variance of the responses of those who answered yes to the x dollar question will be zero if the subsequent tax price asked is less than x since these respondents would all answer yes.[29] From EQ. (3), if there are a relatively small number of respondents who can be reinterviewed, the optimal allocation of these reinterviews between the two strata (using as the strata indicator variable the respondent's answer to the x dollar tax price) would be no reinterviews in the strata in which the responses would have zero variance and all of the avail-

[28] We are also using the assumption that if asked the same question twice in the same survey (without any intervening information) a respondent will give the same answer. This assumption is different from that made by biologists of lab animals which are assumed to respond to some specified level of a stimulus only in a probabilistic fashion. In particular, note Finney's (1978) criticism of a method proposed by Dragstedt and Lang (1928) which implicitly assumes that if an animal did not respond to a specified dose, it will not respond to any dose lower than that one.

[29] This statement holds for the variance of the respondents who said no to x dollars of the subsequent amount is more than x dollars.

able reinterviews in the other strata. The allocation depends on whether the new tax price was greater than or less than x dollars. This scheme can obviously be repeated based on the responses to the initial tax price asked, x_1 dollars, and the next tax price asked, x_2 dollars.

To estimate a demand curve, we will need to know the percent in favor at several different tax prices. The questions are then two. How many tax prices do we need? And, how many different tax prices can we ask a respondent? There is an obvious tradeoff between how well the percentage acceptance at a particular tax price is estimated and the number of different tax prices for which estimates can be obtained given a fixed sample size.[30] The larger the number of tax prices used, the smaller the sample size at each tax price, and, thus, less precision in the estimation of $\bar{p}$ at that point. In most cases, the overall sample size will be limited by the researcher's budget. Also, it should be realized that there are severe response problems with asking people a series of very repetitive questions (Mitchell and Carson, forthcoming). In particular yea-saying may become a problem, and the tax prices asked can no longer be consider to be exogenous so that the desirable incentive compatibility properties of this particular elicitation format are lost. Thus, we cannot recommend asking a single respondent more than one or two follow up tax prices.

For many purposes, it is possible to rule out some parts of the demand curve as being essentially uninteresting, and, thus, it may be possible to restrict the prices asked about to a fairly small range. For instance, a city council may have no interest in proposing a referenda with a tax price so high that it stand no chance of being approved by 50% (or more) of the voters, or a policy analyst may be most interested in an elasticity estimate in the area of some particular tax price. This may not always be the case. In particular, if the researcher is interested in mean consumer surplus, estimates of the extreme quantiles become very important. In such cases, special

[30] In many cases, the survey firm will place constraints on the number of tax prices and the number of "follow up questions/reinterviews". This is particularly true of omnibus surveys.

[30] There are a some similarities between the method we are proposing and the bidding game frequently used in CV surveys (e.g., Randall, Ives, and Eastman, 1974) which has been shown to usually suffer from starting point bias (Mitchell and Carson, forthcoming). In the bidding game, however respondents are bid up (or down) in small increments from an initial starting bid request until maximum willingness to pay is obtained. In the scheme we are proposing, one randomly chosen additional tax price (which may be a large jump from the first request) is asked. Maximum willingness to pay is never obtained, only discrete indicators of it.

experimental designs, larger samples, and/or nonstandard estimation procedures will usually be necessary for precise estimates.[31]

IV. An Empirical Application

California's frequent statewide referenda offer a useful context for a natural experiment whereby we first use our method of simulating referenda in a survey and then compare our hypothetical results with the actual outcome of a referendum vote (at a single tax price) for the same public good.[32]

We have chosen to model our referenda simulation after Proposition 25 ("Clean Water Bond Law of 1984") due to our familiarity with willingness to pay for water quality in a contingent valuation context (Carson and Mitchell, 1986). Proposition 25 authorized a 20 year bond issue of 325 million dollars largely for the purpose of constructing sewage treatment plants.[33] In many cases, the money raised from the sale of the bonds would be used as the state's matching contribution for Federal sewerage treatment grants.[34] The cost of Proposition 25 to the average California household is approximately $4 per year for 20 years, after interest payments are include. The principal and interest on the bonds would be paid out of the state's general fund. Because the exact distribution of repayment obligation is unspecified from the taxpayers point of view, a flat per household tax is assumed for ease of administration in our simulation. This is approximately correct if sewer charges/utility taxes are used to pay off the bonds, somewhat less correct if sales tax revenue is used, and too regressive if the somewhat progressive state income tax is used. A detailed description of Proposition 25 was presented in the California voters pamphlet (Eu and Hamm, 1984) which was mailed to all California voters before the election.[35]

---

[31] In contrast median willingness-to-pay can be estimated with few problems. Such a measure is suggested by the use of a political decision rule such as majority rule rather than the potential Pareto improvement criteria often used in a benefit-cost analysis.

[32] While California is perhaps unique in the number of issues on which state-wide referenda votes are held and the regularity at which these votes occur, referenda, particularly on local issues, are commonly held throughout the country. Magleby (1984) is a comprehensive source on referenda in the United States. The study of California referenda has long been popular with both economists and political scientists. See Wolfinger and Greenstein (1968), Mueller (1969), Deacon and Shapiro (1975), Lutrin and Settle (1975), and Lake (1983).

[33] The referendum was put on the ballot by the state legislature which approved it with only one dissenting vote.

[34] This is the fourth bond issue of this nature put to a referendum vote in California since 1970. The other referenda in 1970, 1974, and 1978 all passed comfortably and authorized 925 million dollars in bonds. Radosevich (1975) provides an analysis of the 1974 referendum from a political science prospective.

[35] It was received by the voters after our referenda simulation took place, although our description of the simulated referenda in the survey questions was based on information given to us by the Secretary of State's office that prepared the voter's pamphlet. There were 16 other propositions being voted on. Proposition 25 received almost no publicity, being overshadowed by the Presidential race and some of the other

Implementation

We arranged to simulate Proposition 25 at different tax prices on a Field Institute California Poll held in early October 1984. The Field Institute is a not-for-profit organization, and the well known California Polls are conducted for a group of leading California newspapers and broadcasters, as well as a consortium of California universities. The California Polls are devoted to examining issues of national and state politics. In this particular California Poll, the focus was on the November 1984 general election, providing an excellent and realistic context for our simulation.

