Institute for Urban Environmental Risk Management
Marquette University, Milwaukee, WI

Technical Report #6

Flood Risk and Contingent Valuation Willingness to Pay Studies: A Methodological Review and Applied Analysis

Margaret C. Daun, candidate for Masters of Science in Applied Economics
David Clark, Ph.D., Department of Economics

August 2000

I. Introduction

Many environmental issues and concerns have become topics most American households are familiar with. Phrases like air pollution, global warming, the *Exxon Valdez* oil spill, Superfund, and ozone action days are now commonplace and demonstrate both policymakers' and the general public's increased cognizance of contemporary environmental debates. Of particular interest to citizens of the Milwaukee metropolitan area are the issues of flooding and more generally, watershed management practices.

Beginning with the metropolitan area's first major late 20th century flooding event in 1964 and subsequently over the course of the past thirty-five years, public awareness and reaction to flooding events has increased markedly. Major flooding events in 1986, 1997 and again in 1998 have contributed to the public's anger and exasperation as residents have attempted to cope with devastating flooding events, particularly the two consecutive floods in '97 and '98. As a result, many metro-Milwaukeeans are not only well-acquainted with some of the aforementioned environmental issues of national interest, but are also well aware of flooding and watershed concerns specific to this area. As evidenced by increased public policy initiatives, media attention and citizen outcry, the populace of metropolitan Milwaukee values flood control and perhaps relatedly, the ecological improvement of area watersheds.

From this arises a fundamental question: How can policy makers and researchers evaluate the public's perceived benefits from the maintenance of current flood risk levels and ecological improvements with the actual monetary cost of plans to accomplish these ends? Benefits from flood control can arise both directly and indirectly to individuals

and the community at large. This reflects the direct and indirect ways the environment impacts humans (see schematic diagram on pg. 7). Direct benefits accrue when the number of future flood victims is reduced. These benefits may be private (i.e. gained by an individual as the individual and/or his/her property is no longer at risk or at a lower risk level) and/or public (the community avoids future repair costs following flooding, fewer relief efforts, etc.) Indirect benefits include commercial (businesses avoiding passing on increased costs due to flooding to their consumers) and altruistic (sense of "doing the right thing" for the whole community, valuing the environment, etc.). Like direct benefits, these indirect benefits can be both private and public. The difficulty of weighing the many possible perceived benefits for the community, including all of the aforementioned public/private, direct/indirect benefits, against flood abatement plan costs is a challenging task. Herein lies this impetus for this analysis.

Specifically, this study intends to estimate the benefits of maintaining the current level of flooding risk, in terms of both spatial scope as well as frequency, and improvements in overall stream water quality that reduce ecological risks. The benefit estimates generated in this study will include private and public, indirect and direct benefits in aggregated and disaggregated form. There are a number of methods in which these benefits may be estimated, or to put it another way, in which non-market goods may be valued. This study uses Contingent Valuation (CV or CVM), specifically Willingness to Pay (WTP) surveys, to estimate these values.

II. The Contingent Valuation Method and Willingness To Pay Estimates

A. Definition of Contingent Valuation

The *Federal Register* (1994: 1142) defines contingent valuation as:

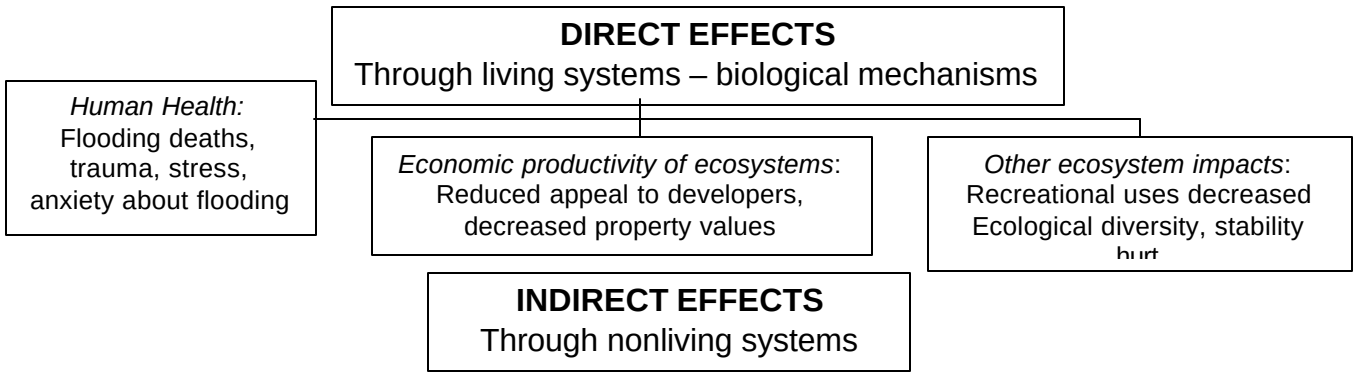
> A survey-based approach to the valuation of non-market goods and services, that relies on a questionnaire for the direct elicitation of information about the value of the good or service in question. The value obtained for the good or service is said to be contingent upon the nature of the constructed (hypothetical or simulated) market and the good or service described in the survey scenario.

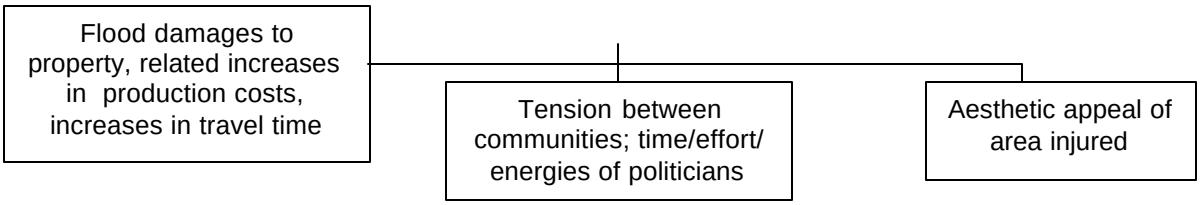
As a traditional competitive market does not exist for the sort of environmental improvements at question in this study, CVM provides a method for obtaining value estimates, specifically the amount that individuals or households are willing to pay for particular environmental goods.

B. Why CVM?

Freeman (1979: 20), summarizes and classifies the numerous channels through which environmental qualities, such as the level of flood risk and the ecological improvement of a watershed, impact humans. He states that these effects may be direct or indirect via other living organisms or inanimate systems. The following schematic diagram (Diagram 2-1) demonstrates the potential sources of direct and indirect effects as described by Freeman.

Diagram 2-1: **The Channels Through Which Environmental Changes Affect Humans**





As intuitively expected, direct effects are most easily observable and thusly, the public is more keenly aware of these impacts. However, as recent floods have attested to, indirect effects involving "materials damage" (i.e. property damage and associated costs) are more salient than other indirect effects. Consequently, in addition to the challenges posed by the valuation of both public/private and direct /indirect benefits, this study also must consider the saliency of indirect and direct environmental impacts and how this will influence people's evaluative processes.

1. The Public and Private Nature of Goods

Unlike purely private goods, for which consumers can directly (and hence truthfully) reveal their demand for through market transactions, many environmental goods are purely public goods. Pure public goods "have no explicitly identifiable individual property rights because consumers cannot be excluded from enjoying them" (Kopp and Portney, 1985). Put another way, public goods are nonrivalrous and nonexcludable. The environmental goods under consideration in this study, (i.e. the maintenance of current levels of flood risk as well as potential associated ecological improvements) have many of the qualities of public goods. However, benefits for the provision of these goods arise from the value individuals assign to these goods (Brown, 1984), which may differ as each individual's risk cognizance and evaluative processes differs. Therefore, the goods being valued in this study have characteristics of both private *and* public goods. Thus, CVM is the most appropriate methodology to utilize.

It is essential to additionally highlight that public goods can generate both use and nonuse benefits (Mitchell and Carson, 1989:55-67). 'Use' benefits include all the current direct and indirect ways an agent expects to make physical use of a public good (i.e. fishing, swimming, aesthetic enjoyment). 'Nonuse' or 'existence' benefits indicate the utility gained by an agent other than that gained from their expected personal use. Simply, existence values are those which do not require an individual to actually visit a recreational site to gain from its maintenance or improvement (Krutilla and Fisher, 1975). Mitchell and Carson (1989) further categorize existence values into two groups, vicarious consumption (by significant others, relatives or close friends, and diffuse others, the general public) and stewardship values (preservation or bequest).

However, as CVM has become more prevalent, numerous economists have sought to build upon the foundational public good valuation ideas of Ciriacy-Wantrup, 1952 and Pigou, 1952. Efforts have been made to more accurately specify the character of both use and existence values and their interrelationship and to furthermore extrapolate various types of existence benefits (i.e. Brookshire, Eubanks, and Randall, 1983; Smith, 1983; Freeman, 1984; Brookshire, Eubanks, and Sorg, 1986; Smith 1986a, 1986b). It is obvious that the maintenance of current flood risk levels and the ecological improvement of stream water quality will generate use and/or existence benefits for individuals, or households more generally.

Additionally and as aforementioned, flood management programs have private good attributes as well as their obvious public good (use and existence) attributes. Households will experience differentiable and rivalrous benefits depending upon how the flood risk level specific to each household is affected by the abatement plan (i.e. risk

reductions are not equal in relative terms). In principle, these public and private benefits or values can be measured by the area under the appropriate Hicks-compensated demand curve. However, although flood abatement plans have private as well as public good attributes, there is no immediately observable or readily accessible market for flood risk reduction, as there is for most private goods. As a consequence, households cannot directly value (i.e. reveal their demand for) the privately accruing benefits of flood management. Therefore, as aforementioned, the CVM and its many variations, is one methodology that can be usefully applied in this study, which requires the valuation of non-market private and non-market public goods[1].

2. CVM Approaches to Obtaining Benefit Information

(a) The Observed and Indirect Method

There are primarily three approaches for obtaining benefit information (Freeman, 1979). The first approach attempts to take advantage of the relationship between private marketed goods and non-marketed, public goods. Very basically, information from the related private goods market is utilized to draw inferences about the value of the public good in question. These "observed/indirect[2]" methods rely on data from situations where consumers make actual traditional market decisions, as consumers do when deciding to go on a trip or purchase a house. The value of the nonmarket good in question must then be inferred from market data available about another good to which it has an estimable linkage. A large number of assumptions are required to make this inference.

[1] Throughout the remainder of this paper, the non-market *private goods*, as explained above, may be included under the umbrella term "public" and/or "non-market" goods.

[2] Mitchell and Carson (1989:75-87) use two dimensions, how preferences are revealed and type of behavioral linkage, to distinguish classes of behavior based benefit estimation methods.

Furthermore, this method measures consumer surplus, rather than option price[3]. Hedonic pricing models are an example of this methodology. The hedonic method assumes that the price of a marketed good is a function of its various characteristics, including the characteristic of interest, usually some environmental good. Following Mitchell and Carson (1989), letting X represent a commodity class, and P_i the price of some good in commodity class X which has characteristics Q_j, generates

$$P_i = P(Q_1 \ldots Q_j \ldots Q_n) \qquad [2\text{-}1]$$

The implicit price of the characteristic of interest Q_j is then determined by differentiating P_i with respect to Q_j. Then this relationship is modified as per specific supply and demand restrictions on Q_j. Finally, consumer surplus is estimated using that demand function. The most common hedonic modeling involves property values where the housing price is assumed to include the value of some environmental service[4].

However, there are serious limitations to the hedonic pricing methodology which make it inappropriate for this study. First, the data must be such that it is possible to control for all relevant characteristics, including structural, neighborhood, and environmental. Where unique resources are already in public hands (i.e. varying levels of flood risk), this may be impossible. Second, due to the relative slowness of housing turnover, it is not easy to find genuinely comparable households in relevant

---

[3] Option price is the ex ante WTP for a specific change in the level of the amenity in question, or in this study, the avoidance of a specified reduction in the level of the amenity. This is usually considered to be the relevant welfare measure in cost-benefit analysis as policymakers must decide what policies to institute ex ante, and citizens payments for the policy change rarely depend on whether the policy actually has its intended impact (ex post). Consumer surplus, in contrast, is an ex post measure.

[4] For example, a hedonic pricing model was utilized in an earlier work by another member of this research project team to examine the impact of flood risk on residential property values (Taylor, 1999); other studies which investigate the impact of flood risk and other environmental risks on residential property values include Freeman, 1979; Schaefer, 1990; Schilling et all, 1987; Barnard, 1978; Park and Miller, 1982; Donnelly, 1989; Speyrer and Ragas, 1991. Hedonic models have also been used in valuing local shorelines and air quality (Brown and Pollakowski, 1977; Harrison and Rubinfeld, 1978; Freeman, 1979; Brookshire et al., 1982).

neighborhoods. Third, people must be aware of the actual physical differences in the levels of the characteristic being valued. Fourth, expectations about changes in the level of the good in question are presumed to enter into the determination of prices, particularly housing prices, but are generally unobservable. Finally, as per Brown and Rosen (1982), the standard assumptions used to identify systems of hedonic pricing supply and demand equations are tautological and incorrect except in extremely rare cases.

(b) The Observed and Direct Method

The second, "observed/direct," approach estimates benefit preferences as they are revealed through the voting process, by posing a question about the provision of a certain level of a public good as a referendum issue. The referendum will ask whether or not a voter is willing to support a program to supply a specific public good according to a specific payment scheme. Presuming an informed voter, the decision will be based on whether or not the voter perceives that the marginal utility of the program is greater than the marginal utility of the payment amount. However, this method is inappropriate for this study as it is not possible to generate such a politically enforced referendum.

(c) The Hypothetical and Indirect Method

The third, "hypothetical/indirect," approach utilizes a survey or interview process and asks individuals to reveal directly their willingness to pay (WTP) for a stated level of a public good or alternatively, asks what the individual would be willing to accept (WTA) in the form of a compensating payment, to forego a stated level of a public good. Due to the hypothetical nature and flexibility of CVM like WTP and WTA, ex ante judgements, which include existence values, are obtainable. Furthermore, both WTP and

WTA elicit explicit monetary valuations of public goods and therefore the outcomes of these studies are of particular value to the policymaker and the researcher alike.

(d) WTA or WTP?

The issue of whether to elicit a WTP or a WTA valuation is of central importance to this and any other CV survey. This is essentially a theoretical question of property rights. Put simply, does the agent have the right to sell the amenity in question, or if he/she wants to enjoy it, does he/she have to buy it? Clearly, as contingent valuation deals with goods that are held at least to a degree collectively, this is a difficult question to answer. For CV, perceived property rights and entitlements may be more important than legal ones in many cases.

The perceived current property right in most environmental CV cases is the status quo. Consequently, if the environmental amenity in question is an improvement upon the status quo, then the CV question should be elicited in the WTP format. In this manner the respondent is being asked to value the improvement to the status quo. Conversely, if the respondent is being asked to compare the current status quo to a new status quo where the environmental quality has degraded, then the CV question should be elicited in the WTA format.