The topics we asked about were:

(1) Was the respondent aware of Proposition 25? [AWARE25]

(2) How did the respondent intend to vote on Proposition 25 (a water quality bond issue), given the brief description of the referendum which would appear on the ballot in the November election? [VOTEREF][36]

(3) If the referendum were to cost their household $4 a year for 20 years (the amount implied in the legislative analysis contained in the larger election information pamphlet mailed out by the California Secretary of State's office in mid-October), how would the respondent vote? [VOTE4]

(4) [Depending on the response to the question above] How did the respondent intend to vote on one of three randomly assigned lower amounts (if the vote at $4 was no) or one of six randomly assigned higher amounts (if the vote at $4 was yes)? [VOTE1, VOTE2, VOTE3, VOTE5, VOTE7, VOTE10, VOTE15, VOTE25, VOTE50]. The assignment scheme for dollar amounts is discussed in more detail below.

propositions. It received cursory endorsement from most of the media.

[36] We were able to repeat the first two questions one month later in the California Poll taken immediately before the November General Election. There was approximately a 15% increase in awareness of Proposition 25 (from 30%) and almost no change in VOTEREF.

Field California Poll

The resulting Field California Poll consisted of 1022 respondents sampled by random digit dialing and interviewed by telephone. Assuming that every California household containing one or more registered votes has a telephone (listed or unlisted) and that there were no systematic non-response effects, this sample can be treated as one obtained through simple random sampling.[37] The VOTEREF and VOTE4 questions are from this sample, as are the attitudinal and demographic variables used in examining question (d).

The sampling design for the VOTE1, VOTE2, VOTE3, VOTE5, VOTE7, VOTE10, VOTE15, VOTE25, and VOTE50 questions was based on the six independent replicants making up the Field California Poll and the response to the VOTE4 question. To estimate the sensitivity of the percentage of California voters in favor of Proposition 25 to the tax price, the following experiment within a survey (Fienberg and Tanur, 1985) was conducted. If a respondent answered "no" to VOTE4, then that respondent was asked one of the following: VOTE1 [would you vote for the referendum if the cost were $1 per year], VOTE2 [cost $2 a year], or VOTE3 [cost $3 a year].[38] The particular treatment depended on which of the six replicants the respondent was in, with two replicants each assigned to VOTE1, VOTE2, and VOTE3.[39] From these responses, the percent who would vote in favor of the proposition at each of the dollar amounts for which questions can be easily calculated.

The issues we wish to address here are:

(a) Is the response to VOTEREF independent of VOTE4?

(b) Does providing the respondent with information on the cost of the referendum result in reducing the number of "don't know" responses?

---

[37] There was a simple stratification between Northern and Southern California and some very minor clustering within telephone area codes. These effects are likely to be small due to the broadness of the strata and low inter-cluster correlations and thus are ignored. A detailed description of the sampling plan for this survey is available from the Field Institute.

[38] There are really three strata. The third being those who stated "don't know" to VOTE4. These respondents are for now considered to be against the referendum at any price. (The original design called for the "don't knows" to be treated as "no," but this was not done in the survey execution.)

[39] "Don't know" responses to VOTE4 were not asked any additional willingness to vote questions.

(c) How sensitive is the percent of California citizens who are willing to vote for Proposition 25 to the cost of that referendum (i.e., what is the price elasticity of voter approval at different points)?

(d) What demographic and attitudinal characteristics are associated with the response to VOTE4 (or VOTEREF)?

Independence of VOTEREF and VOTE4

To a political scientist or a sociologist, there would be no reason to ask whether a respondent's answers to VOTEREF and VOTE4 are independent. They would assume not. In an economist's world, though, questions about a person's wants are clearly suspect, unless they are tied to the cost of fulfilling those wants. We know from past experience that questions like VOTEREF are good predictors of actual outcomes on referenda [Magleby (1984)] and this is true for Proposition 25.

Proposition 25 passed in the November election receiving 73% of the vote. For our survey, dropping all the "undecideds" to the VOTEREF question and recalculating the percentages gives 83% for and 17% against. However, studies of voting behavior in bond issues indicate that many of our undecided voters are going to vote and that more than half will vote "no" (thus preserving the status quo). Splitting the undecideds 50/50 results in 75% in favor; 60/40 against results in 73% in favor; while splitting them 70/30 results in 70% in favor. The 95% confidence interval based on sampling variation for percentages in this region (for our size sample) is approximately plus or minus 3%. As is usual, non-sampling errors, particularly in the form of uncertainty over how the don't know's are going to vote and *who* is going to actually vote, are likely to be the major source of error in predicting actual voting outcomes.

VOTEREF and VOTE4 each have three possible responses, "yes," "no," and "don't know". A 3 x 3 contingency table can be formed by cross classifying the responses to these two questions. This table with the actual frequencies, as well as row, column, and cell percentages and marginals is given in Table I.

Denoting the probability that an individual falls into the ith row and the jth column of table I by P_{ij}, we can state our hypothesis formally:

$$H_1 : P_{ij} = P_{i.}P_{.j} \text{ for } all\ i \text{ and } j$$
$$K_1 : P_{ij} \neq P_{i.}P_{.j} \text{ for } all\ i \text{ and } j\ .$$

For large n, a test statistic based on the difference between observed and expected values (under H_1) can be defined, which has a $X^2_{(df=[i-1][j-1])}$ distribution under the null hypothesis (Fienberg, 1977).[40]

TABLE I ABOUT HERE

Difference Between Percent of "Don't Know" Responses

Providing a respondent with information on how much the referendum will cost him or her if it passes should reduce uncertainty about the referendum, and, as such, should reduce the number of "don't know" responses.[41] This hypothesis can be stated as:

$$H_2: D_1 = D_2$$
$$K_2: D_1 > D_2,$$

where D_1 is the number of "don't knows" responses to VOTEREF and D_2 is the the number of "don't know" responses to VOTE4. D_1 and D_2 can be thought of as resulting from different binomial distributions, and the test of H_2 versus K_2 written in terms of the binomial parameter θ where the estimate of θ is given by D/n. For VOTEREF, θ_1 = .2456 (standard deviation = .4306), and for VOTE4, θ_2 = .0832 (standard deviation = .2762).

[40] This test statistic can be defined as:

$$\sum_i \sum_j \frac{[observed - expected]^2}{expected}.$$

[41] "Don't knows" on a question like this, where the outcome (cleaner water) is perceived as socially desirable, are generally of two types: (1) those who are unsure of the cost of passing the referendum; and (2) those who are stating "don't know instead of stating a "no." Providing respondents with the cost information should remove much (although not all) of the uncertainty associated with the referendum's impact. Most of the remaining responses will be of the "don't know-means-no" type. This categorization can also be justified using economic theory in which consumers do not make purchases unless the utility gained from making the purchase is sufficiently above the utility lost at the margin from spending the money.

Since the binomial distribution has the regenerative property, we can perform an exact test of H_2 versus K_2. The probability of the observed outcome under H_2 is less than .0001, so we reject H_2.[42]

Estimation of the Price Elasticity of Approval

The statistician analyzing a bioassay generally uses some form of quantal response model. Usually a logit, ($ln\,[1/(1-p\,]$), or a probit, the inverse of the standard normal cumulative distribution, response function is assumed.[43] In both cases, the maximum likelihood estimates can be obtained using iteratively reweighted least squares, and both are members of the family of generalized linear models (McCullagh and Nelder, 1983). We will report only the logit results here. Of more importance is the functional form for tax price. Here both biologists and economists make two typical choices: linear or logarithmic. We will estimate both before assessing the need for other functional forms. At each of the j tax prices (j=1, ..., 10), y_j of the n_j respondents interviewed indicated a willingness to pay to pay the specified tax price. (Different ways to define the y_j and n_j are considered below).

The log-likelihood function for the logistic response model can be written as,

$$L(p,y) = \sum_{j=1}^{10} [y_j ln\,(p_j/(1-p_j)) + n_j ln\,(1-p_j)]\,, \tag{5}$$

where the y_j are considered to be distributed binomially with index n_j and parameter p_j. We assume that

$$g(p) = ln\,(p_j/(1-p_j)) = \eta_j = \sum_{k=1}^{2} x_{jk}\beta_k\,, \tag{6}$$

where x_{j1} is a constant term and x_{j2} is either the j_{th} tax price or the natural log of the j_{th} tax price. We can now rewrite the log-likelihood function in terms of β,.

$$L(\beta,y) = \sum_{j=1}^{10} \sum_{k=1}^{2} y_j x_{jk} \beta_{jk} - \sum_{j=1}^{10} n_j ln\,(1 + \exp\,(\sum_{k=1}^{2} x_{jk}\beta_k))\,. \tag{7}$$

---

[42] As n is large, the normal approximation can also be used. This gives a *Z–value* of 10.15.

[43] None of our observations are in the extreme tails, so it will be impossible to distinguish between the one of two response functions given our sample size. There are generalizations of the logit/probit curves for dose response functions which involve estimating additional parameters, if these do not fit the data (Prentice, 1976).

The logistic response curve may be fit using many standard statistical packages. We used GLIM [Baker and Nelder (1978); McCullagh and Nelder (1983)], because it has convenient facilities for handling the different binomial denominators, n_j. β is estimated by calculating,

$$\hat{\beta} = (X'WX)^{-1}X'Wz \ , \tag{8}$$

where

$$z_j = \hat{\eta}_j + \frac{y - n\hat{p}_j}{[n_j\hat{p}_j(1-\hat{p}_j)]} \ , \tag{9}$$

and W^{-1} is a diagonal matrix with elements $(\hat{p}_j(1-\hat{p}_j)/n_j)$. The covariance matrix of $\hat{\beta}$ is given by $(X'WX)^{-1}$, since the scale σ^2, of the binomial distribution is assumed to be one. The statistic, $\sqrt{n}(\hat{\beta} - \beta)$, is distributed $N(0, \dot{n}(X'WX)^{-1})$, plus a bias term of order $O_k(\dot{n}^{-1/2})$, which goes to zero at the rate of one over the square root of $\dot{n}$, where $\dot{n}$ is the minimum over j of the n_j.

The number of observations, n_j at each price plays a role in the weight matrix, W, as an indicator of the precision with which p_j is estimated. Otherwise, it only plays a part in estimating the p_j. Three intuitive definitions can be used for n_j. The first, N*, is the number of respondents who were actually asked the $VOTE_j$ question. The second N**, is the number of respondents who either explicitly or implicitly answered the $VOTE_j$ question. This is N* plus, for $VOTE_j$ tax prices greater \$4, all the respondents not willing to pay \$4, and, for $VOTE_j$ tax prices less than \$4, all those respondents who were willing to pay \$4. Our third definition of n_j, N***, comes from the role that n_j, the group sample size under simple random sampling, plays in EQ. (7). The variance of the simple random sample estimator $\bar{p}_j$ is,

$$VAR(\bar{p}_j(Simple)) = [\bar{p}_j(1-\bar{p}_j)]/n_j \ . \tag{10}$$

Our estimate of p_j is obtained using double sampling. Its variance is a variant of the stratified sampling estimator (Cochran, 1977; 334-335),

$$VAR(\hat{p}_{j(Double)}) = \frac{W_1^2\bar{p}_{1j}(1-\bar{p}_{1j})}{n_{1j}-1} + \frac{W_2^2\bar{p}_{2j}(1-\bar{p}_{2j})}{n_{2j}-1} \ , \tag{11}$$

where $\bar{p}_{1j}$ and $\bar{p}_{2j}$ are the simple random sample estimates of the j_{th} p in strata 1 and 2, respectively. The values W_1 and W_2 are the ratios of the number of observations in strata 1 and 2 to the number of observations in the entire sample (these are the same for all j), and n_{1j} and n_{2j} are the number of observations from the two strata sample for the $VOTE_j$ question.[44] We could always take the minimum of n_{1j} and n_{2j} to be two, in which case the variance contribution from one of the two strata will always be zero, because either $\hat{p}_{1j}$ will be zero or $\hat{p}_{2j}$ will be one.[45] However, even the allocation of these two observations is unnecessary as revealed preference theory tells us what the value of either p_{1j} or p_{2j} will be.