Theoretically, WTP and WTA for an equal amount of an identical environmental good should be relatively close. However, it has been found consistently in empirical studies that WTA>WTP by a large margin[5]. Mitchell and Carson (1989) posit four hypotheses to explain the WTA>WTP difference. First, respondents to WTA are motivated to give inflated WTA or infinite WTA because they perceive the property

[5] Mitchell and Carson (1989: 30-41) describe the WTP vs. WTA debate in detail. See also Willig (1976), Randall and Stoll (1980) for seminal work in the area and Haneman's recent theoretical work.

rights implied by WTA to be implausible or illegitimate (i.e. “I refuse to sell.” or “I would not give up this good for anything less than a nearly infinite amount of compensation.”) Second, with uncertainty, time limitations for the optimization of decision-making, and risk aversion, respondents will tend to give lower WTP and higher WTA. Third, a theory known as prospect theory, developed by Kahneman and Tversky (1979), may explain the difference. Very basically, the fixed-quantity nature of public goods provision accentuates the difference between a loss of an amenity (WTA) and a gain (WTP) of the same magnitude. The respondent will tend to value the loss more than the gain, as the former implies that the good is an entitlement to be forfeited while the latter merely a potential improvement (this is even further exacerbated in the dichotomous choice elicitation format as this format implies a one-time discrete choice between two states of the world). Fourth and finally, Hanneman (1986) finds that the theoretical differences between WTP and WTA suggested earlier will be very small for public goods with many close substitutes and very large (perhaps infinite mathematically) for public goods with few substitution possibilities.

This study asks respondents to value two types of goods, the ecological improvement of the river in question and the maintenance of current levels of flooding risk. The former is a clear example of the sort of good requiring a WTP question since it involves an improvement from the status quo. The latter seems to require a WTA question. However, Mitchell and Carson (1989: 40-41) state that when an expenditure is needed to maintain the level of a public good currently available, and when the level of that public good will deteriorate without the expenditure, the appropriate elicitation is WTP. Therefore, the flood risk questions will also be posed using the WTP format.

C. History of CVM

The use of a direct interview to measure values associated with public environmental goods was first suggested by Ciriacy-Wantrup (1947). His 1952 text, *Resource Conservation: Economics and Policies*, expanded and more fully developed his methodology. However, the seminal application of the CVM was conducted by Davis in the early 1960's (1963a, 1963b, 1964) in his survey of 121 hunters and recreationalists in the Maine backwoods. Ridker (1967) applied CV in a study of air pollution benefits. As highlighted by Mitchell and Carson (1989), Ridker made a prophetic observation following the conclusion of his research:

> It now seems evident that a much narrower, deeper, and psychologically sophisticated questionnaire is required to measure and untangle the various determinants of cleaning costs. Such a questionnaire would require substantially more time and expenditure – perhaps three to four times greater – than went into this essentially exploratory study. Even then the possibility of success may not be very high, for such a project would raise problems in survey design that have not yet been solved. (Ridker, 1967: 84)

A number of economists in the 1970s continued to expand the applications of the CV approach. Studies included valuations of hunting rights (Hammack and Brown, 1974), reduction in recreational congestion (Cicchetti and Smith, 1973, 1976a, 1976b), urban parks (Darling, 1973), pollution control and air visibility (Eastman, Randall, and Hoffer, 1974, 1978; Randall, Ives, and Eastman, 1974), heart attack risk reduction (Acton, 1973), and improved beachside water quality (Hanemann, 1978; Binkley and Haneman, 1978). As expected, applications, results comparisons, and empirical investigations of CVM continued and expanded throughout the 1980's. Mitchell and Carson (1989: 9-14, 307-315) provide a more detailed and complete listing of numerous

CV studies, the public goods they valued and some of their characteristics, as well as descriptions of "much of the pioneering methodological work (12-13)."

It is also crucial to point out that complementing researchers' and academia's growing interest in CVM, policymakers and government agencies also began to increasingly advocate and employ the use of CVM. In 1972, the Water Resources Council published its "Principles and Standards for Water and Related Land Resources Planning" in the *Federal Register*. This document and its subsequently revised version (1979) established crucial guidelines and formal clarifications for federal project benefit evaluations and specified acceptable methodologies, based on consumer preference economic theory. It recommended three methodologies for benefit estimations: travel cost, unit day,[6] and contingent valuation (see above definition). In regards to CV, the document (1973) stated:

> The general measurement standard of the value of goods and services is defined as the willingness of users to pay for each increment of output from a plan.

As elucidated by Thunberg (1988:2),

> For flood control projects, benefits could logically be derived from both a consumer WTP to avoid damages to property and to avoid the social and psychological dislocations associated with flooding.

Additionally, following these developments, the U.S. Army Corps of Engineers and its Institute for Water Resources began to employ CV extensively (Mitchell and Carson, 1989). Furthermore, beginning in the 1970's, the Environmental Protection Agency (EPA) has devoted large funding amounts to the exploration of CVM and its

[6] Travel or time costs may serve as proxies for market prices in estimating the demand for publicly provided, site-specific recreational services or other publicly provided nature areas. Travel or time data is obtained by surveys. These data may be used to develop of a demand curve for the site as a function of distance or time. If a monetary value is assigned for each mile or time unit (minute, quarter or half hour, etc.), a consumer surplus calculation can be made for easily be made. For a basic description of this

potential applications in policymaking cost-benefit analyses (Smith, 1984). To this end, the EPA commissioned an in-depth assessment of CVM, which convened in Palo Alto, California in July of 1984 and included a review panel of leading CV economists and psychologists (Mitchell and Carson, 1989). The panel concluded (Cummings et al., 1986) that the method showed much promise and gave the highest priority for further research to the development of an overall framework.

However, the legislation with the most expansive impact on CV was the Comprehensive Environmental Response, Compensation, and Liabilities Act of 1980 (CERCLA or Superfund). The act provides regulations to assess damages for the injury, destruction or loss of natural resources. In order to identify the "best available procedures to determine such damages" the regulations

> …shall take into account consideration factors, including, but not limited to, replacement value, use value, and the ability of the ecosystem or resource to recover.[7]

The Department of the Interior (DOI) subsequently embarked on a lengthy decision making process in an effort to determine the definition of the appropriate measure of damages.[8] This precursory debate foreshadowed lengthy and protracted legal and legislative battles, which continue today over the use of CVM for damage assessments.

This legal and legislative battle has been centered about two poles. The first pole is characterized by individuals, groups, organizations, and institutions with generally a liberal political philosophy and whose primary objective is the preservation, protection and restoration of the environment and its natural resources. These groups clearly

---

method, see Freeman (1979); for applications and discussion see McConnell (1985), Mendelsohn and Brown (1983), Binkley and Hanneman (1978), and Bockstael, Hanneman, and Strand (1985).

[7] CERCLA section 301(c), 42 U.S.C § 9651 (c).

[8] "Resolution of the definition of damage to be used in this proposed rule is of such importance that it has already been discussed at length…" Department of the Interior, Proposed rulemaking, "Natural resource Damage Assessments," 50 Fed. Reg. 52126, 52147 (Dec. 20, 1985)(1985 Interior Proposed Rule).

advocate legislation and legal decisions which favor the legitimization and the comprehensive, wide-reaching acceptance of CVM and its related methodologies, such that nonuse, option, and existence values are captured in damage and benefit estimates. These groups argue that as CV methodologies continually improve and as universal methodological guidelines are set (by the very legislation being debated), CVM will be a wholly reliable and accurate damage/benefit estimator. The opposite pole, on the other hand, is generally characterized by business interests and environmental conservatives. They advocate strict legal and legislative limitations on the usage of CVM and its related methodologies. They argue that as the scientific community debates the accuracy of CVM estimates, particularly given potential sources of bias (see section II.D.), and concerns that CVM inflates damage and/or benefit estimates, the methodology should be used only as a last resort. As intuitively expected, political trends and interest group pressures have played large roles and continue to significantly impact this continuing debate.

The preliminary ruling of the DOI in 1985 created a measure of damages with two central components: a 'lesser of' rule and a hierarchy of economic methodologies. The 'lesser of' rule provided that the amount which may be recovered for damages be either the cost of restoration, replacement or diminution of direct use values, whichever was less. The hierarchy of measurement methodologies for the measurement of the diminution of direct use values was specified thusly: measure the loss directly through the diminution in market price; if the resource was not traded in a market, the appraisal method was to be attempted next; if neither of these first two methods were feasible, nonmarket methodologies were considered last, including travel costs, hedonic pricing,

unit value, factor income, and CV. However, CV was to be used only to "to estimate option and existence values…if the authorized official determines that no other valuation technique will be feasible."[9] In its final ruling in 1986, the DOI further limited the use of CV by stating that the estimation of existence and option values could only be conducted when use values could not be determined.[10]

As summarized by Kopp and Pease (1997: 14-15),

> The combination of the 'lesser of' rule and the hierarchy of measurement methodologies and the restrictions on estimating passive [non] use values *guarantee undervaluation* [emphasis added] of injuries to natural resources… Undoubtedly, this approach was appealing to those likely to be held liable for the costs, but was certain to raise concerns among come trustees and other interested groups. With the considerable attention the Interior's rulemaking had generated, it is not surprising that litigation ensued and that several of the areas challenged concerned measure of damages and use of CV.

Kopp and Pease (1997) describe in great detail the resultant litigation and its policy and legal implications for resource preservation and CVM. Following litigation the DOI was forced to revise its damage assessment policies, acknowledging CV as a valid damage assessment methodology and revoking its hierarchy of economic methodologies.[11]

While the DOI continued to revise its 1986 rulings, the National Oceanic & Atmospheric Administration (NOAA) was forced to begin to develop damage assessment regulations of its own. The catalyst for this regulatory activity was the 1989 *Exxon Valdez* oil spill in Prince William Sound, Alaska. Congress subsequently passed the Oil Pollution Act (OPA) in August of 1990, which created a new compensation recovery mechanism designed to more fully address the expansive impacts of oil spills. Furthermore, the OPA required the NOAA to generate damage assessment regulations for

[9] 1985 Interior Proposed Rule at 52143; 43 C.F.R. § 11.83
[10] 43 C.F.R. § 11.83(d)(5).

spillage incidents. Kopp and Pease (1993:18) identify seven socio-political factors which caused the debate over CV to intensify. Following the example of the EPA, on April 8, 1992, the NOAA convened an expert panel, chaired by two Nobel laureates, to evaluate the reliability of CV. The Panel's report, released on January 15, 1993, concluded that if CV studies follow the Panel's suggested guidelines, the resulting estimates are "reliable enough to be the starting point of a judicial process of damage assessment, including lost passive-use values." Kopp and Pease (1997:18-23) detail the subsequent hotly debated and litigious evolution of NOAA regulations.

Kopp and Pease (1997) close their discussion of the legal-legislative-political CV controversy by highlighting the current interagency effort to consider appropriate guidelines for CV estimation of nonuse values. Ronald Reagan's Executive Order 12291 of February 19, 1981, requires agencies proposing new regulations to provide the Office of Management and Budget (OMB) a regulatory impact analysis (RIA) for such regulations, a central part of which is a formal cost-benefit analysis of the proposed regulation.[12] Seeing that it is generally recognized that total value estimated can only be obtained using CV, as the EPA and other agencies have moved toward benefit calculations based on a total value concept for RIAs, CV regulations and rulemakings are of crucial importance. To consider the appropriate guidelines for CV estimates of nonuse values and the role of these estimates in RIAs, an interagency working group has been created. The group is comprised of OMB, EPA, the Department of Energy, the President's Council of Economic Advisors (CEA) and several other agencies. Clearly,

---

[11] Department of the Interior, Notice of Proposed Rulemaking, "Natural Resource damage Assessments," 56 Fed. Reg. 19752, 19759 (Apr. 29, 1991)(1991 Interior Proposed Rule).

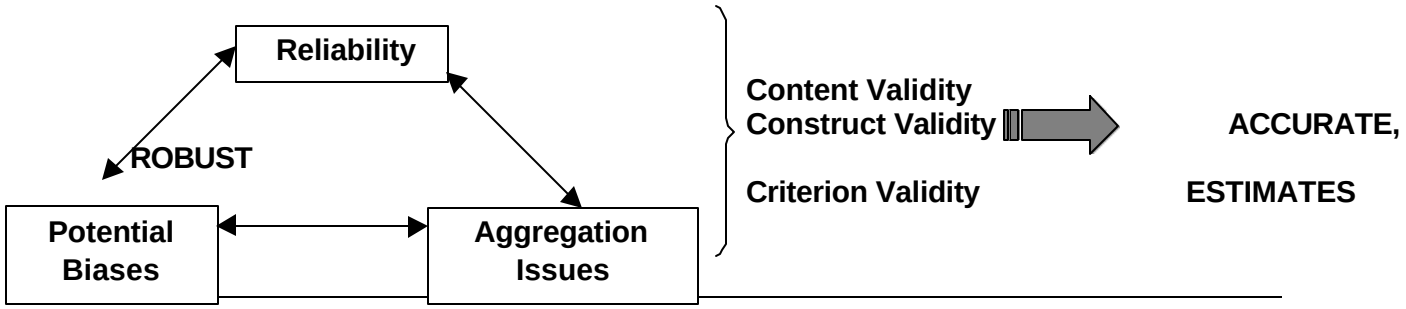
CVM is an integrally important aspect of public policy formation and will continue to be of concern for relevant federal, state and local agencies and policymakers.

D. CVM and potential sources of bias

Having discussed in relative detail the history of CVM and the current legal socio-political debate surrounding CVM, it is crucial to address the question: What creates skepticism about CVM? As Ridker's quote earlier alluded to, skepticism for CVM is grounded in both operational and methodological problems (various biases), the implications of psychology for CVM, the lack of precise guidelines for CVM implementation, and the lack of specific standards for evaluating CVM results.

Mitchell and Carson (1989:120-125, 211-293) detail the relationship between accuracy in benefit or damage estimates and its dependence upon the validity of the data, analyses, and overall methodology that are used in measurement. Validity (of which there are three types, see below discussion) in turn, depends on how the data and methodologies deal with potential biases, aggregation issues,[13] and reliability:

Diagram 2-2: **Basic Interrelations Between Accuracy, Validity, Reliability, Biases, and Aggregation**



[12] The executive order stated that "[r]egulatory action should not be undertaken unless the potential benefits to society for the regulation outweigh the potential costs" and "[r]egulatory objectives shall be chosen to maximize the net benefits to society." 46 Fed. Reg. 13193 *et seq.* (Feb. 19, 1981)

[13] Once a correct theoretical measure has been obtained for a sample of individuals through CVM, two key aggregation questions are then raised: is it possible (i.e. statistically valid and meaningful) to add up the WTP amounts obtained for individuals to get an aggregate benefit estimate for the relevant population? and is it possible to aggregate these amounts across subcomponents to obtain the relevant total benefit? (Samuelson, 1954; Mitchell and Carson, 1989)

Mitchell and Carson (1989) specify approximately five general sources for biases in CV studies:

(1) incentives to misrepresent responses (i.e. strategic biases);
(2) implied value cues (i.e. starting point biases, range biases and relational biases);
(3) scenario misspecification (i.e. payment vehicle biases, elicitation biases, and symbolic biases);
(4) improper sampling design or execution (i.e. population choice bias, sampling frame bias, sample nonresponse bias, sample selection bias); and
(5) improper benefit aggregation (i.e. temporal selection bias and sequence aggregation bias).

Bishop, Champ, Brown and McCollum (1993) offer basic guidelines, echoed by many other economists (Freeman, 1979; Mitchell and Carson, 1989; etc.) in order to increase the robustness of CV and WTP studies. As reflected in the above schematic, the authors first highlight that accurate estimates of WTP can be made even when unreliability exists by simply increasing sampling size and utilizing theoretically viable sampling techniques. Even when unreliability exists in observed WTP, this is tolerable so long as bias is within tolerable bounds. Of greater concern to WTP studies is on the validity of estimated average WTP for correctly defined aggregates of individuals.

Content validity addresses whether the design and execution of a study and its survey methodology are "conducive to the revelation of WTP" (Bishop et al, 1993:65). Most simply, a valid study will craft a hypothetical scenario such that the theoretical consumer will reveal true WTP. Fischoff and Furby (1988:148) state that a successful CV framework will fulfill the requirements of a satisfactory transaction, which they define as one "involving individuals who are fully informed, uncoerced, and able to identify their own best interests." Therefore, three aspects of the transaction must be defined with careful, adequate explicitness and understood correctly by survey

respondents: the good to be valued, the payment vehicle, and the context of the transaction. Focus groups, pretesting, and the availability of a uniform 'frequently asked questions' response guide for interviewers can therefore greatly improve the content validity of a WTP study. Content validity is consequently a largely relative issue, with differing implications for every analysis[14].

Construct validity addresses how the variable of interest or the value being estimated (i.e. WTP) is related to other variables as predicted by theory. Simply put, when using regression analyses, do the variables which theory indicates would influence WTP have the expected signs, statistical significance and appropriate relative magnitudes? This is therefore an issue of correct model specification and econometric methodology.

Criterion validity is assessed by comparing the contingent value (i.e. WTP estimate) to the criterion, "which [should be] unequivocally closer to the theoretical construct (Mitchell and Carson 1989: 192)." Basically, this involves comparing the WTP estimates found in the analysis of one sample with another estimate generated from a 'simulated market' or an actual referenda. The closer the WTP estimate to the criterion, the more valid it is judged to be. However, it is clear that for non-use values it is extremely difficult, if not impossible to create a 'simulated market[15].'

E. Relevant Literature: WTP Studies and Flood Risk

As explicated in detail above, there are numerous ways to derive a valuation of the maintenance of current flood risk levels and/or the ecological improvements of

[14] Focus groups, pretesting, and a FAQ page were all utilized in this study. The appendix includes the completed survey and the FAQ guide. Section III discusses survey development and the relevant potential biases.

watersheds. These include analysis of the relationship between private marketed goods and non-marketed, public goods, analysis of individuals' preferences as they are revealed through the voting process (referendum), and the utilization of a survey or interview process that asks individuals to reveal directly their willingness to pay (WTP) for a stated level of a public good. This study uses a CVM survey, sampling the residents of two watersheds, in order to estimate the population's mean aggregate WTP for the maintenance of status quo flooding risk and/or ecological improvements of the watersheds.

The willingness to pay model has its roots in the work of Ciriacy-Wantrup (1947, 1952) and more substantially Davis (1963a, 1963b, 1964). As aforementioned, the survey WTP approach has been used extensively in the early stages of CVM development (Ridker, 1967; Hammack and Brown, 1974; Cicchetti and Smith, 1973, 1976a, 1976b; Darling, 1973; Eastman, Randall, and Hoffer, 1974, 1978; Randall, Ives, and Eastman, 1974; Hanemann, 1978; Binkley and Haneman, 1978; etc.). The use of surveys to estimate WTP is well established (Freeman, 1979; Mitchell and Carson, 1989; Freeman, 1993). The existing body of literature regarding the ability of CV surveys to estimate WTP is expansive.

The vast majority of this literature addresses the WTP for certain levels of a public environmental good. CVM surveys have been used extensively to estimate a monetary value for non-marketed natural resource goods and/or services (Cummings and Harrison, 1992; Carson et al., 1994) and for determining the value of environmental damages (see earlier discussion, section II.C). Another large portion of the literature

---

[15] Hedonic pricing models and benefit transfer models are two possible methods with which a simulated market could be created against which to compare WTP estimates.

addresses WTP survey elicitation issues (i.e. dichotomous-choice or discrete-choice (DC), open-ended (OE), "cheap talk" formats[16]), usually and not surprisingly in terms of some environmental public good (Langford et al.,1998; Haab and McConnell, 1998; Halvorsen and Sælensminde, 1998; Green et al., 1998; Alvarez-Farizo et al., 1999; Byrnes, Jones, Goodman, 1999; McLeod and Bergland, 1999; Cummings and Taylor, 1999; Werner, 1999, etc.) Another subset of WTP literature focuses on critiquing WTP estimation validity and its ability to accurately predict *actual* payments (Acutt and Mason, 1998; Willis and Powe, 1998; Taylor, 1998; Johannesson et al., 1999, etc.)

Of particular interest to this study is the literature which addresses environmental risks and WTP, specifically flooding risks, as well as literature which explores DC and OE elicitation methods, as the survey was structured with iterative bid DC questions and a final OE question. First, the literature addressing risk reduction WTP studies will be discussed and second, studies addressing the DC and OE elicitation approaches will be treated.

1. WTP and Risk Reduction

As highlighted by Thunberg (1988:6), few applications of CVM have been made to flood control specifically (Thunberg, 1988; Shabman and Stephenson, 1996; Johnston, Swallow and Weaver, 1999; and a few nebulous early applications by the U.S. Army Corps of Engineers[17]). More authors have used WTP estimates to evaluate preferences

[16] "Cheap talk" questionnaire formats refer to the explicit discussion of the hypothetical bias problem (what it is and why it may occur) with WTP survey respondents in lieu of references to budgetary constraints and substitutes (Cummings and Taylor, 1999).

[17] With the 1936 Flood control Act (FCA), the Army Corps of Engineers was established and as their policy and research evolved, benefits from flood protection came to be defined as the present value of property damages avoided. This definition has expanded following a formal clarification of project benefit evaluation in 1972 (a WTP measure of property damage costs avoided *and* psychological dislocations associated with flooding). The Corps' application of this new broader WTP has been sporadic and the methodology used inconsistent.

for reducing/maintaining other types of environmental risk levels: WTP for public hazardous waste risk reduction (Smith and Devousages, 1987; Romer, Pommerehne, and Feld, 1998), a comparison of valuation methodologies for health risks (Berger, Blomquist, Kenkel, and Tolley, 1987), valuing changes in generic environmental risks (Freeman, 1991), risk and self-protection (Schogren and Crocker, 1991), and WTP for air quality damage contrasted against conjoint valuation methods (Mundy and McLean, 1998). The three aforementioned articles dealing directly with flood risk and watershed management will be described below as well as the two studies that address specifically the reduction of public risks due to hazardous waste.

Thunberg (1988) surveys a residential neighborhood in an area in Southeast Roanoke, Virginia comprised of 142 properties, located within the floodplain of the Roanoke River during a two week period in 1987. Guided by earlier investigations of the floodplain by the Army Corps of Engineers, Thunberg, estimates WTP using log-log, lin-log and purely linear functional forms. As all of Thunberg's specified models are linear in their parameters, Ordinary Least Squares is used as the estimation method. In his models Thunberg includes the following explanatory variables: a flood zone (FZ) indexed variable based upon frequency of flood occurrence for a particular parcel, an expected structure plus contents value benefits (ESVB) variable defined as the product of the subjective assessment of the probability of flood control (DPROB) and the structure plus contents value of the home, an expected anxiety relief benefit (EANXB) variable defined as the product of DPROB and the individual's level of anxiety, a tenancy status (TENN) dummy variable, a community effects (COMM) variable defined as the individual's subjective assessment of the potential benefits for the community at large, an

expectations for property value changes (PROPV) variable defined as the individual's subjective assessment of potential increased in property values, a dummy variable controlling for landowners that have immediate plans to sell their property (SELL), a flood insurance premium expectations (EFIP) variable defined as the landowner's subjective assessment of the potential change in flood insurance premiums, and an income (INC) variable.

Thunberg (1988:85) states that the "hypothesis that WTP is positively related to flood zone is consistent with an expected future flow of protection services definition of flood control benefits." However, Thunberg empirically rejects this hypothesis. Regardless of whether Thunberg was estimating WTP using a lump sum or annual payment scheme, the flood zone variable was significant at the 10% level with a *negative* sign. Thunberg offers three hypotheses for these unexpected and surprising findings: strategic behavior (when a respondent perceives an incentive to behave strategically, over or underbidding WTP, depending upon the uncertainty of the payment vehicle and risk adverseness[18]), self insurance by locational choice (i.e. flood plain landowners select a flood-plain location on the basis of a relative trade-off between consumption opportunities and safety from flooding), feelings of victimization, frustration and distrust of government. Thunberg indicates that these findings are largely due to empirical issues and model specification problems and argues that the theoretically positive relationship between WTP and the reduction of flood risk is not questioned by his findings.

[18] If the respondent's expected payment obligation is perceived to be less than the amount that he states, he will tend to overbid; if the expected obligation is larger than the stated amount, he will tend to underbid or free ride (Mitchell and Carson, 1989).

Thunberg's results have one critical implication for this study: it is crucially important to assess the explanatory power of attitudinal/behavioral variables. Including such variables may allow for the identification of respondents that believe that they have self-insured by their locational choice (or, more likely, that others *should have* chosen their locations more carefully) and for the assessment of a respondent's relative feelings of frustration, victimization, and trust or mistrust of government.

Lenord Shabman and Kurt Stephenson in their 1996 article, "Searching for the Correct Benefit Estimate: Empirical Evidence for an Alternative Perspective," investigate flood risk reduction benefits for a flood control project in Roanoke, Virginia (the same geographical area investigated by Thunberg). Their analysis compares residential flood risk reduction benefit estimates from the property damages avoided (PDA), hedonic price, and CVM techniques. These estimates are then contrasted against actual voting behavior on a referendum considering the flood control project. Their research references and compares previous empirical investigations of the flood risk reduction benefits from the construction of the Roanoke flood control project (U.S. Army Corps of Engineers, 1984; Thunberg, 1988; Thunberg and Shabman, 1991; Driscoll, Dietz, and Alwag, 1994).

An area in southeast Roanoke, VA consisting of homogeneous, residential properties was the study site. This particular neighborhood consisted of 134 homes. The U.S. Army Corps of Engineers (1984) used the PDA method[19]. Driscoll, Dietz, and Alwag (1994) estimated a hedonic price equation using 99 property transactions sold at

[19] The property damages avoided technique calculates the repair costs to a specific property with and without a flood risk reduction project for a given flood event. The difference between the repair costs is the annual flood risk reduction benefit estimate for that flood. Then each benefit estimate is weighted according to the likelihood of the flood event in a year and sums over all possible floods to calculate the

fair market prices between 1980 and 1990, excluding properties not repaired after a flood event in 1985. The final estimated equation indicated that land traders were willing to pay higher prices for property that was less likely to be flooded after the 1985 flood. In other words, after 1985 homebuyers became more flood risk sensitive. Thunberg (1988) used the CVM survey method to elicit WTP responses for flood risk reduction (see above discussion of his paper). Two hypothetical payment methods were presented to the respondent: WTP in a onetime assessment (lump-sum) and WTP each year for fifteen years.

Shabman and Stephenson then compare the three benefit estimates resulting from the three aforementioned methodologies amongst themselves and to revealed voting behavior[20]. For an identical set of residential property owners, there was a PDA, an hedonic price estimate, an annual and lump-sum WTP bid, and evidence of the owner's voting behavior. A total of 73 observations were included in the analysis. The following table summarizes the comparisons:

Table 2-1: **Comparison of Mean Benefit Estimates in $**

Flood Zone*	**Hedonic**	**PDA**	**CVM** – includes all bids	**CVM** – excludes protest and uncertain bids
FZ < .05	4.424	2,412	115	230
.02<FZ≤.05	2,331	908	203	381
.01<FZ≤.01	914	419	980	1,225
.002<FZ≤.01	316	99	223	322
.0001<FZ≤.002	196	15	240	429
ALL FZ	1,333	597	314	520

* The flood zone = the probability of being flooded in any given year

annual expected benefits for that property. The present value of the expected annual benefits is discounted, resulting in the total benefit to that property.

[20] Voting records were used to determine whether or not a property owner voted in the Roanoke flood control referendum in which the flood control project in question was identical to the one described to them a year and a half earlier in the WTP study. The annual payment WTP question was similar to the way in which citizens were actually asked to pay for the project.

Source: Shabman and Stephenson (1996)

As is clear from the above table, certain patterns in the magnitudes of the estimations are apparent. By far, the hedonic price technique generated the largest estimates, with the gap between the estimates greatest for the most flood prone areas. By extension, these hedonic estimates indicate that people were willing to pay a substantial premium for reductions in flood risk. Of the 73 CVM respondents, 18 were protest bidders and 11 were uncertain about their bids. These were coded as zero bids. Regardless of whether or not the CVM include or exclude the protest and uncertain bids the same general pattern is revealed: average WTP bids were substantially lower than PDA and hedonic estimates. When comparing the WTP bids with actual voting behavior, the authors note that only 28 of the 73 CVM respondents voted in the referendum. The authors demonstrate that overall, nonregistered voters were more likely to offer a positive WTP. In addition, their WTP was higher, on average, than the other voter groups (registered and voted, registered and did not vote). In conclusion, Shabman and Stephenson state that the hypothetical CVM bids did not translate into actual choice behavior.

The authors draw several interesting implications from their comparisons:

1) Their findings vividly illustrate the protracted debates over which estimation technique provides the correct valuation of nonmarketed goods. More basically, their work highlights two fundamental questions about CVM: is CVM reliable and is it accurate?
2) This debate rests on the premise that there exists a single "true" value for which an accurate and unbiased estimate is being made. This search is itself based on (i) the neoclassical utility maximization model's behavioral postulate that individuals have stable, precise, and well-defined preferences, and (ii) that only the limitations of economists' nonmarket valuation techniques stand as a barrier to calculating the true value. Shabman and Stephenson, noting research on choice by psychologists, state that "there may be no single true behavior if preferences vary across time and between choice-making circumstances" (441). In other words, is it possible for individuals to actually overcome complex cognitive difficulties? These cognitive

difficulties are apparent both in WTP surveys, which are inherently complex in their construct, but also in hedonic price studies. When housing purchases are in question, people are faced with a complex and uncertain decision environment (home buyers must evaluate a myriad of factors – number of rooms, lot size, the messy neighbors, quality of local schools, the condition of the roof and the plumbing, etc., etc.)

3) Moral considerations and equity concerns substantially influence WTP bids.
4) The difference between stating an opinion (preference revealed in the WTP bid) and making a choice (acting by voting for the referendum) is not a trivial one.

The authors conclude by stating that the search for a "true," "correct," or "unbiased" benefit is a futile one and that differences in benefit estimates from different techniques are to be expected due to choice-making circumstances (444).

Shabman and Stephenson's work has three critical implications for analysis at hand[21]. First and foremost, the most conservative estimation technique, namely a WTP CVM analysis, should be used as it is least likely to overestimate a true benefit estimate. Second, the estimation of a "true" WTP can be much more than a futile exercise is it attempts to take into consideration the moral, ethical, civic, and equity concerns and cognitive difficulties faced by respondents. This echoes the implications gleaned from Thunberg (1988). Third and finally, cognitive challenges faced by respondents should not be underestimated.

In Robert Johnson, Stephen Swallow, and Thomas Weaver's 1999 article, "Estimating Willingness to Pay and Resource Tradeoffs with Different Payment Mechanisms: An Evaluation of a Funding Guarantee for Watershed Management," they contend that the public's willingness to provide funds for environmental improvements, as well as their view of the likelihood of outcomes from policy changes, may be significantly impacted by the chosen funding mechanisms (e.g. taxes, bond issues, earmarked contributions).

Seeing that the typical CVM survey generally considers the payment mechanism as a simple means to assess realistic monetary tradeoffs and measure dollar-denominated welfare impacts, their study specifically addresses the issue of whether the respondents' faith in the payment mechanism impacts WTP. In other words, typical CVM surveys do not allow for the possibility that respondents may assume that the government is inefficient in its ability to restrict new funds to the specific program in question, albeit hypothetically. Very basically, and mirroring Thunberg's (1988) conclusion, do respondents' mistrust of government, as demonstrated through their payment mechanism preferences, significantly impact WTP? The authors further state that even in studies where the payment mechanisms are designed consistently within the CVM literature's guidelines to minimize hypothetical or payment-vehicle bias, respondents may nonetheless act under an assumption of some level of inefficiency in government agencies' use of collected funds. The study area used is a relatively rural area in southwestern Rhode Island, valued for its pristine environment and natural resource base. The study asked 301 survey respondents to value various watershed management policy packages, varying across 6 different payment mechanisms (total usable responses = 1800 out of a possible 1806).