It is possible to set the value of $VAR(\hat{p}_{j[Double]})$ equal to the formula (EQ. 10) for the variance of $VAR(\bar{p}_{j[Simple]})$ by substituting $\hat{p}_{j(Double)}$ for $\bar{p}_j$, so that,

$$VAR(\hat{p}_{j(Double)}) = [\hat{p}_{j(Double)}(1-\hat{p}_{j(Double)})]/n_j \ , \tag{12}$$

and then solve for the value of n_j, N***, which satisfies the equation. This is the appropriate value of n_j to use in the estimation of the demand curve.

Table II displays p_j, N*, N**, and N***. Comparison shows the large increase in efficiency of N*** over the number of observations represented by N*. Note how the effect of double sampling is most pronounced for tax prices near $4. For these tax prices, the separation of respondents into two strata of pure types is most successful. It becomes less successful as the strata sampled become more mixed with those who are willing and not willing to pay the specified tax price. The results presented below are for N*** since this is the technically correct definition of effective sample size. The results with N* and N** are given in footnotes and are very similar.

TABLE II ABOUT HERE

As the stimulus variable, ln(price), provides a better fit to the data than the model incorporating a linear price stimulus. For the logit equations, the price model has a X^2 of 69.74

[44] A simple numerical example may be helpful here. The total number of interviews was 1022. Of these, 831 were willing to pay $4 and 191 were not, so W_1 equals 191/1022 and W_2 equals 831/1022. $\bar{p}_{1j}$ equals zero, so the first term of EQ. (11) drops out. We asked 127 [$n_{2(\$15)}$] in strata 2 the VOTE15 question and $\bar{p}_{2(\$15)}$ equals .803, so $VAR(p_{\$15(Double)})$ equals ([831/1022]^2 [(.803)(1-.803)/(127-1)]).

[45] If $VOTE_j$ represents a higher tax price than VOTE4, strata 1 will contribute zero to the variance estimate since p_{1j} is always zero, and if $VOTE_j$ represents a lower tax price than VOTE4, strata 2 will contribute zero to the variance estimate since p_{2j} is always one.

(df=8) while the ln(price) model has a X^2 of 8.38 (df=8).[46] Since the scale of the binomial distribution is one, the expected value of X^2 if the model is appropriate is 8, indicating that the ln(price) model fits well while the linear price model does not.

This can best be seen graphically. Figure 1 is graphed in the manner of a bioassay to emphasize the nature of tax price price as the stimulus variable under the researcher's control. The ln(price) model fits the data everywhere while the linear price model cannot really be said to fit well anywhere except in a very small region.[47] The actual points are marked with small x's.

FIGURE 1 ABOUT HERE

The estimated equation is

$$E[\ln((p_j)/(1-p_j))] = 2.256 - .6010*\ln(price), \qquad (13)$$

where the asymptotic standard error of the intercept term is .0758 (t=29.76) and the asymptotic standard error of the ln(price) term is .03757 (t=15.99).[48] The tax price elasticity of voter approval ε_v from EQ. (1) can be defined here as:

$$\frac{\partial \ln V_p}{\partial \ln T_p} = \frac{\partial \ln([1+exp(-X\beta)]^{-1})}{\partial x_k} = \frac{\exp(-X\beta)}{(1+\exp(-X\beta))} * \beta_k, \qquad (14)$$

where x_k is ln(price). Using the coefficients from EQ. (13) and evaluating EQ. (14) at a tax price of \$4, gives a point estimate of ε_v as -.1167. The approximate 95% confidence interval for ε_v is [-.0848, -.1508].[49]

---

[46] The X^2 statistics are 2.36 and 16.47 for N*, and 7.31 and 89.76 for the N** models, where the larger number in both cases is for the linear price model.

[47] One of the advantages of this approach with its small number of data points (though not underlying observations) and only one stimulus variable is that bad fits are obvious and desirable transformations more easily seen.

[48] An estimation of the regression diagnostics suggested by Pregibon (1981) for logistic regression reveals no obvious problems with EQ. (13). For the N* ln(price) model, the coefficients are 2.264 (.1337) and -.5968 (.0606), and for N** are 2.236 (.0625) and -.5888 (.0303), where the numbers in parentheses are the accompanying asymptotic standard errors.

[49] The confidence interval for the elasticity is calculated using a very conservative method (i.e., larger than the standard asymptotic expansion) due to Hauck (1983) which is valid for any dose level. It is likely to be significantly smaller for the \$4 tax price because of the large number of actual observations observed at that point. For the N* ln(price) model, the upper 95% bound, the point estimate, and lower 95% bound for ε_v are [-.0848, -.1146, -.1508] while for N** they are [-.0866, -.1145, -.1424]. While the use of double sampling is quite important, the actual choice of the definition of the n_j is not tremendously important as ε_v is estimated far more precisely under any of these three

Figure 2 displays the actual demand curve for the public good represented by Proposition 25 as a function of price in the traditional economic manner with price on the vertical axis and quantity on the horizontal axis.[50] Median willingness to pay is approximately $50, a number substantially, but not grossly higher than the $25 found by Carson and Mitchell (1986) for a fishable to swimmable quality water change.[51] We believe that if the average tax price had been $50 there would have been substantial *public* controversy over whether California was getting a good deal for its money, and that this controversy would have caused a significant reduction in the percent of yes votes at the higher tax prices.

FIGURE 2 ABOUT HERE

While the demand curve we have estimated is a demand curve from the politician's perspective, it looks as if it could have been lifted from one of the standard elementary "principles" texts. Indeed we could have estimated the same demand curve using the individual observations and declared the same (approximate) coefficients to those of a representative voter. It is important to note how such a curve is estimated. Following standard principles books, the demand for a good is a function of price, income, and taste. At each of the different tax prices, we have a random sample of voters from the same population. Thus, income and taste variables are held constant. The only thing which has changed is the tax price. Because of the controlled application of tax price as a stimulus, there are none of the usual separability or simultaneity problems associated with estimating demand functions. Essentially what we are estimating is a response surface under controlled conditions. As such, the extension to the case, where different quantities of the public good in question are used as well as different tax prices, is straightforward from both a conceptual and practical standpoint.[52]

definitions of N than almost any elasticity estimate we can find in the public good's literature.