Their results indicated that the institutional context of a 100% funding guarantee, as implied by the payment vehicle, influenced respondents ranking of various policy packages[22]. Respondents presented with a funding guarantee chose to accept the described package at a different rate than respondent for whom no guarantee was

[21] It should also be noted that their sample size does not meet the minimum requirements set by Mitchell and Carson (1989).

[22] Their analysis is both intricate and lengthy and thus, not appropriate for detailed dissection in this paper. Only the primary conclusions and implications for this study will therefore be discussed.

provided, *ceteris paribus.* In addition, the funding guarantee positively influenced respondents' WTP for changes in specific watershed program attributes. In sum, the authors alert researchers assessing the validity of CVM to ensure the characteristics of the payment vehicles do not themselves generate *valid* differences between WTP estimates.

These conclusions of Johnston, Swallow, and Weaver highlight three critical points for the WTP estimates in this study. First, the choice of a WTP CVM study for this research is viable and the resulting estimations cannot be predetermined to be invalid due to our selection of a payment vehicle. Furthermore, it is therefore highly appropriate to purposely describe the payment vehicle in general terms[23]. Finally, a variable measuring respondents' relative trust/mistrust of government must be investigated or alternatively, potential payment vehicle bias investigated[23].

Smith and Desvouges (1987), use a CVM survey instrument to assess individual's WTP for reductions in the risk levels from exposure to hazardous materials. Their primary sample consisted of approximately 485 households in the former city of West Berlin (Federal Republic of Germany). They drew eight sub-samples from this larger whole sample and assigned each sub-sample a different initial risk level. Each respondent was offered two risk reductions: the first reduced risk by 50% and the second by an additional 60%. The value of a risk reduction (i.e. expressed willingness to pay) was found to be constant irrespective of the initial baseline risk level and the size of the risk reduction offered, seemingly contradicting standard expected utility theory.[24] The authors posit that these puzzling results point to the embedding phenomena widely discussed in CVM literature (see below).

---

[23] See further discussion below in section III.B and section IV.D.1.d.

Romer, Pommerehne and Feld (1998) expand upon Smith and Desvouges (1987) and Thunberg (1988). The authors also report the results of a CV survey on the reduction of hazardous waste risk in the former city of West Berlin (Federal Republic of Germany), with a total of 166 household interviews. The authors focus special attention on the possibility for individual risk averting behavior, potential embedding phenomena, comprised of scope and sequence (i.e. anchoring) aspects[25], which they point to as a probable causes of the 'surprising' conclusions reached by Smith and Desvouges, as well as the treatment of zero and protest bids.

Initially, using mean bid analysis, the authors found that even at the 10% significance level (t-value:1.25) they could not reject the hypothesis that WTP for a higher level of risk reduction starting from a higher initial risk level was independently higher than a lower risk reduction starting from a lower initial risk level. However, when the respondents who indicated in their responses to an open-ended question that they considered the initial risk levels to be significantly exaggerated were deleted (17 total respondents), the hypothesis of WTP for higher levels of risk reduction being equal to or less than WTP for lower levels of risk reduction can be rejected at the 5% level of significance (t-value 1.66).

Based upon these initial results, the authors go on to use non-parametric statistical tests to investigate the similarity of the means of these WTP for the two types of risk in

---

[24] It would be expected that offering a second risk reduction would result in lower WTP if decreasing marginal utility of the risk reduction were to hold.

[25] In Romer, Feld, and Pommerehne's study, the same level of risk reduction was valued differently depending upon whether the risk reduction level was offered as the first or second reduction. Also, second risk reductions offered in a follow-up question format generated protest and strategic bids reflecting the skepticism of the respondents concerning the questionnaire validity (anchoring). Scope effects occurred when different levels of risk reduction did not elicit proportionately larger or smaller WTP responses. Most generally, embedding effects occur when the value of the first WTP question and the differentials between the WTP values in subsequent questions influence the respondent's replies in a DC format.

question. Next, the authors use multiple regression analysis in order to include other major determinants of WTP other than baseline risk levels. The authors then offer two possibilities to cope with zero and protest bids: (the deletion of protest bids from the sample is considered and dismissed as far too rudimentary) include a dummy variable to acknowledge the existence of free riders as identified through a OE question or alternatively, a probit model as suggested by Heckman (1976, 1979).

Based upon the sum of their analyses, the authors find considerable evidence for anchoring and scope effects in individual WTP for public risk reducing measures. They state that the size of WTP for risk reduction may depend on other activities aimed at reducing the risk level relevant for the individual. If these other risk-averting behaviors are possible, then WTP does not generally need to increase as the offered risk reduction increases.

Their conclusions (and the findings of Smith and Desvouges, 1987) highlight three key insights for this study. First, when considering respondents that are valuing both the maintenance of current levels of flood risk reduction *and* the related ecological improvements, it is essential to test for embedding bias (Path C survey respondents only, see survey in appendix). Second, as it is generally prohibitively costly to relocate households, it can be assumed that respondents in this study do not have other risk-averting alternatives at their disposal. Third, the treatment of protest bids is a crucial consideration.

As the above discussion highlights, WTP estimates for risk reduction, particularly flood risk reduction, have not consistently yielded the a priori positive relationship as conventional utility theory would indicate. Thunberg (1988), Smith (1985), Smith and

Desvouges (1987) and Romer, Pommerehene and Feld (1998) find a negative relation. The authors attribute this to a number of potential causes, which when considered, do not conflict with traditional utility maximization theory: strategic behavior bias, self insurance by locational choice, embedding effects, both in scope and anchoring effects, and alternative individual-specific risk averting behaviors. These results have numerous implications for this research as elucidated above. Furthermore, the research of Shabman and Stephenson (1996) compares benefit estimates for flood risk reduction resulting from three techniques and their findings reaffirm this study's choice of a WTP CVM survey. Their work also reveals that it is critically important to include attitudinal and behavioral variables in this study's benefit estimations. Finally, Johnston, Swallow, and Weaver (1999) highlighted the importance of the survey's payment vehicle considering respondents' assumption of government inefficiencies. This implies that a variable measuring a respondent's relative trust/mistrust of government or payment vehicle bias should be investigated.

2. Dichotomous Choice vs. Open Ended Elicitation

Next, the studies addressing OE and DC elicitation methods will be addressed. Recently, there has been extensive debate and much written scholarship concerning the choice of elicitation format for WTP CVM surveys. The issue is a critical one, both practically in the construction and implementation of the survey itself and empirically in the development of the appropriate mathematical model and choice of the estimation technique. While Mitchell and Carson (1989) list nine different variations of elicitation methods, these can be generalized into four basic elicitation methods found in the literature. These are:

1) the single dichotomous choice (DC) format where a respondent is asked whether they are willing or not to pay a specific amount[26];
2) the 'bounded' DC format which uses follow-up WTP questions based upon the respondent's earlier answers[27];
3) the single open-ended (OE) question where a respondent is simply asked his/her maximum willingness to pay for the good in question without being offered any initial value; and
4) the iterative bidding (IB) format where respondents are presented with an initial or starting point bid and then are bid upwards until a maximum willingness to pay is reached.

Huang and Smith (1998) and Bennett and Tranter (1998) both describe the most up to date empirical investigations of the OE vs. DC elicitation question and provide excellent summaries of the key issues of concern (see Tables 2-1 and 2-2 below). In order to best evaluate the four elicitation formats, each will be discussed briefly in terms of their merits, potential limitations as well as potential biases.

(a) Single DC

As discussed earlier (section III.C), the single DC format was recommended by the U.S. National Oceanic and Atmospheric Administration (NOAA) Panel in 1993 (Kopp and Pease, 1997). The panel stated that the DC format is the most realistic elicitation format when coupled with a tax payment vehicle within a referendum context. The NOAA panel also warned that the DC elicitation format could lead to an upward bias due to 'yea-saying' tendencies.[28] The propensity for DC elicitation techniques to generate WTP estimations greater than those resulting from OE elicitations has become virtual gospel in much of the relevant literature. However, this conclusion has begun to

[26] For example, "Would you be willing to pay $50 on an annual basis for the next 10 years for the good in question?"
[27] For example, a respondent is asked initially if he/she would be willing to pay $100 for good in question and the respondent answers in the negative. The respondent could next be asked if he/she would be willing to pay $50 (or $75 or $23 depending on how the follow-up questions are determined or 'bounded').
[28] People tending to answer yes to WTP amounts offered them; this is of especial concern in bounded DC and IB formats as respondents may then tend to continue to answer yes repeatedly without appropriate reflection, referred to as 'anchoring' effects.

be disputed and re-evaluated by a number of researchers. Many of these researchers question the conclusion that DC > OE (Halvorsen and Sælensminde, 1998; Huang and Smith, 1998), stating that this phenomena is due to the treatment of errors, most of which arises from model misspecifications. Others expand previous hypotheses, offering various explanations of why this phenomena may, in fact, be true (Lunander, 1998; Bohara et al., 1998; Green et al., 1998). These authors argue that DC > OE can be caused by free riding and related incentives, total project cost, perceived size of the group contributing, and/or psychometric anchoring. In addition, it has been suggested that DC formats encounter additional problems when individuals are unfamiliar with the good to be valued. Thus, respondents are said to have cognitive difficulties in accurately understanding and subsequently valuing the good (referred to also as hypothetical bias; see Blumenschien at al.1998). This is particularly problematic when there are high nonuse values to be considered by the respondent.

(b) Bounded DC Format

Essentially, this format combines DC and IB formats as it utilizes one or more follow-up WTP questions, most usually with a substantial gap between bids (Hanemann et al., 1991; Kanninen, 1993; Alberini, 1995). This main drawback of this model is that it may create a starting point or initial bid bias (i.e. the initial question or bid may bias the response to the second question). In addition, 'yea-saying' is also problematic as alluded to earlier.

(c) OE Format

This format can be especially useful in exploratory or pilot studies, where the researcher's knowledge about what the sample's respondents' range of WTP is limited.

Furthermore, empirical estimation of WTP is more straightforward than that from DC data. In addition, estimation derived from OE elicitation is less subject to the statistical biases which can potentially result from logit or probit analyses (used in DC elicitation). However, the OE format has been criticized as presenting the respondent with significant cognitive and psychological difficulties (Bennett and Tranter, 1998) due to the completely 'open-ended' nature of the elicitation (i.e. no guidance). On the other hand, Mitchell and Carson (1989) point out that when person to person interviews (phone or face-to-face) are used, allowing for clarifications (i.e. FAQs), when reference values for similar goods are offered, and when respondents are familiar with the good that they are being asked to value, the OE format is desirable[29].

(d) IB format

The similarities between bounded DC and the IB elicitation format, which offers a respondent an initial (usually randomly determined) bid and then follows with further bids designed to eventually identify the respondent's maximum WTP, are clear. The greatest advantage of the IB method is that maximum WTP is eventually determined. As occurs in the bounded DC method, the primary concern with IB is the potential for starting point bias. Table 2-2 below summarizes these merits and limitations (Bennett and Tranter, 1998) of all four elicitation types.

Table 2-2:
The Primary Merits and Limitations of Four Common WTP Elicitation Formats

Format	Merits	Limitations

[29] Phone interviews, FAQs, benchmarking values of similar goods are all used in this study's survey. In addition, due to the heightened media attention surrounding flooding and the intensity of the public's awareness surrounding watershed management issues, most respondents are expected to have some familiarity with the good to be valued.

Open Ended	• Less starting point bias • Straightforward estimation • Good for small samples/pilots	• Not suited to mail surveys • Cognitive difficulty (particularly non-use values) • High non-response • Maximum WTP not obtained • Lack of respondent faith in their bids • Possible bias from strategic bidding?
Dichotomous Choice	• Mirrors market decisions • Suited to mail surveys • Easy probit/logit analysis • Questions more easily answered	• Possible bias from statistical analysis • Difficult bid level choice • 'Yea-saying' tendency • Needs large sample?
Bounded DC	• Mirrors market decisions • Suited to mail/telephone surveys • Questions more easily answered • Statistically more efficient • Bid choice less critical	• 'Yea-saying' tendency • More complex estimation • Possible bias from statistical analysis • Possible bias from first bid response
Iterative Bidding (IB)	• Straightforward estimation • Maximum WTP obtained	• Starting point bias • Not suited to mail surveys

Source: Bennett and Tranter (1998)

Due to the apparent "widespread belief" (Huang and Smith, 1998) that DC WTP estimates > OE WTP estimates, it is essential to more closely examine the recent empirical investigations of the relative impacts of both DC and OE elicitation methods. A myriad of studies have compared WTP from these two types of elicitation methods and nearly all find that DC leads to systematically higher estimates of WTP than does OE (Brown et al., 1996; Boyle et al., 1996; Ready, Buzby, and Hu, 1996; Kristöm, 1993; Reaves, Kramer, and Holmes, 1997; Kramer and Mercer, 1997; Balistreri et al., 1996; Holmes and Kramer, 1995; McFadden, 1994; Kealy and Turner, 1993; Johnson, Shechter, 1991; Seller, Stoll, and Chavas, 1985). The scholarship indicates primarily two categories of explanation: 1) differences in how the respondents perceive and therefore respond to the various valuation formats (i.e. strategic behavior bias, anchoring, yea-saying, etc.) and 2) differences in the statistical efficiency and robustness of the WTP estimates. Additional explanations as to why DC CVM studies lead to higher WTP

estimates are offered by Magat, Viscusi, and Huber (1988). They suggest that differences in WTP estimates may also be due to either

a) the respondents maximizing the difference between a price and their reservation price in the common act of purchasing. Thus, respondents may anchor their stated WTP to some perceived cost and consequently underestimate their true WTP for the good; and/or respondents' tendencies
b) to be cautious, WTP for the good will approach their reservation price from below.

Table 2 below summarizes the findings of a number of studies that specifically compared DC and OE elicitation formats.