[50] Equation (13) predicts 80% approval at $4 and predicts 73% approval (the actual outcome of Proposition 25) at $8. Given the higher propensity of higher income people to vote and the proportional/slightly progressive nature of the tax price, $8 may have been closer than $4 to the average tax price paid by those *actually voting*.

[51] One of the major problems with this and many other referenda is that it is not really clear what people were buying with their money.

[52] The actual monetary cost of conducting such a survey, however, might be quite sizeable due to the need for a much larger initial sample

Characteristics Associated with Yes on VOTE4

Past work on the determinants of willingness to pay for water quality (Carson and Mitchell, 1986) suggests the following demographic or attitudinal variables (self-identification as an environmentalist, concern over water pollution, participation in water-based recreation, and income) are associated with higher willingness to pay for water quality improvements. Other researchers have occasionally reported significant age, race or sex correlations. Since there are often partisan differences on referenda, we also selected political ideology and presidential preferences as additional possible correlates.

The hypothesis of independence of VOTE4 with each of these variables can be tested using the X^2 test described earlier. Before doing this, we dichotomize VOTE4 so that it equals one if the respondent was willing to pay $4 and equals zero if the respondent did not indicate a willingness to pay $4.

For the 2 x 2 table of preference for Reagan by VOTE4, we have a X^2 value of .124 (df=1) indicating that the null hypothesis of no association cannot be rejected. The 2 x 2 table with Republican/Non-Republican [PARTY], shows some support for the hypothesis that partisan differences influence VOTE4, as the X^2 value is 4.28 (df=1) which has a p-value of .04. For the water based recreation variable [WATUSE], the hypothesis of independence is rejected [X^2 = 4.91 (df=1; p-value=.03)]. The hypothesis of independence is also rejected with respect to an economic-environmental tradeoff variable, ECEV, [X^2 = 33.8 (df=4; p-value=.001)], with respect to confidence in the state legislature to provide environmental protection, CONFLEG, [X_2 = 28.18 (df=3; p-value=.001)], and with respect to self-identification as an environmentalist, ENVIST, [X^2 = 21.38, (df=1; p-value=.001)].[53] We found no race or sex effects.

Income was measured in this survey using the California Poll's standard series of categories. Unfortunately, over 30% of the respondents were in the highest income category

size.

[53] WATUSE takes two values; 1 when the response was "engage in water-based recreation", and 0 when the response was "do not engage in water-based recreation". ECEV is a five point scale with 1 equal to "very pro-environmental with respect to economic trade-offs" and 5 equal to "very pro-economic growth with respect to environmental trade-offs." CONFLEG is a four point scale with 1 equal to "great confidence in the state legislature in providing environmental protection" and 4 equal to "no confidence in the state legislature."

(total household income greater than or equal to 40 thousand dollars) with most of the detail available for low income households. At the $4 tax price, the income elasticity appears to be close to zero. This is perhaps not too surprising since the cost of the referendum at $4 is not burdensome. As one increases the tax price from $4 to $50, the income elasticity appears to increase but still remains fairly small.[54] This is not surprising since sewer treatment plants are not generally considered luxury goods.

V. Concluding Remarks: Knowledge and the Act of Voting

Perhaps the strongest criticism of willingness to pay estimates obtained using the contingent valuation has been put forth by Bishop, Heberlein, and Kealy (1983) and Freeman (1986). In the words of Freeman, "contingent valuation works best where we need it least." By this, Freeman was referring to the fact that people are most likely to give reasoned and informed willingness-to-pay answers about public goods and levels of public goods with which they have had the most experience. In particular, he refers to recreational demand where methods (e.g., travel cost analysis [Clawson and Knetsch, 1966]) based on complementarity with marketed goods already exist for benefits estimation. Bishop, Heberlein, and Kealy (1983) explore this theme further noting the long process of comparison and consultation a consumer is likely to go through before purchasing a large durable item such as an automobile.

This is undoubtably true of purchases such as automobiles. However, this paper puts forth a different model, one which likens the decision making process to the mental process an individual goes through before casting a vote in an election. It is important to note that Bishop, Heberlein, and Kealy, and Freeman have raised their criticisms in response to the hypothetical nature of contingent valuation surveys and the fact that responses given in a contingent valuation survey are not binding.

---

[54] We hesitate to speculate on its exact functional form due to data problems and estimation problems, and the fact that our experimental design was not intended for this purpose. A typical CV survey uses modified "standard" demographic questions in order to obtain more detail, particularly with regard to income.

There is a vast literature in political science which suggests that people are not tremendously well informed when they cast votes.[55] Only a few days before the election, less than 50% of the likely voters were aware that Proposition 25 was on the ballot. This was in spite of newspaper endorsements and a detailed voter's pamphlet from the Secretary of State's office. Magleby (1984) cites research showing that fewer than 30% of California voters read the voters pamphlet, despite the fact that it is free and is delivered to their home. The descriptions of the propositions are sometimes complicated, often full of legal jargon, and require a reading level above that possessed by many of the voters. Yet, people frequently vote on decisions which have a tremendous influence on their lives.

We assert that the half hour to an hour contingent valuation survey interview with its tremendous detail on what the problem is, what the respondent is purchasing, what the alternatives are, and how the solutions can be implemented results in responses that are far more informed than those of typical voters. The method we propose here obtains results closer to the actual act of voting, but certainly not the results of *well informed* voters as is the case in a standard contingent valuation survey.[56] It is much cheaper than a full fledged contingent valuation survey. The telephone survey questions on which our estimate of ε_v is based cost less than ten thousand dollars, whereas the same one thousand surveys done using in depth 45 minute personal interviews as is desirable for a standard contingent valuation survey cost between seventy-five and one hundred thousand dollars.

A standard contingent valuation survey, however, provides much more information. It provides direct estimates of the Hicksian compensating surplus, as well as the actual amounts at which a respondent would change from a yes to a no vote. Typically, the provision of different quantities of the public good in question is also explored. For many policy purposes, however,

[55] They do, however, vote in a rational fashion and in a manner consistent with their preferences on that issue. They may though hold contradictory positions the impications of which they have not confronted, and are often confused about the positions of candidates on minor issues and of the exact implications of specific legislation. See Mueller (1969), Bendict (1980), Crosby, Gill, and Taylor (1981), and Magleby (1984) for discussions. Oppenheimer (1985) argues that it may be rational in many instances for voters to be ill informed due to the small probability that they will affect referenda outcome and the nonnegligible cost of gathering information.

[56] If one is interested in predicting *actual* voter behavior as opposed to informed voter behavior, the method we prose in this paper may be preferred to a standard contingent valuation survey on that criteria.

estimates of ε_v are sufficient.

REFERENCES

Amemiya, Takeshi (1981), "Qualitative Response Models: A Survey," *Journal of Economic Literature,* 19, 1483-1536.

Baker, R.J. and J.A. Nelder (1978), *The GLIM System Release 3: Generalized Linear Interactive Modeling* (Oxford, Numerical Algorithm Group/Royal Statistical Society).

Benedict, Robert (1980), "A Test of the Direct Legislation Process: The Presence of an Informed Voter?," *The Social Science Journal,* 17, 21-39.

Bergstrom, Theodore C., Daniel L. Rubinfeld, and Perry Shapiro (1982), *Micro-Based Estimates of Demand Functions for Local School Expenditures," Econometrica,* 50, pp. 1183-1205.

Bergstrom, Theodore C. and Richard P. Goodman (1973), "Private Demands for Public Goods," *American Economic Review,* 63, 780-796.

Bettman, James R. (1979), *An Information Processing Theory of Consumer Choice* (Reading, MA: Addison-Wesley).

Bishop, Richard C. and Thomas A. Heberlein (1979), "Measuring Values of Extra-Market Goods: Are Indirect Measures Biased?," *American Journal of Agricultural Economics,* 61, pp. 926-930.

Bishop, Richard C. and Thomas A. Heberlein (1980), "Simulated Markets, Hypothetical Markets, and Travel Cost Analysis: Alternative Methods of Estimating Outdoor Recreation Demand," Staff Paper Series no. 187 (University of Wisconsin, Department of Agricultural Economics).

Bishop Richard C. Thomas A. Heberlein, and Mary Jo Kealy (1983), "Hypothetical Bias in Contingent Valuation: Results From a Simulated Market," *Natural Resources Journal,* 23, 619-633.

Black, Duncan (1958), *The Theory of Committees and Elections* (Cambridge: Cambridge University Press).

Borcherding, T.E. and R.T. Deacon (1972), "The Demand for the Services of Non-Federal Governments," *American Economic Review,* 62, 842-853.

Bowen, H.R. (1943), "The Interpretation of Voting in the Allocation of Economic Resources," *Quarterly Journal of Economics, 58, 27-43.*

Brookshire, David S., Mark A. Thayer, William P. Schulze, and Ralph C. d'Arge (1982), "Valuing Public Goods: A Comparison of Survey and Hedonic Approaches," *American Economic Review, 72, 165-176.*

Buchanan, J.M. and G. Tullock (1962), *The Calculus of Consent* (Ann Arbor: University of Michigan Press).

Carson, Richard T. and Robert Cameron Mitchell (1986), "The Value of Clean Water: The Public's Willingness to Pay for Boatable, Fishable, and Swimmable Quality Water," Discussion Paper no. QE85-08, revised, Resources for the Future.

Clawson, Marion and Jack Knetsch (1966), *Economics of Outdoor Recreation* (Baltimore: Johns Hopkins University Press for Resources for the Future).

Cochran, William G. (1977), *Sampling Techniques, 3rd. ed.* (New York: John Wiley).

Cox, D.R. (1970), *Analysis of Binary Data,* (London: Meuthen).

Courant, Paul N., Edward M. Gramlich, and Daniel L. Rubinfeld (1980), "Why Voters Support Tax Limitation Amendments: The Michigan Case," *National Tax Journal,* 33, 1-20.

Crosby, Lawrence A., James D. Gill, and James R. Taylor, "Consumer/Voter Behavior in the Passage of the Michigan Container Law," *Journal of Marketing,* 45, 19-32.

Cummings, Ronald G., David S. Brookshire, and William D. Schulze, eds. (1986), *Valuing Environmental Goods: A State of the Arts Assessment of the Contingent Valuation Method* (Totowa, NJ: Rowman and Allanheld).

Cummings, Ronald G., Louis Anthony Cox, Jr., A. Myrick Freeman, III (1984), "General Methods for Benefit Assessment," in *Evaluation of the State-of-the Art in Benefits Assessment Methods for Public Policy Purposes,* a report to the Division of Policy Research and Analysis, National Science Foundation, by Arthur D. Little.

Deacon, Robert and Perry Shapiro (1975), "Private Preference for Collective Goods Revealed Through Voting on Referenda," *American Economic Review,* 65, 943-955.

Deming, W. Edwards (1977), "An Essay on Screening, or on Two-Phase Sampling, Applied to Surveys of a Community," *International Statistical Review,* 45, 29-37.

Desvousges, William H., V. Kerry Smith, and Matthew P. McGivney (1983), " A Comparison of Alternative Approaches for Estimating Recreation and Related Benefits of Water Quality Improvements," (Washington, DC: U.S. Environmental Protection Agency, Office of Policy Analysis, EPA-230-05-83-001).

Down, Anthony (1957), *An Economic Theory of Democracy* (New York: Harper and Row). 45, 29-37.

Dragstedt, C.A. and V.F. Lang (1928), "Respiratory Stimulants in Acute Cocaine Poisoning in Rabbits," *Journal of Pharmacology and Experimental Therapeutics,* 32, 215-222.

Enelow, James M. and Melvin J. Hinich (1984), *The Spatial Theory of Voting: An Introduction* (Cambridge: Cambridge University Press).

Eu, March Fong and William G. Hamm (1984), *California Ballot Pamphlet: 1984 General Election November 6* (Sacramento: California Secretary of State).

Fienberg, Stephen E. and Judith M. Tanur (1985), "A Long and Honorable Tradition: Interwining Concepts and Constructs in Experimental Design and Sample Surveys," paper presented at the International Statistical Institute Meeting, Amsterdam.