Table 2-3: **A Summary of DC and OE Comparison Studies**

Authors	DC/OE WTP	Results Based on Test	Model Assumption	Error Assumption	Commodity
Reaves, Kramer, and Holmes (1997)	1.29(PD)* 1.22(PI)	Yes, with no significant difference regardless of treatment of protest responses	OE results are simple average values; DC are for double bounded DC with logit models a function proposed cost. Asked of different respondents	Parametric	Maintain red woodpecker by restoring the Francis Marion National Forest
Kramer and Mercer (1997)	.68	Yes, with no significant difference	OE and double-bounded DC asked of different respondents. Both models use log of cost	Parametric	Tropical rain-forest protection thru a "donation" to the UN
Balistreri et al. (1996)	1.23-1.26 (DC/OE) 1.54-1.58 (DC/Auction)	Yes, but with a significant difference that appears robust given model specification	OE DC, English auction (actual experiment) asked of different respondents	Parametric (logit and normal)	Insurance against financial loss in context of lab experiment
Boyle et al. (1996)	1.2-1.4	Yes, with a significant difference for one of four samples	OE and DC asked of different respondents	Parametric	Moose hunting Prevention of oil spills
Ready, Buzby, and Hu (1996)	3.6-4.4 (DC vs. payment card)	Yes	OE and DC asked of different respondents	Parametric	Grapefruit with different risk due to pesticides
Holmes and Kramer (1995)	9.4 (DC vs. payment card)	Yes, with a significant difference	Payment card and DC asked of different respondents	Normal	Protect spruce-fir forests in South Appalachia
McFadden (1994)	DC > OE (order of magnitude) 5.0-10.6(Medians)	Yes, with a significant difference	OE and DC asked of different respondents	Nonparametric and parametric	Preserving wilderness

Kealy and Turner (1993)	1.03-1.1 1.4-2.6	Yes, with a signif-icant difference	Both questions asked of each respondent	Parametric (normal)	candy bar Reduction in acid rain damage in Adriondack
Kriström (1993)	DC > OE, mean not computed	Test of difference in distributions	A subset of sample re-ceived both OE and DC	Nonparametric	Preserving forest in Sweden
Johnson, Bregenzer and Shelby (1990)	1.6	No	OE and DC asked of different respondents	Normal and logit	Access to white-water river recreation

*PD refers to protest observations deleted, PI to protest observations included Source: Huang and Smith (1998)

Despite the many potential behavioral causes for the consistently higher DC estimations of WTP, two recent papers find that these causes are secondary to statistical or empirical issues. Both Huang and Smith (1998) and Halvorsen and Sælensminde (1998) in their respective analyses of the aforementioned studies' conclusions and further supported by their own analyses, determine that the deviations between DC and OE estimations of WTP were due primarily to statistical shortcomings, which had critical impacts on model variance. Specifically, even very small changes in the assumptions about the distribution and functional form of the random utility function[30] dramatically influenced the size of the proportionate error in estimates of the conditional mean or median WTP in both studies. Distributional assumptions are necessary to obtain an estimate of WTP. However, these assumptions about the random utility function "do not give a very good description of the data" (Halvorsen and Sælensminde, 1998). As a result, Huang and Smith (1998) demonstrate that the range of errors from their various model specifications can "easily span most of the reported differences between DC and OE results." Halvorsen and Sælensminde (1998) also demonstrate that different

[30] Basic economic theory indicates that an individual will maximize his/her utility subject to a budget constraint. When attempting to value nonmarket goods, the utility function is assumed to be comprised of a vector of market goods and a vector of the nonmarketed good in question. Furthermore, the error term of this utility function is assumed to be random, i.e. not correlated with any of the explanatory variables or homoscedastic. For CVM studies, violations of the assumption of homoscedastic variance may yield

assumptions about the random utility function's distribution and functional form leads to instability in several estimations (log-normal, log-logistic, log-linear, and multiplicative functions). In summary, although many studies indicate that respondent behaviors lead to biases which then result in DC WTP estimations that are consistently higher than OE, these two papers counter these conclusions and re-emphasize the importance and far-reaching impact of robust model specification.

---

substantially biased WTP estimated. Specifically, if the distribution of the utility function is allowed to be heteroscedastic with respect to certain variables, WTP estimates will fall or increase dramatically.

I. Survey Formation and Implications

A. Focus Groups

This study will use a CV WTP survey conducted in November, 1999 – May, 2000 to estimate the WTP for the maintenance of status quo flooding risk levels and/or corresponding ecological improvements to the watersheds for residents of two metropolitan Milwaukee watersheds, the Menomonee River and Oak Creek watersheds. The survey was developed over approximately a year and a half. During this development period, eight focus groups were conducted by the University of Wisconsin Survey Center in order to explore residents' feelings and thoughts about local flooding and the ecological quality of the rivers. Mitchell and Carson (1989) strongly advocate the use of focus groups and pretesting in order to best structure WTP questions. The primary objective of the focus groups was to uncover the best way to ask the highly technical WTP questions. Eight specific objectives were articulated, in order to probe participants about their:

1) Feelings of physical and emotional connectedness to the river and creek;
2) Perceptions about the health of the river and creek;
3) Perceptions about flooding;
4) Understanding of WTP questions;
5) Salient beliefs about the river and creek related issues;
6) Perceptions about their capacity to get information about this topic from the mass media;
7) Beliefs about nature and quality of this information about this topic from the mass media;
8) Value placed on the prevention of flooding and ecological improvement relative to other community issues.

Furthermore, the focus groups provided the research team with invaluable information about the participants' understanding of the contingent valuation scenario

which was developed in order to elicit the WTP response. Mitchell and Carson (1989) state that in order to obtain data which can be used to estimate a benefit equation, CV scenarios must define and communicate to respondents the following:

1) *The reference level of utility.* Usually this is done by defining the level of income available to the respondent and by describing the property rights situation with respect to the particular public good.
    a) *The property right.* In the case of collectively held goods, respondents should understand that they are currently paying for a given level of supply. The scenario should clearly indicate whether the levels valued are improvements over the status quo, or potential declines in the absence of sufficient payments.
    b) *Current disposable income.* Respondents should take taxes and fixed long-term obligations into account in giving the WTP amount for the good. If the household is the unit of analysis [as is the case in this study], the reference income should be the household's rather than the respondent's income.
2) *The nature of the public good.* Unlike ordinary survey questions, the nature of the good and the changes to be valued must be specified in detail in a CV survey. It is essential to make sure that respondents do not inadvertently assume that one or more related improvements are included in the change that they are valuing when this is not the case. Conversely, if the intent is to value several public goods, this should be made clear to the respondents.
3) *The relevant prices of other goods.* Where the change (or a lack thereof) in the public good will significantly affect the prices of other goods, the impact of this change should be communicated to the respondents.
4) *Conditions for the provision of the good and payments for it.* When and for how long the good is going to be provided should clearly be specified if this is not obvious. Respondents should understand the frequency of payments required and whether or not the payments will be required over a long period of time in order to maintain the quantity or quality change. They should also understand who will have access to the good and who else will pay for it, if it is provided.
5) *The nature of the WTP amount desired.* The scenario should be designed to ensure that the respondents express their consumer's surplus for the good and not some other type of value such as a "fair price."

The focus group enabled further improvements to the survey structure and question formats in pursuit of the aforementioned ends. In particular, the following modifications

or improvements were made to the survey's WTP question formats, as indicated by the focus group results[31]:

- The participants from the Oak Creek area were routinely unable to provide a WTP amount for the reduction of flood risk in their area because they did not believe that river flooding was a problem in their community. Oak Creek residents were much more concerned about the overall ecological health of the river. Therefore, the final survey was constructed with three separate question paths:
  Path A: "Flood Path," Menomonee River residents only, asked about WTP for flood risk only
  Path B: "Environment Path," Menomonee River and Oak Creek residents, asked about WTP for improvements to the ecological health of the river only
  Path C: "Combined Path," Menomonee River residents only, asked about WTP for both flood risk and ecological improvements.
- The verbiage setting up the CV scenario, describing the current flooding and/or ecological problem, and describing the proposed project was reworked to be more accessible to the layperson (please see the included survey, pp. 18-28). It was also clarified that the unit under consideration is the entire household and not the individual.
- Participants frequently requested details and clarifications about the proposed project itself (i.e. how long to be completed? how much is the total cost? would this provide a long-term or short-term solution?). Therefore, uniform responses to these frequently asked questions (FAQ) were generated and the FAQ sheet subsequently made available to the interviewers.
- Participants also asked about the payment vehicle and how the cost of the project would be distributed. Thus, in the final survey the phrase was used, "keep in mind that the expense of this project would be spread out over among all the residents of the Menomonee River/Oak Creek watershed for the next 20 years."
- Additionally, participants indicated that the form of taxation was a sensitive issue given that the tax burden in the areas of being surveyed is high relative to other areas. Thus, in the final survey the description of the type of taxation was deliberately vague indicating only that the expense would be borne by others also, but avoiding any specific indication of how that tax would be collected.
- In addition to describing the problems and proposed project, participants were also provided with several "benchmarks." These benchmarks described other public goods and services that the respondents were currently funding and the amount of average annual household contributions to these goods. Participants indicated that these reference values were helpful, but were either too controversial or difficult to relate to (i.e. a highly controversial local tax being used to fund a new professional baseball stadium and the NASA space program). Therefore, in the final survey benchmarks were used that were non-controversial and were specific to the four-county Southeastern Wisconsin area in which the survey was conducted. Annual household payments for the following goods were included as benchmarks in the final

[31] The survey and FAQ page are included in the appendix.

survey (rounded to the nearest dollar): ambulance service $8; parks and recreation $54; highway construction, maintenance, and administration $95; law enforcement and fire protection $203; public education $1500.

- Respondents across the board had a difficult time understanding the description of the flood risk and the potential for the flood plain to expand. The focus groups explored three alternative descriptions of the flood risk levels and based upon the responses, the final risk description combined two of the three risk descriptions and is as follows:

  Scientists estimate that if current development trends continue, the risk of flooding will be roughly 3 to 5 times higher than it currently is. That is, the risk of flooding in any given year will increase from 1% to 5% for most households in the flood plain. However, in some parts of the flood plain, such as Piggsville and the eastern part of the village of Wauwatosa, there will be an even greater risk of flooding. In addition, up to 30 home and other buildings that are not at risk now will be at risk of flooding as the flood plain expands.
- Given the frequency of recent flooding events in Southeastern Wisconsin, participants, as expected, did not have difficulty providing a WTP to open-ended questions. Nor did respondents seem to have a problem understanding the ecological good or giving a WTP for that good. However, given that most respondents may have minimal experience with the ecological good in question and that the survey will need to be divided into three sub-paths, dichotomous choice questions may be needed in the final survey.
- In order to identify protest bids during later analysis, the final survey also asked a follow-up question of respondents indicating a $0 WTP, asking them to identify what their main reasons were for their zero bid.

B. Pretesting

After the revisions to the WTP questions and the survey in general, as guided by the focus groups, the survey was then pretested by the University of Wisconsin Survey Center. These pretests were used to approximate the time needed for each survey, identify further issues or clarifications needed for the FAQ form, and to finalize the elicitation format to be used. Few modifications were made to the WTP questions following the alterations motivated by the focus groups. As guided by previous research into elicitation formats, particularly Mitchell and Carson (1989), a bounded iterative bid format concluding with a final open-ended maximum willingness to pay question was used in this survey. The combination of three basic elicitation techniques (iterative

bidding, bounded dichotomous choice, and open-ended) was chosen for the five following reasons:

1) to most closely mirror market decisions (this proposed plan would most likely come before respondents in the form of a referendum);
2) minimize cognitive difficulties or, in other words, to maximize the ease of a respondents' complete and accurate understanding of the hypothetical scenario;
3) minimize non-response (i.e. uncertain, unsure, don't know and logically inconsistent responses) and protest bids (given by respondent who is not truly willing to forgo the good, but rather is objecting to the evaluation process itself);
4) minimize strategic overbidding (where the given WTP differs from true WTP in an attempt to influence the provision of the good or the respondent's level of payment for the good); and
5) to obtain a maximum WTP, which was required for the testing of psychological models by the risk communication members of the research team.

C. Implications of Survey Format, CV Scenario, and WTP Elicitation

The nature of the good being valued in this study, the format and method used in the survey, the description and phrasing of the CV scenario, and the WTP elicitation method all can potentially contribute to systematic error or biases. As Mitchell and Carson (1989) point out, susceptibility to particular biases will vary greatly depending upon the "nature of the amenity being valued." Consequently, it is possible to evaluate what biases are most likely to occur in this study, their potential empirical impacts, and to then critique whether appropriate measures were taken to minimize their occurrence and effects.

Mitchell and Carson point out that it is first essential to enhance reliability as much as possible and therefore reduce variance as much as possible by designing the CV scenario so that respondents find it both realistic and credible. To this end, they suggest and advocate strongly using extensive focus groups and pretesting, two methods used in this study. They also suggest that interviewers avoid placing pressure on respondents and

offer the respondent more opportunity to reflect on the topic before valuing the good[32]. Again, by using a well accredited survey lab and by positioning the WTP questions later in the survey, this study followed these recommendations to increase reliability.

The authors also point out that the very elements that enhance realism may increase bias. They term this the "realism-bias tradeoff" (228). Therefore, Mitchell and Carson point out that it is meaningless to speak about "hypothetical bias" in CV surveys, as there is no unique bias produced by the hypothetical nature of CV surveys. As aforementioned, the good to be provided, the type of market, the payment vehicle, and the elicitation method all can contribute to bias. Each bias considered most germane to this study is described below, including a definition, its empirical implications, approaches to minimize its effects,[33] and those steps undertaken in this study. In short, based upon the suggestions of Mitchell and Carson (1989) and other researchers, every effort was made in the creation of the survey as a whole, with particular attention afforded to the WTP questions, during the focus groups, pretesting, and resultant revisions, in order to avoid as much as possible any biases.

1. *Incentives to missrepresent responses*: biases in this class occur when a respondent misrepresents his/her true WTP.
   a) *Strategic bias*: respondent gives a WTP amount that differs from true WTP in an attempt to influence the provision of the good and/or respondent's level of payment for the good. Mitchell and Carson point out that strategic bias is only problematic under certain "worst-case" conditions and that such conditions are very rare in most CV studies. The authors suggest staying away from the use of a highly abstract or implausible payment vehicle in order to avoid this bias. As it is uncertain if the respondent is strategically over or underbidding it is difficult to detect and/or correct for this bias. However, as guided by the focus groups and pretesting, this study carefully crafted the CV scenario such that the payment vehicle, timeframe, and distribution of the costs were quite sufficiently clear and as realistic as possible.

[32] The severe impacts of not allowing respondents sufficient time to reflect on the topic are discussed in sections IV.C.4., V., and VI.

[33] A full typology, summarized in table format, is available in Mitchell and Carson (1989: 236-237, 263).

2. *Implied value cues*: these biases occur when elements of the contingent market are treated by respondents as providing information about the "correct" value for the good.

   a) *Starting point bias*: where the elicitation method or payment vehicle directly or indirectly introduces a potential WTP amount that influences the WTP amount given by a respondent (may be accentuated by a tendency to yea-saying). This type of bias is most prevalent when dichotomous choice or iterative bidding formats are used since they directly offer the respondent a proposed amount that the respondent must then accept or reject. In an iterative bid format, even when a respondent rejects an initial bid, starting points well above the respondent's true WTP will tend to increase the revealed WTP and vice versa. Mitchell and Carson highlight that no "valid method exists to adjust the findings obtained by the bidding game to compensate for the effects of the starting point" (241). The authors therefore do not conclude that the iterative bid format is the preferred elicitation format. This study in response to these concerns begins its WTP questions with a simple DC question asking respondents if they are willing to pay *anything* for the proposed project. Furthermore, this study also uses an open-ended question in addition to iterative questions in order to specify a respondent's maximum WTP[34].

   c) *Relational bias*: where the description of the good presents information about its relationship to other public or private commodities that influences a respondent's WTP amount. Mitchell and Carson warn that the use of benchmarks may induce relational bias. The authors acknowledge that benchmarks can be used to remind respondents of their budget constraint, and they suggest that any benchmarks should be not directly related to the amenity in question. In studies about environmental benefits they suggest benchmarks about police and fire protection, the space program, education, and roads and highways. This study utilized only benchmarks completely unrelated to the good being valued.