Finney, David (1971), *Probit Analysis, 3rd ed.* (Cambridge: Cambridge University Press).

Finney, David (1978), *Statistical Methods in Biological Assay, 3rd ed.* (New York: McMillian).

Freeman, A. Myrick, III (1979), *The Benefits of Environmental Improvement: In Theory and Practice* (Baltimore: Johns Hopkins University Press for Resources for the Future).

Gibson, Betty B. (1980), "Estimating Demand Elasticities for Public Goods from Survey Data," *American Economic Review,* 70, 1069-1076.

Gramlich, Edward and Daniel Rubinfeld (1982), "Using Micro Data to Estimate Public Spending Demand Functions and Test the Tiebout and Median Voter Hypotheses," *Journal of Political Economy,* 90, 336-360.

Greenley, Douglas A., Richard G. Walsh, and Robert A. Young (1981), "Option Value: Empirical Evidence from a Case Study of Recreation and Water Quality," *Quarterly Journal of Economics,* 96, 657-672.

Groves, Theodore (1973), "Incentives in Teams," *Econometrica,* 45, 783-809.

Groves, Theodore and Adolfo Todo-Rovira (1986), "Estimation of Demand for Local Public Goods: An Alternative to the Median Voter Model," paper presented at the NBER Universities Research Conference of Government Expenditures Programs, Cambridge, November.

Hanemann, Michael W. (1984), "Discrete/Continuous Models of Consumer Demand," *Econometrica,* 52, 541-562.

Harrison, David and Daniel L. Rubinfeld (1978),"The Distribution of Benefits from Improvements in Urban Air Quality," *Journal of Environmental Economics and Management,* 5, 313-332.

Hoehn, John and Alan Randall (forthcoming), "A Satisfactory Benefit-Cost Indicator for Contingent Valuation," *Journal of Environmental Economics and Management.*

Hauck, W.W. (1983), "A Note on Confidence Bands for the Logistic Response Curve," *The American Statistician,* 37, 158-160.

Jones-Lee, M.W., M. Hammerton, and R.R. Philips (1985), "The Value of Safety: Results from a National Survey," *The Economic Journal,* 95, 49-72.

Kahneman, Daniel, Paul Slovic, and Amos Tversky (1982), *Judgement Under Uncertainty: Heuristics and Biases* (Cambridge: Cambridge University Press).

Knetsch, Jack L. and Robert K. Davis (1966), "Comparisons of Methods for Resource Evaluation," in Allen V. Kneese and Stephen C. Smith, eds., *Water Research* (Baltimore: Johns Hopkins University Press for Resources for the Future).

Lankford, R. Hamilton (1985), "Preferences of Citizens for Public Expenditures on Elementary and Secondary Education," *Journal of Econometrics,* 27, 1-20.

Lutrin, Carl. E. and Allen K. Settle (1975), "The Public and Ecology: The Role of Initiatives in California's Environmental Politics," *Western Political Quarterly,* 28, 352-371.

Magleby, David B. (1984), *Direct Legislation: Voting on Ballot Propositions in the United States* (Baltimore: Johns Hopkins University Press).

McCullagh, P. and J. Nelder (1983), *Generalized Linear Models* (London: Chapman and Hall).

McFadden, Daniel (1973), "Conditional Logit Analysis of Qualitative Choice Behavior," in P Zarembka, ed., *Frontiers in Econometrics* (New York: Academic Press).

McFadden, Daniel (1976), "The Revealed Preferences of a Government Bureaucracy: Empirical Evidence," *Bell Journal of Economics,* 7, 55-72.

Mead, R. and R.N. Curnow (1983), *Statistical Methods in Agriculture and Experimental biology* (London: Chapman and Hall).

Mitchell, Robert Cameron and Richard T. Carson (forthcoming), *Using Surveys to Value the Benefits for Public Goods: The Contingent Valuation Method* (Baltimore: Resources for the Future/Johns Hopkins University Press).

Morgan, James N. (1978), "Multiple Motives, Group Decisions, Uncertainty, Ignorance, and Confusion: A Realistic Economics of the Consumer Requires Some Psychology," *American Economic Review Papers and Proceedings,* 68, 58-63.

Mueller, Dennis C. (1979), *Public Choice* (Cambridge: Cambridge University Press).

Mueller, John E. (1969), "Voting on the Propositions: Ballot Patterns and Historical Trends in California," *American Political Science Review,* 63, 1197-1212.

Neyman, J. (1938), "Contributions to the Theory of Sampling Human Populations," *Journal of the American Statistical Association,* 33, 101-116.

Noam, Eli M. (1982), "Demand Functions and the Valuation of Public Goods," *Public Choice,* 33, 101-116.

Oppenheimer, Joe A. (1985), "Public Choice and Three Ethical Properties of Politics," *Public Choice,* 45, 241-255.

Peterson, George (1975), "Voter Demand for Public School Expenditures," in John E. Jackson, ed., *Public Needs and Private Behavior in Metropolitan Areas* (Washington: The Urban Institute).

Pregibon, D. (1981), "Logistic Regression Diagnostics," *Annals of Statistics,* 9, 704-724.

Prentice, R.L. (1976), "A Generalization of the Probit and Logit Methods for Dose Response Curves," *Biometrics,* 32, 761-768.

Radosevich, Ted C. (1975), "Electoral Analysis of the Clean Water Bond Law of 1974: Patterns of Support in a Continuing Environmental Issue," Working Paper no. 15, Institute of Governmental Studies, University of California, Berkeley..

Randall, Alan, Berry C. Ives, and Clyde Eastman (1974), "Bidding Games for Valuation of Aesthetic Environmental Improvements," *Journal of Environmental Economics and Management,* 1, 132-149.

Randall, Alan, Glenn C. Blomquist, John P. Hoehn, and John R. Stoll (1985), "National Aggregate Benefits of Air and Water Pollution Control," Interim Report to the U.S. Environmental Protection Agency.

Ridker, Ronald G. (1967), Economic Cost of Air Pollution (New York: Praeger).

Romer, Thomas and Howard Rosenthal (1979), "The Elusive Median Voter," *Journal of Public Economics,* 12, 143-179.

Samuelson, Paul A. (1954), "The Pure Theory of Public Expenditure," *Review of Economics and Statistics,* 36, 387-389.

Tiebout, C.M. (1956), "A Pure Theory of Local Public Expenditures," *Journal of Political Economy,* 64, 415-424.

Vaughan, William J. and Clifford S. Russell (1982), *Freshwater Recreational Fishing: The National Benefits of Water Pollution Control* (Baltimore: Resources for the Future/Johns Hopkins University Press).

Wolfinger, Raymond and Fred Greenstein (1968), "The Repeal of Fair Housing in California: An Analysis of Referendum Voting," *American Political Science Review,* 62, 753-769.

Zeckhauser, Richard (1973), "Voting Systems, Honest Preferences, and Pareto, Optimality," *American Political Science Review,* 67, 934-946.

Table I

VOTEREF by VOTE4

Frequency Cell Percent Row Percent Column Percent	YES	NO	DON'T KNOW	TOTAL
YES	597 58.4% 95.9% 71.8%	33 3.2% 5.1% 31.1%	13 1.3% 2.0% 15.3%	643 62.9%
NO	62 6.1% 48.4% 7.5%	57 5.6% 44.5% 53.8%	9 0.9% 7.0% 10.6%	128 12.5%
DON'T KNOW	172 16.8% 68.5% 20.7%	16 1.6% 6.4% 15.1%	63 6.2% 25.1% 74.1%	251 24.6%
TOTAL	831 81.3%	106 10.4%	85 8.3%	1022 100.0%

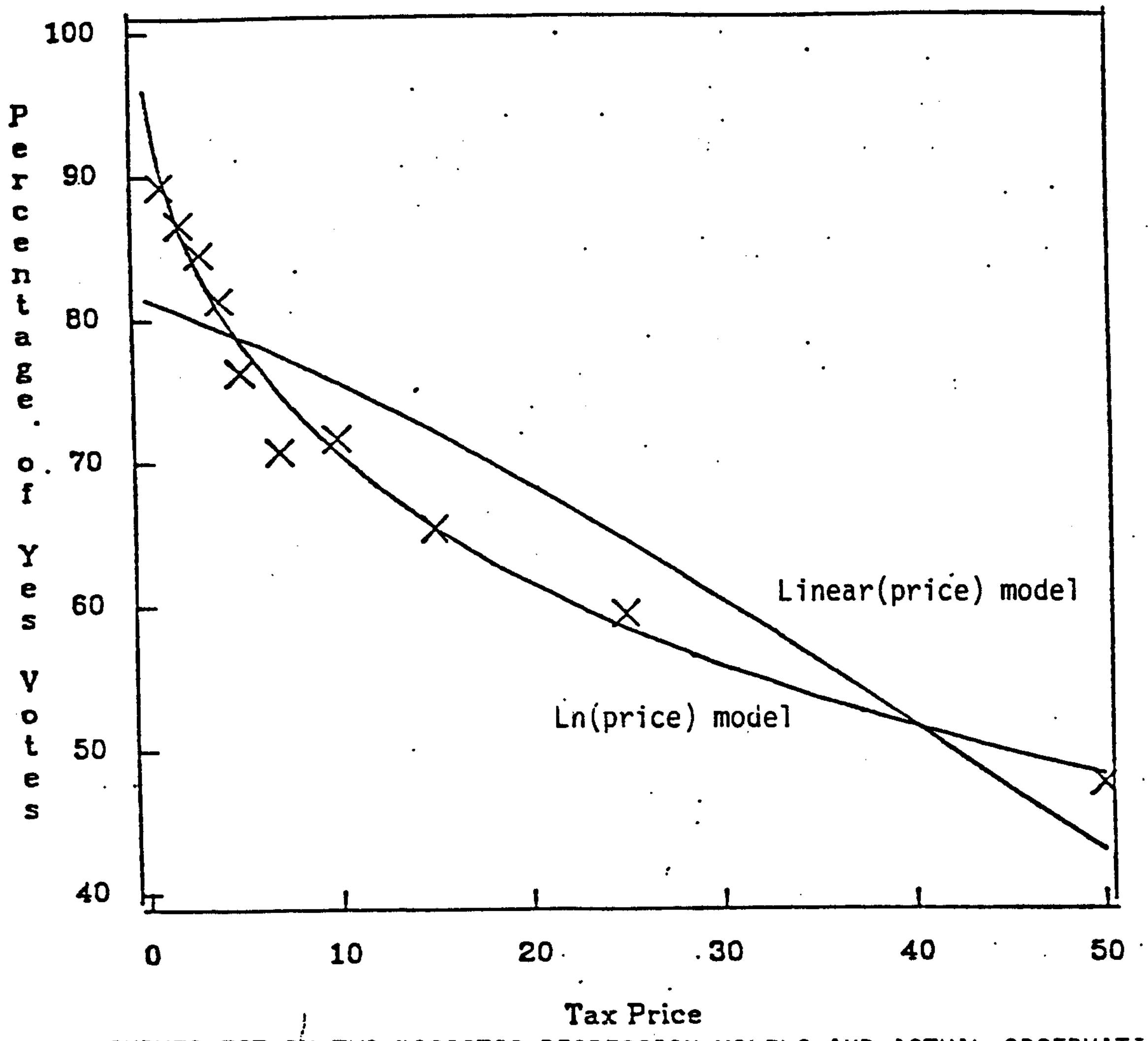
Table II

Percent in Favor of Proposition 25

Tax Price	Percent in Favor	N*	N**	N***
1	.893	28	926	302
2	.867	38	936	604
3	.846	40	938	1005
4	.813	1022	1022	1022
5	.763	146	337	682
7	.706	144	335	391
10	.717	144	335	421
15	.653	127	318	273
25	.595	127	318	260
50	.478	136	327	210

Figure 1

TAX PRICE BY PERCENTAGE OF YES VOTES



CURVES FIT BY TWO LOGISTIC REGRESSION MODELS AND ACTUAL OBSERVATIONS

Figure 2

PERCENTAGE OF YES VOTES BY TAX PRICE