3. *Scenario misspecification*: biases in this category occur when a respondent does not respond to the correct contingent scenario (the *intended* scenario is correct and the errors occur because the respondent does not understand the scenario as the researcher intended it).

   a) *Symbolic bias*: respondent values a symbolic entity instead of the researcher's intended good. In order to avoid this bias, the interviewing procedure and the wording of the scenario must overcome the respondent's natural inclination to think of the CV survey as an ordinary public opinion survey. When respondents respond to the symbol instead of the amenity in question, revealed WTP for the improvement of local air quality would be virtually identical to revealed WTP for the improvement of national or global air quality, for example. In addition, topics that create powerful emotional feelings may make it difficult for respondents to focus on the valuation-relevant aspects of the scenario (i.e. nuclear waste). These

[34] This will be statistically tested for by including the value of the initial bid given to the respondent in the estimations. If this variable is significant it implies that the final WTP was positively or negatively related to the starting point of the DC modeling process.

topics are not appropriate for a CV survey. Again, the meticulously crafted CV scenario and the type of amenity being valued should not induce this bias.
b) *Part-Whole bias*: respondent values a larger or smaller entity than the researcher's intended good (is similar to symbolic bias and related to aggregation issues). This bias could affect a respondent's valuation of the amenity in terms of its geography or the policy package of which it is a part. For example, a respondent when asked his/her WTP for water quality improvements in a local river basin could be unable to isolate that river basin in his/her mind. The respondent may then value a larger range of waters than intended by the researcher. Similarly, a respondent may treat a proposed policy as symbolic of a larger policy package and assign to the proposed policy the value that they have for the entire policy package. In this study, the saliency of the amenity to be valued is high and the description very detailed, which should avoid this bias.
c) *Payment vehicle bias*: payment vehicle is either misperceived or itself not valued in a way intended by the researcher. In instances where this bias occurs, the WTP is influenced by the type of payment vehicle. However, the authors point out that no misspecification exists, nor bias, if the researcher understands the relevant effects of the payment vehicles on WTP and knowingly accepts them, as is the case in this study. This study only generally defined the payment vehicle so as to minimize any potential bias.
d) *Question order bias*: where a sequence of questions, which should not have an effect, does have an effect on a respondent's WTP amount [this would only be of concern for the Combined Path]. Mitchell and Carson indicate that when there is a natural sequence to the provision of the goods, this should prescribe the order of the questions. This is the case for this study as the impact of a flood control project would be more immediate than the effects of a plan for the ecological improvement of the river, which would most likely be gradually felt over an extended period of time. In order to overcome this potential bias, respondents in Path C were asked to value the flood risk maintenance and ecological improvement goods jointly, as a policy package.

4. *Sample design and execution biases*
Mitchell and Carson highlight a few key points about nonresponse problems and biases. First, some level of nonresponse to WTP questions is inevitable. Second, given the choice between a respondent offering an unconsidered guess at WTP or having them say they don't know how much the good is worth to them, the latter behavior is preferable. Third, nonresponse rates in CV surveys as high as 20 to 30 percent is not uncommon for WTP questions in certain circumstances. Fourth, item nonresponses on WTP questions fall into four categories: don't know, refusals, protest zeros, and responses that fail to meet an edit for logical consistency[35], with protest bids posing the greatest problem. Fifth, in order to identify protest bids, the technique of asking respondents who gave zero bids why they did so is indispensable (this is done in this study).
a) *Sample nonresponse bias*: where the sample statistics calculated by using those elements from which a valid WTP response was obtained differ significantly from

[35] See section IV.C.3.

the population parameters on any observed characteristic related to WTP. In other words, this occurs when different categories of respondents, such as those with greater or lesser education or incomes, tend to have different nonresponse rates, and these categories of people also tend to hold different values for many public goods. It is possible to use weighting to correct for item nonresponse, however the preferred techniques use the available information from other questions answered by the respondents to impute the missing values (Kalton, 1983; David et al., 1986)[36]. Mitchell and Carson recommend the use of imputation, but it is important to note that imputation reduces bias, but not variance, as imputed values should not be treated as real values for the purpose of calculating sample size[35].

5. *Inference biases*
   a) *Multiple public goods sequence aggregation bias*: where the WTP amounts for public goods that are substitutes or complements are added together to value a policy package containing those amenities, despite the fact that the amenities were valued in an order (for example, independently) different from the appropriate sequence [again, only an issue for the Combined Path]. In order to avoid this type of bias, this study utilized three different question paths and the WTP amount for Path C respondents elicited as a policy package[37].

D. Finalized Survey

The finalized survey was in the field from November 1999 to May 2000. Approximately 1000 households participated in telephone interviews administered by the University of Wisconsin Survey Lab in the Menomonee River and Oak Creek watersheds. As aforementioned, the survey was administered using three distinct question paths, the Flood Path (Path A), the Environment Path (Path B), and the Combined Path (Path C)[38]. Of primary interest to this study is the Flood Path and the Combined Path. In addition to WTP questions, the survey also included questions that address the respondents' attitudes, beliefs, values, and motivations. Other questions

[36] For a discussion of other correction methods, see Mitchell and Carson (1989: 272-276).
[37] For a more detailed description of aggregation issues, see Mitchell and Carson (1989: 41-48, 286-287), Hoehn and Randall (1982), Randall et al. (1985) and Tolley et al. (1983, 1985).
[38] Thus, it is possible to test for potential embedding effects by combining the samples from paths A and C, including a dummy variable for path C, and investigating this variable's significance and sign. It is also possible to compare the two paths' mean WTP. However, both of these techniques fail to satisfy the *ceteris paribus* condition.

explored respondents' behavioral characteristics and ethics. Lastly, basic demographic information about the respondent was collected in the survey. Of primary importance to this study are the WTP responses and demographic information. To a more limited extent, attitudinal and behavioral information is also useful (see section IV below).

II. Model Specification

A. Theoretical Model

Basic utility theory is used in this study to guide the development of a theoretical model to describe a household's willingness to pay for the maintenance of current levels of flood control. The theoretical model can then be used to aid in the identification of variables for empirical analysis and in the formation of hypotheses. Following Freeman (1979), an individual (or household in this case) will maximize utility subject to a budget constraint. In this study, the household's utility can be described by a vector of market goods, X, and a nonmarket good, Z. From Samuelson (1954), the value of the nonmarketed good, which is not priced and can only be provided in a fixed amount, is given by the household's WTP to receive the nonmarketed good or the consumer surplus.

$$\text{Maximize } U(X,Z) \qquad \text{Subject to: } \sum_i P_i , X_i < Y \qquad [4\text{-}1]$$

where Y is income and P is a vector of prices for the marketed good. Solving generates a demand function:

$$X_i = X_i (P,Z,Y) \qquad [4\text{-}2]$$

From equation 4-2 it is possible to see that the level of a nonmarketed good can enter as an argument in the demand for a marketed good. However, because the nonmarketed good is not priced it is not possible to similarly derive a demand function for the nonmarketed good from the utility maximization system. Minimizing expenditures conditional on utility, the dual of the individual's utility maximization, can be stated as follows:

$$\text{Minimize} \sum_i P_{i\,,}\, X_i = M \qquad \text{Subject to: } U(X,Z) = U^* \qquad [4\text{-}3]$$

where U^* is a reference level of utility and M is the minimum money expenditure required to attain U^*. By solving 4-3, the household's expenditure function results:

$$E = E\,(P, Z, U^*) \qquad [4\text{-}4]$$

To generate the Hicks-compensated demand function for the nonmarketed good, equation 4-4 can be differentiated with respect to a given price

$$\chi^* = E_\chi\,(P,Z,\, U^*) \qquad [4\text{-}5]$$

Both equivalent variation and compensating variation surplus measures can be evaluated through equation 4-5 depending on the reference level of utility (Thunberg, 1988). To generate the Hicks-compensated inverse demand function for the nonmarketed good, equation 4-4 is differentiated with respect to Z. According to Mishan (1976), this is the theoretically appropriate surplus measure for welfare comparisons.

$$m^* = E_Z\,(P, Z, U^*) \qquad [4\text{-}6]$$

In terms of money income transfer required to maintain the household's utility at U^*, equation 4-6 gives the marginal WTP for a change in the level of Z. The benefit to the individual for a change in Z is therefore given by

$$WTP = \int_{Z_0}^{Z_1} E_Z\,(P, Z, U^*)\, dZ \qquad [4\text{-}7]$$

Any proposed flood project, provided at a zero price, can only provide a fixed flow of flood hazard mitigation services. Although the proposed project in this study would maintain the current level of flood risk, it is still impossible to know what the exact resultant level of flood protection services provided by such a project for each household would be. The valuation of the flood control project proposed in this study depends upon each household's subjective assessment of the flooding risk, which in turn depends on

proximity to the river, location in the flood plain, and other factors, *and* the extent to which the proposed project will impact that risk. The question is therefore, what is the appropriate theoretical model of decision making under uncertainty?

The expected utility model, developed by Von Neumann and Morgenstern (1947), addressed this very question. The authors argued that the rational decision maker, when faced with choices with uncertain outcomes, will maximize expected utility of consumption. Many risk reduction WTP studies use this expected utility model (Brookshire et al., 1985; Smith and Desvouges, 1987). However, Carson and Mitchell (1989) as well as other researchers[39] indicate that the EU model's hypothesis are not appropriate for the analysis of decision making involving losses rather than gains. Therefore, the EU model was not utilized in this study to describe WTP for the proposed flood control project. Similar to Thunberg (1988:12), this study

> Assumes that flood plain landowners maximize the utility of the expected value of an investment in flood protection. The context of decision making under uncertainty is maintained as individuals [households] are assumed to form subjective judgements over future flood event and flood consequences.

In other words, the watershed landowner is assumed to hold subjective beliefs about the likelihood, causes, and effects of flood events in coming years. The landowner can also therefore be said to hold subjective expectations for property effects, nonproperty effects, and a subjective probability distribution for flooding. The proposed flood control project provides a fixed flow of flood protection services as aforementioned. These protective services accrue to the landowner as long as the landowner remains in the watershed. This study assumes that the landowner seeks to maximize his/her utility of the expected present value of consumption. In the absence of the proposed flood control plan the

[39] See also Schoemaker (1982), Smith and Desvouges (1985), and Kahneman and Tversky (1979).

household's utility, U^o, is conditional upon expected flooding effects, price levels, income, and general economic conditions. If the current level of flood control would degenerate over time and the household must pay out in order to maintain the current welfare position, the landowner's WTP for a river flood control project can be derived by forming a lifetime expenditure function

$$E = E\,(P, FP, U^o) \qquad [4\text{-}8]$$

Where E is the minimum level of lifetime expenditures required to attain U^o, P is a vector of prices over the time horizon for consumption goods, FP is the flow of flood protection services, and U^o is the household's current level of welfare. Differentiating with respect to FP

$$\partial E / \partial FP = E_{FP}\,(P, FP, U^o) \qquad [4\text{-}9]$$

then integrating from zero to a specific level of flood protection (FP')

$$WTP = \int_0^{FP'} E_{FP}\,(P, FP, U^o)\,dFP \qquad [4\text{-}10]$$

Equation 4-10 describes the maximum WTP that would maintain the household's current welfare position, U^o. As all the variables are considered over time, equation 4-10 measures the household's WTP as if the payment were lump sum. Consequently, the WTP decision represents an investment in a flood control project with the expectation of receiving a future flow of flood protection services in order to maintain the current level of flood risk. In other words, investment in a flood control project will be a function of the level of flood protection services provided. Following Thunberg (1988), flood levels are denoted as (i), the subjective property and nonproperty effects are therefore defined as $j(PE_i)$ and $k(NPE_i)$ respectively. The household's subjective probability distribution for varying flood levels can be defined as m(i). The expected net present value of the

property and nonproperty effects that are avoided with the proposed project is the WTP, or value of flood control investment. Therefore, the WTP can be stated as a general function of subjective probabilities, m(i), subjective property, $j(PE_i)$, and nonproperty, $k(NPE_i)$, effects, and the household's investment position, IP (i.e. budget constraint).

$$WTP = f(\, m(i), j(PE_i), k(NPE_i), IP) \qquad [4\text{-}11]$$

And more simply,

$$WTP = f(SP, SPE, SNPE, IP) \qquad [4\text{-}12]$$

where SP = subjective probabilities or the household's subjective judgement of the with and without flood control project chance of flooding, SPE = subjective property or the household's subjective evaluation of the expected protection of personal property provided by the project, SNPE = subjective nonproperty effects or the household's subjective estimation of expected community effects and other expected nonproperty related personal gains (i.e. anxiety relief, increased recreational possibilities, etc.) from the project, and the IP = investment position or a combination of household income and other factors impacting investment (discount rates, etc.).

B. Empirical Model

The generalized model specified above (4-12) guides the selection of variables useful in empirical analysis for this study. It is critical to first note that in this study the WTP for flood control is not elicited as a one-time lump sum payment, but rather as an annual payment over a twenty-year period. The WTP estimated in this study is the annual WTP function, rather than the total WTP function[40]. Therefore, only the household's income needs to be considered in lieu of an investment position. The empirical variables indicated by the generalized model (equation 4-12) can be grouped

into three categories: demographic variables (including income and community characteristics), attitudes and values, and risk variables (including a calculated flood risk measurement or FRM unique to every household).

$$WTP = f(\text{INC, some measure of flood risk, demographics, attitudes/values, and other risk attributes}) \quad [4\text{-}13]$$

Demographic variables of interest include the following: income (INC), age, educational level, gender, marital status, how long the residents of the household have been in the area, length of predicted stay, tenancy status, number of dependents in the household, race, and other community characteristics (i.e. is the community affluent, rural, urban, general racial make-up?). Attitudinal/value variables of interest include the following: political party affiliation, environmental beliefs and activities (including donations, organization memberships, etc.), anxiety about flooding, feelings towards agencies involved in flood control, trust/mistrust of government agencies, general knowledge about flooding and the environment, and feelings of connection to the community. Risk variables include the following: the calculated FRM, the household's distance to the river, the household's perceived nearness to the floodplain, and the household's past experiences with flooding.

C. Data Limitations

Households totaling 999 were surveyed in the Menomonee River and Oak Creek watersheds. Respondents were asked questions following one of three unique question paths. Of the 999 respondents, 303 followed the Flood Path (Path A), 432 followed the Environment Path (Path B), and 264 followed the Combined Path (Path C)[41]. These three

[40] Hence, total maximum WTP would be the discounted present value of this stream of payments.
[41] All Path A and C respondents were Menomonee River watershed residents, while all Path B respondents were Oak Creek watershed residents for the reasons elucidated earlier.

subsets of data were further impacted and/or reduced due to the following five factors: two unavoidable technical limitations, respondents' unwillingness to provide income information, WTP nonresponse and protest bidding, and the number of questions asked respondents. As this study is primarily concerned with WTP for flood risk reduction, only Paths A and C are of empirical interest and will be discussed further.

1. Technical Data Limitations

(a) Geocoding

First, the Path B data set was reduced due to the inability of the geographical information system being utilized in this study to geocode the addresses[42]. As it is impossible to generate any type of risk measurement without a geocoded address, the data from respondents with 'un-geocodable' addresses were removed from the data set. As aforementioned, all of the 303 Path A respondents' addresses could be geocoded and thus, no data points were lost in Path A. However, in Path C, 26 addresses could not be geocoded and thus, the sample was reduced to a total of 238. It is critical to note that unfortunately, the addresses which were not geocoded are not randomly located throughout the geographic area being investigated. Instead, the addresses are concentrated in two towns, Menomonee Falls and Germantown. Therefore, it is possible and indeed likely that the estimations resulting from Path C with suffer from spatial bias.

(b) Flood Risk Measure Calculation

The second data limitation stems from the calculation of the Flood Risk Measure (FRM). The calculation of the FRM, as developed by the engineering members of the

[42] Geocoding involves identifying the respondents' reported addresses as a specific and unique points in a coordinate space. In other words, geocoding allows each household to be precisely geographically located. The coordinate space which after geocoding includes each household's location, also includes the Menomonee River watershed, the Oak Creek watershed, major roads, highways, railroads, etc. Geocoding

research team and as defined through hydrologic modeling techniques, requires the floodplain to be digitized (i.e. located in the coordinate space as a polygon). The flood plain has been fully digitized for all large branches of the Menomonee River and Oak Creek watersheds' rivers. However, in areas of the watershed where the river is only a relatively small tributary or stream, the flood plain has not been digitized for that smaller portion of the river. Thus, it is impossible to calculate an FRM for households that lie near these smaller portions of the rivers[43]. This reduces the number of usable observations for both Paths A and C. Path A looses 145 responses and Path C looses 133 observations. This will create spatial bias in any estimation of WTP using the FRM for both Paths A and C[44].

In order to overcome this potential bias problem, a proxy variable for the FRM was also calculated. This proxy variable, DISTRVR, is the distance from the household to the nearest river, stream, or tributary. This does not exclude any data points due to the relative smallness of the river or stream in certain portions of the watersheds, as was the case with the FRM. This is because 100% of the river and its tributaries are identified in the coordinate space. Thus, for Path A and C estimations using the DISTRVR proxy variable, the sample sizes remained at 303 and 238. It is important to highlight that even when using the DISTRVR variable, Path C still has lost 28 observations, while Path A has lost zero observations.

2. Missing Income Data

---

therefore allows the spatial relationships between the respondent's household location and the watershed, floodplain, and river to be investigated.

[43] Currently, the engineering research team is working to complete the digitization of the floodplain.

[44] Ultimately, the FRM is not used in any of the final model specifications. The spatial bias resulting from the loss of significant numbers of data points rendered the estimations using the FRM with minimal explanatory power and few significant independent variables meeting theoretical expectations. Therefore, a proxy variable for flood risk, the household's distance to the river (DISTRVR), is used.

Respondents were directly asked to provide their income in the previous year. When a response was given, a continuous income variable was generated. If respondents were unwilling to provide their exact income, they were asked to provide a range (see the survey included in the appendix, questions 159 and 159a). If respondents were unwilling to provide either a specific income or a range they were considered income item nonresponsdent[45]. The vast majority of respondents (surprisingly) provided their actual incomes.

Despite that fact, it is still necessary to somehow calculate an income measure for the remaining respondents, rather than discarding these data points. It is possible for those that provided only an income range and for those who were item nonrespondent to impute an income value. The underlying assumption of this imputation process is that the distribution of the incomes reported continuously is identical to those reporting only a range or refusing to respond altogether. For every unique sample, missing income values were re-imputed using the following procedure. First, for all those respondents reporting continuous income data, that data was grouped into the correct corresponding ranges and the mean income for each range computed. Second, these calculated range means were then assigned as the income value for the respondents corresponding to that range. Third, the entire sample's mean income was calculated and assigned to all respondents who refused to provide any income information. The distributional assumption underlying this procedure assures that the imputed incomes will not distort or bias the data (however, it should be also noted that it adds no new information either).

3. WTP Nonresponse and Protest Bids

[45] Approximately less than 7% of the respondents in both Path A and C were income item nonrespondent – a very low rate of nonresponse relative to other CVM WTP studies.

(a) Nonresponse

Mitchell and Carson (1989:267) state that

> no matter what sampling plan and survey method are used in a contingent valuation survey, some level of nonresponse to the WTP questions is virtually inevitable, with the consequence that the number of those who give valid WTP amounts will be smaller than the number of originally chosen sample elements.

Item nonresponse occurs when a respondent answers some or most of the questions on a survey, but fails to answer a particular question of interest, such as the WTP question[46]. Mitchell and Carson (1989) indicate that nonresponse rates of 20 to 30 percent for the WTP questions are not uncommon where the sample includes people of all educational and age levels (i.e. is random), the scenario is complex, and the object of valuation is an amenity that people are not accustomed to valuing in dollars. Item nonresponses fall into four categories: (1) don't knows or uncertains, (2) refusals, (3) responses that fail to meet an edit for minimal consistency[47], and (4) protest bids. Following CVM literature and the accepted methodology, respondents with WTP item nonresponses falling into the first three categories must be removed from the data set. Eliminating these nonresponses reduces the number of observations for both Path A and C by eight and five respectively. Path A then has 295 observations while Path C has 233 observations.

(b) Identification of Protest Bidders

Protest bidders must next be identified. Mitchell and Carson (1989) describe protest bidders as respondents that offer a $0 bid because they reject the evaluation process itself. It is crucial to distinguish between protest bidders and those respondents who truly prefer to forgo the good in question rather than to have to pay for it. To this

[46] As alluded to previously, item nonresponse for questions which ask for the respondents are quite common, although this was not a major problem in this study.

end, Mitchell and Carson (1989), following Desvousges, Smith, and McGiveny (1983), highly recommend the technique of asking respondents indicating a $0 WTP, in an open-ended format, why they did so.

Respondents in the survey used in this study were asked the following question (either question 100 or 104), which began a series of WTP elicitation questions that ended with a final maximum WTP question.

Suppose the plan I just described were one the next ballot. Would you be willing to pay anything for this plan?

If respondents indicated that they would *not* be WTP anything, they were then asked this open-ended follow up question (q100a or q104a)

Could you briefly tell me your main reason for not wanting to provide funds for a project such as this?

Therefore, by using the narrative responses to these follow-up questions, it is possible to distinguish between protest and other $0 bids. Mitchell and Carson (1989) point out that the determination of whether or not a $0 bid is a protest bid or a legitimate valuation is ultimately a subjective one[48]. For Path A, of the 295 respondents, 15 were determined to be protest bidders (or 5.084%). Of the 150 Path A respondents with a calculated FRM, 6 were determined to be protest bidders (or 4.000%). For Path C, of the 233 respondents, 15 were determined to be protest bidders (or 6.438%). Of the 100 Path C respondents with a calculated FRM, 3 were determined to be protest bidders (or 3.000%).

The above discussion concerning the data set can be summarized in the following table:

[47] Where the responses are logically inconsistent or entirely illogical (i.e. low income respondents giving WTP amounts that represent implausible large portions of their income).

[48] These narrative responses are included in the appendix and those indicating protest bidding are in boldface.

Table 4-1: **Summary of Data Sets**

TOTAL DATA SET (paths A, B, and C) = **999**

PATH A	**PATH C**
Total respondents = **303** Of those 303, **geocoded** observations = **303**	Total respondents = **264** Of those 264, **geocoded** observations = **238**
After **removing WTP item nonresponses** falling into categories 1-3 from Paths A and C, Total sample = **295**	Total sample = **233**
Of those 295 respondents, **protest bids = 15**, or **5.084%**	Of those 233 respondents, **protest bids = 15**, or **6.438%**
OR	
Of those 295 respondents, those **with FRMs = 150** Of those 150 respondents, **protest bids = 6**, or **4.000%**	Of those 233 respondents, those **with FRMs = 100** Of those 100 respondents, **protest bids = 3**, or **3.000%**

(c) Treatment of Protest Bidders

Next, the treatment of protest bidders must be decided upon. This is a twofold question. First, it must be determined whether or not protest bidders will be left in the sample when estimating the WTP function. Second and furthermore, it must be determined if this makes any appreciable difference in the estimation (aside from questions of bias). If protest bids are left in the sample, a dummy variable is then generated (where if the $0 WTP bid is a protest bid, the dummy variable, PROTEST in this study, is set equal to 1). When this dummy variable is included as a regressor in the estimations, it would be expected to have a negative (and most likely significant) coefficient. Romer, Pommerehne, and Feld (1998) argue that simply discarding protest

bidders is a “rudimentary” approach and strongly advocate the use of dummy variables instead. Mitchell and Carson (1989:267-269) seem to concur.

When considering protest bid treatments for this study, it is important to note, as the above table and discussion illustrate, that the percentage of protest bidders for any of the samples considered is quite small (3-6.4%). Notwithstanding these small percentages, when the final models were specified and estimated, they were estimated both including and excluding the protest bidders. When protest bids were excluded there was no significant changes in the magnitudes, signs, or significance of the other explanatory variables. Therefore, the estimation results reported below include the PROTEST dummy variable.

(d) Implications of Excluding Protest Bidders

Second, although in the final estimated models the protest bidder respondents were included in the sample, it is necessary to discuss further the potential empirical impacts of excluding protest bidders. When protest bidders are excluded from the sample and are therefore unobservable, the sample becomes nonrandom, drawn from a subpopulation of a wider population. This type of sample is referred to as a censored sample (Maddala, 1983; Greene, 1999). The implication is that the error term is correlated with the explanatory variables, creating what is called sample selection bias. In order to overcome this problem, a two-step estimation procedure, known as the Heckman two-step or ‘Heckit,’ can be utilized (Heckman, 1976; Kennedy, 1998). Very basically, a probit model is used in the first step in order to predict whether or not a respondent will reply to the WTP question with a protest bid. This information, in

particular the expected value of the error term, is then incorporated into the OLS estimation[49].

Usage of the Heckman estimation has two important caveats. First, as Greene (1981, 1983, 1995) and Kennedy (1998) both point out, while the Heckit procedure does generate consistent estimates, solving the selectivity bias problem, it is inefficient. It introduces a measurement error problem in the second step, since an estimate of the expected value of the error term is employed. Second, Kennedy (1998:256), states that for small samples "it appears that the Heckman procedure can often do more harm than good and that…OLS is surprisingly efficient." Stolzenberg and Relles (1990), Hartman (1991), Zuehlke and Zeman (1991), and Nawata (1993) in their various Monte Carlo studies, find that relative to subsample OLS the Heckit procedure does not perform well when the sample size is small, the amount of censoring is small, the correlation between the errors of the regression and the selection equations is small, and the degree of collinearity between the explanatory variables in the regression and the selection equations is high. All of these criteria apply to the regressions in this study. Due to this and the inefficiency of the two-step Heckman estimation method, it is not utilized in this study.

4. Number of Questions Asked of Respondents

Respondents in Path A were asked a total of over 145 questions, including various demographic questions, the WTP elicitations, and approximately 80 questions directed at attitudinal, behavioral, and value characteristics. From these 145+ questions, approximately 75 potential explanatory variables were extracted. The survey took approximately 30 to 45 minutes to administer. In addition, before the WTP questions

[49] See Greene (1995, 1981, 1983) and Heckman (1976, 1979) for a detailed description of the estimation procedure.

were iterated, the respondents had already been asked approximately 85 questions and had been interviewed for approximately 25 minutes. In sharp contrast, respondents in Path C were asked only a total of approximately 30 questions, including *only* very basic demographic information and the WTP elicitations. From these 30 questions only approximately 15 potential explanatory variables could be extracted. The survey took approximately 8-10 minutes to administer. Furthermore, the respondents had been asked a total of *only three* questions before the WTP questions were iterated.

The implications of this final data limitation are grave. First, with no potential attitudinal, behavioral, or value explanatory variables, the estimation of a Path C WTP function is seriously impeded. Second and just as detrimental, the respondents are given literally no time at all to contemplate and reflect on the watershed issues they are being interviewed about before the WTP questions are posed. Put simply, unlike Path A respondents who have about 25 minutes to think – and be probed – about the issue at hand before being asked their WTP, Path C respondents have to virtually pull an amount out of thin air. Mitchell and Carson (1989:97) write

> The CV researcher's objective is to obtain the respondents' consumer surplus for the amenity – the maximum amount the good is worth to the respondent before he would prefer to go without it. It might be thought that the best way to do this would be to ask the respondent what maximum price he would be willing to pay for the described good, and record the answer. Unfortunately, respondents often find it difficult to pick a value out of the air, as it were, without some form of assistance, just as they are hard pressed to determine the highest price they are willing to pay for items at a garage sale in the absence of price tags.

This implies that Path C respondents faced great cognitive difficulties in providing their WTP, particularly considering that these respondents were asked to value a policy program *package*, including both maintaining current flooding rates (a nonmarket good they may have some initial understanding of) *and* improving the ecology of the

watershed (an extremely nebulous nonmarket good). This draws into question the reliability of their WTP responses. Later, citing the work of cognitive psychologists, Mitchell and Carson (1989:115) note

> 'The behavior itself of answering questions may act as a sort of primer which makes some cognitions more *accessible* or salient than others.' In the case of the anchoring heuristic [or starting point bias], another tendency identified by cognitive psychologists, people make estimates by starting from an initial value, which may be suggested by the formulation of the problem, and then adjust that value to yield the final answer[50].

Still later in the text Mitchell and Carson (219) also state, "Another reliability-enhancing device is to offer the respondent more opportunity to think about the topic before posing the valuation questions." This indicates that unlike Path C, the Path A respondents, being given much more time to contemplate the amenity to be valued, may not suffer from such grave cognitive difficulties. Therefore, the estimations of Path A WTP would not be expected to suffer from starting point bias and would also be expected to be reliable.

D. Variable Definitions and Functional Forms

1. Variable Definitions

Variables included in the empirical models were obtained from three sources. The vast majority of the variables were gleaned from the responses to survey questions (the survey is replicated in its entirety in the appendix). The community characteristics were obtained through a database program, CensusCounts®, which groups and organizes data from the 1990 U.S. Census.[51] Finally, the engineering members of the research team calculated variables measuring flood risk. Each variable used in the final estimations is described below and its *a priori* sign given, if known.

---

[50] The forthcoming discussion will demonstrate that aside from the spatial bias problem initially faced with the Path C data set, this additional cognitive hurdle the respondents faced did indeed lead to embedding and starting point bias, rendering Path C estimations void of virtually all explanatory power.

(a) Demographic Variables

INC – The effect of income on WTP is expected to be positive and is measured as the household's 1998 or 1999[52] income from all sources (survey question 159). The respondents were specifically asked

And, just roughly, what was your total household income from all sources, before taxes?

If respondents were unwilling to provided that figure they were then asked (q159a)

Then, could you tell me in which of the following groups your total household income falls, from all sources, last year, before taxes?
(1) Under $10,000
(2) $10,000 to less than $20,000
(3) $20,000 to less than $30,000
(4) $30,000 to less than $40,000
(5) $40,000 to less than $50,000
(6) $50,000 to less than $60,000
(7) $60,000 to less than $70,000
(8) $70,000 to less than $80,000
(9) $80,000 or more?

As detailed above, for respondents either providing only an income range or no data at all, incomes were imputed.

AGE – The respondents were asked (q156) in what year they were born. Their ages were then calculated. *A priori* it is not possible to know how a respondent's age may impact WTP. However, in general it is possible to hypothesize that as people grow older, they become more politically conservative and therefore their WTP would decrease. Consequently, the sign of the AGE variable may be expected to be negative.

ED – The effect of education on WTP is unknown, but generally expected to be positive. As people accumulate education, it is thought that people tend, on average, to become more politically liberal. A more liberal political philosophy would be expected to effect WTP positively. Also, as people accumulate education, the better able they are

[51] Clearly, the age of this data reduces its usefulness.

to understand and evaluate the nonmarket good in question. Respondents were specifically asked (q157)

What is the highest grade or year of school that you have completed?
(1) Eighth grade or less
(2) Some high school
(3) High school grad or GED certificate
(4) Some technical school or vocational training
(5) Technical school graduate
(6) Some college or associate degree
(7) College degree
(8) Postgrad or professional degree
(9) Other [specify]

MARRY – Respondents were asked (q155) about their marital status. Specifically they were asked

What is your marital or cohabitation status? Are you currently married, separated, divorced, widowed, or single?

A dummy variable was coded from the responses, where 1 = married and 0 otherwise. It is possible that married couples are in general more politically conservative and thus, would be WTP less. Alternatively, it could also be that married couples in general tend to own homes at a rate greater than those who are not married and thus feel that they would have more to loose in a flooding situation and consequently would be WTP more. The sign of the MARRY variable is concluded to be unknown, *a priori.*

TENURE – The 'tenure' variable asks (q146) respondents how long they have resided in their household. This would be expected to positively impact WTP as the longer a respondent has resided in his/her household, the greater the degree of civic responsibility felt and the greater the desire to protect one's place of residence.

MVG – This dummy variable is coded 1 for respondents indicating that they plan on moving away from the Milwaukee area in the next two years, and 0 otherwise.

---

[52] Whether the year was 1998 or 1999 depends on when the survey was administered. The survey was in the field from November 1999 to May 2000.

Clearly, if respondents are planning to move away, this would be expected to negatively effect WTP. Respondents were specifically asked (q148)

Do you expect to move away from the Milwaukee are in the next two years?

Other demographic variables were investigated including, age, gender, and race, but none were found to be significant in any of the estimations.

(b) Measures of Flood Risk (calculated and perceived)

FRM – the engineering research team using a geographical information system (GIS) and hydrologic modeling calculated the FRM variable[53]. The FRM depends on the delineation of the floodplain, which is most commonly based upon the flow with a recurrence interval T_r = 100 years. Put simply, at the boundaries of the flood plain, the risk of flooding is $r = 1/ T_r = 1/100 = 0.01$. Properties located within and outside of this "100 year floodplain" are under different risks of flooding and it thus necessary to express flood risk in a relational form. Using hydrologic models, the engineering members of the research team developed such a relational measure, the Flood Risk Measure (FRM). The probability of flooding decreases as one moves away from the river. Due to the calculation of the FRM, the FRM increases as the risk of flooding decreases and thus, the FRM would be expected to negatively impact WTP. As aforementioned, the FRM was calculated for every geocoded household possible.

However, unfortunately every estimation that used the FRM suffered from sample selection bias, with no estimations generating robust results. Therefore, no estimations including the FRM variable are reported.

DISTRVR – Due to the detrimental sample reduction as caused by unavailability of the FRM, this variable was used as a proxy for a relational flood risk measure. It is a simple measure (i.e. not relational) of the household's distance to the river. As aforementioned, as distance to the river increases, flood risk decreases. Therefore a negative relationship between DISTRVR and WTP would be expected.

LOSSFRMFLD – Respondents were asked if they had ever suffered a loss because of flooding. Respondents answering positively were coded as 1, others as 0. Households that have previously suffered losses due to flooding events would be expected to have increased levels of anxiety and sensitivities towards potential flooding, and would perceive their household to have an increased chance of flooding. Thus, this dummy variable would be expected to positively impact WTP. Specifically, respondents were asked (q149)

Have you ever suffered a loss because of the flooding of a dwelling you lived in, that is, it had standing water in a basement or other area that came from water flowing in from the outside?

(c) Community Characteristics

POPDENSITY – Information about population densities was obtained from the CensusCounts® database. The densities were calculated as 1000s of people per square mile and based upon the 1990 census data. *A priori,* it is difficult to say whether this variable would negatively or positively impact WTP. It could negatively impact WTP if residents of more densely populated areas engage in free riding behavior, or if they

[53] For a more detailed discussion of the calculation of the FRM used in this study, see Bartosova et al.

perceive the responsibility of the funding of flood control projects to be that of the government's *because* the area is densely populated. However, it is more likely that there is a positive relationship between POPDENSITY and WTP because as urbanization occurs and population densities increase, flood risk increases.

Numerous other community characteristics were investigated, including wealth, racial makeup, area, average value of homes, and average education levels. None of these variables were found to be significant in any of the estimated models.

(d) Attitudinal, Behavioral, and Value Attributes

ACTION – Respondents willing to support organizations that are active in watershed management issues would be more likely to give higher WTP and the more strongly that they feel they would lend their support, the higher their expected WTP. Therefore, the relationship between this variable and WTP is expected to be strongly positive[54]. The respondents were specifically asked (q89)

For the following statement, please tell me whether you strongly agree, agree, feel neutral, disagree, or strongly disagree. Give time or money in support of a citizen organization concerned about the matter.
(1) Strongly agree
(2) Agree
(3) Feel neutral
(4) Disagree
(5) Strongly disagree
(6) Don't know

POLITICS – The political philosophies with which respondents align themselves would be expected to influence WTP. Respondents with a liberal philosophy would tend

---

(1999). Since ultimately the FRM is not used in the final models a more detailed description is not necessary here.

[54] Interestingly, 29.7% people answering that question responded with agree, 30.1% with feel neutral, and 34.3% with disagree. In sum, 94.1% of the respondents indicated either agree, neutral, or disagree. This highlights that concerns about this variable's potential to serve as a de facto or proxy WTP measure (making it inappropriate as an explanatory variable) are unfounded. Although there will clearly be a strong correlation between a willingness to support an organization and WTP, the *form* of that support (i.e. volunteer time or money) and the *amount* of that support are not captured by ACTION, but are captured in WTP. Furthermore, the WTP measure is phrased in terms of the respondent's willingness to *vote* in favor of a plan on the next ballot.

to be more attuned to and concerned with environmental issues and thus, would be more likely to provide a higher WTP, while respondents with a conservative philosophy would not. Specifically the respondents were asked (q154)

These days we here the words "liberal" and "conservative" talked about a lot. Would you say you are: liberal, somewhat liberal, middle of the road, somewhat conservative, or conservative?
(1) Liberal
(2) Somewhat liberal
(3) Middle of the road
(4) Somewhat conservative
(5) Conservative
(6) Don't know

It would therefore be expected that POLITICS would be negatively related to WTP *a priori*.

(e) Other Variables of Empirical Importance

PROTEST – As discussed previously (in sections IV.C.3.b-d.), protest bids were not removed from the sample. Rather, a dummy variable, PROTEST, was generated. This variable was coded as 1 for any WTP response determined to be a protest bid and 0 for all others. This variable would be expected to negatively impact WTP.

INITIALBID – The value of the first bid offered the respondent in the initial WTP question (q100b and q104b), which begins the iterative bidding process, is included in the WTP estimations in order to test for starting point bias. If this variable is significant it implies that the initial bid positively or negatively influenced the final maximum WTP, and thus, the estimation suffers from starting point bias. As aforementioned (sections III.C. and IV.C.4), it is anticipated that the WTP estimation for PATH C will suffer from this bias.

PATH C – A dummy variable was generated to be used in the estimations which combined the Path A and C data sets. This dummy variable was coded 1 = Path C respondents and 0 = Path A respondents. This variable can then be used to somewhat

crudely investigate potential embedding. The null hypothesis being that embedding exists. If the Path C dummy variable's coefficient is found to be positively signed and significant, this would conclusively indicate embedding, and the null could not be rejected. If the Path C variable's coefficient was estimated and found to be insignificant, of either a positive or negative sign, it would still not be possible to reject the null. However, if the Path C variable's coefficient was found to be negative and significant, the null hypothesis would be rejected and it could be concluded that there is no embedding present. Therefore, *a priori* it is not possible to know what the sign of this variable will be.

(f) Dependent Variable

WTP – The respondents were asked their maximum willingness to pay after a series of iterative bidding questions and bounded by $0 and $500. If respondents offered a bid greater than $500, the respondent was reminded that the contribution was to be made annually for 20 years. Treatment of the nonresponse and protest bids was addressed above. Valid $0 WTP responses were recoded as $.01 to allow the variable to be transformed logarithmically so that elasticities could be obtained.

The variables and their expected signs are summarized in the following table:

Table 4-2: **Variable Names, Definitions, Sources, and Predicted Signs**

Variable Name	Definition	Source	Predicted Sign
INCOME	reported and imputed 1999 income	survey, #159	+
AGE	reported age	survey, #156	– ?
ED	reported education level (1-8)	survey, #157	+?
MARRY	1= married, 0 otherwise	survey, #155	–?
TENURE	reported # of years in dwelling	survey, #146	+
MVG	1=moving in next two years, 0 otherwise	survey, #148	–
FRM	Flood Risk Measurement	GIS	–

DISTRVR	household distance from nearest river	GIS	–
LOSSFRMFLD	1=respondent has suffered a loss from a flooding event, 0 otherwise	survey, #149	+
POPDENSITY	1990 population per square mile	CensusCounts®	+?
ACTION	willingness to support an involved citizen organization with time or money	survey, #89	+
POLITICS	reported political philosophy	survey #154	–
PROTEST	1=$0 protest bid, 0 otherwise	generated	–
INITIALBID	first iterative WTP bid	survey #100b, 104b	?
PATH C	1=Path C respondent, 0 otherwise	generated	?
WTP	reported max WTP	survey #101, 105	dependent

2. Functional Forms

This analysis assumes a utility function that is positive but increasing at a decreasing rate implying that the WTP function is similarly shaped

$$\ln Y = \beta_0 + \beta_i \ln X_i + \beta_j X_j \qquad [4\text{-}14]$$

This functional form (equation 4-14) is utilized for empirical estimation for this study. Additionally, it is of interest to estimate a variant of this functional form, using non-logged dependent and income variables. This linear functional form will be estimated as well. The statistical models to be estimated are therefore as follows:

$$\ln WTP_{PathA} = \alpha + \beta_1 \ln INC + \beta_2 DISTRVR + \beta_3 PROTEST + \beta_4 AGE + \beta_5 MARRY + \beta_6 TENURE + \beta_7 POPDENSITY + \beta_8 POLITICS + \beta_9 ACTION + \varepsilon \qquad [4\text{-}15]$$

$$\ln WTP_{PathC} = \alpha + \beta_1 \ln INC + \beta_2 DISTRVR + \beta_3 PROTEST + \beta_4 AGE + \beta_5 MARRY + \beta_6 TENURE + \beta_7 POPDENSITY + \beta_8 POLITICS + \varepsilon \qquad [4\text{-}16]$$

These specified statistical models are linear in their parameters and therefore, ordinary least squares (OLS) estimation is appropriate.[55]

[55] The appendix includes estimations of equations 4-15 and 4-16 with the INITIALBID variable included. The appendix also includes, for comparison purposes and to investigate embedding, the following estimations:

$$\ln WTP_{PathA} = \alpha + \beta_1 \ln INC + \beta_2 DISTRVR + \beta_3 PROTEST + \beta_4 AGE + \beta_5 MARRY + \beta_6 TENURE + \beta_7 POPDENSITY + \beta_8 POLITICS + \beta_9 MVG + \beta_{10} ED + \varepsilon \qquad [appendix\text{-}1]$$

V. Empirical Results

The following table summarizes the estimations of the above equations:

Table 5-1: Ordinary Least Squares Estimates
Dependent Variable = **ln(WTP)**

EQUATION:	**4-15**			**4-16**		
VARIABLE	coefficient	t-statistic	p-value	coefficient	t-statistic	p-value
CONSTANT	-16.24439	-3.778118	0.0002	-18.12886	-3.472377	0.0006
ln(INC)	1.559599	4.146796	0.0000	2.021404	4.609592	0.0000
DISTRVR	-0.000555	-1.414783	0.1583	0.000228	0.445091	0.6567
PROTEST	-6.846328	-7.580717	0.0000	-5.710148	-4.811984	0.0000
AGE	-0.031911	-1.799268	0.0731	-0.012399	-0.522873	0.6016
MARRY	-0.506344	-1.075435	0.2831	-0.654675	-1.166870	0.2446
TENURE	0.042946	2.095494	0.0371	-0.006526	-0.264067	0.7920
POPDENSITY	4.47E-05	0.986747	0.3246	-6.82E05	-1.464361	0.1446
POLITICS	-2.04805	-1.184260	0.2373	-0.109060	-0.509181	0.6112
ACTION	1.290081	4.841934	0.0000			
# of observations	295			233		
Adjusted R^2	.284575			.212483		
S.E. of Regression	3.351679			3.736392		
Mean WTP	\$79.58$^{\blacklozenge}$, \$83.56$^{\#}$			\$78.74$^{\blacklozenge}$, \$82.64$^{\#}$		

Where $^{\blacklozenge}$ indicates mean WTP calculated with protest bids included
$^{\#}$ indicates mean WTP calculated without protest bids

$$\ln WTP_{PathC} = \alpha + \beta_1 \ln INC + \beta_2 DISTRVR + \beta_3 PROTEST + \beta_4 AGE + \beta_5 MARRY + \beta_6 TENURE + \beta_7 POPDENSITY + \beta_8 POLITICS + \beta_9 MVG + \beta_{10} ED + \varepsilon \quad \text{[appendix-2]}$$

$$\ln WTP_{Combopath} = \alpha + \beta_1 \ln INC + \beta_2 DISTRVR + \beta_3 PROTEST + \beta_4 AGE + \beta_5 MARRY + \beta_6 TENURE + \beta_7 POPDENSITY + \beta_8 POLITICS + \beta_9 MVG + \beta_{10} ED + \beta_{11} PATHC + \beta_{12} PATH\ C*\ln INC + \beta_{12} PATHC*DISTRVR + \beta_{13} PATHC*AGE + \beta_{14} PATHC*MARRY + \beta_{15} PATHC*TENURE + \beta_{16} PATHC*POPDENSITY + \beta_{17} PATHC*POLITICS + \varepsilon \quad \text{[appendix-3]}$$

The ultimate objective of this empirical investigation is to calculate benefit estimates for the maintenance of current levels of flood risk, by using the information provided by respondents in the WTP survey. This requires the use of a variable which measures the spatially mutable flood risk (i.e. a household's level flood risk can change – say for instance, if a flood management plan was enacted – however, a household's distance from the river is fixed). However, due to the effects of the spatial bias resulting from the limitations of the FRM, it is not possible to generate such benefit estimates. Despite this, it is possible to use a proxy for the FRM, DISTRVR, in order to investigate how the WTP variable is impacted by various changes.

The specification of equation 4-15 was motivated primarily by theory and relevant literature, as discussed. Other variables were explored, including numerous attitudinal variables (i.e. feelings about taxation, connectedness to the community, perceptions about program effectiveness, past flooding experiences, etc.), demographic variables (i.e. education, gender, and race), and community characteristics (i.e. blockgroup, average value of households). However, none of these additional variables were significant with the expected signs and very few added to the explanatory power of the estimation. Equation 4-16 was then estimated using a specification that was nearly identical to the final specification of equation 4-15.[56]

Although equations 4-15 and 4-16 are nearly identical, equation 4-16 has fewer significant coefficients (even with one-tailed t-tests and considering only marginal significance), with two variables displaying incorrect signs, and the DISTRVR variable is

highly insignificant. In contrast, Equation 4-15's estimation has more significant variables, with all variables displaying the anticipated signs, and the DISTRVR variable *is* significant. The model's adjusted R-squared is higher and its standard error, Akaike and Schwartz criteria are lower than that for equation 4-16. However, the two models' results are generally quite similar.[57] It is therefore important to discuss the results of these estimations considering a number of issues. First, spatial bias is not expected to be present in the Path A estimations. It is however, expected to be present in Path C estimations. This may partially explain why DISTRVR is significant in equation 4-15, but is not significant in 4-16.

Second, Path C estimations had fewer explanatory variables from which to draw upon for specification purposes. There are no potentially explanatory variables for Path C that reflect respondents' attitudes, beliefs, values, and/or behaviors. Thus, many more attitudinal and value variables were investigated for Path A and the final model for equation 4-15 includes the ACTION variable, which is not available for Path C. ACTION is highly significant and displays the anticipated sign.[58] The variable LOSSFRMFLD was included in Path C estimations in an effort to overcome this by gauging respondents' probable feelings about flooding. This attempt was unsuccessful as LOSSFRMFLD was consistently found to be highly insignificant with an incorrect sign.

---

[56] Appendix D includes estimations of both equations 4-15 and 4-16, but including the INITIALBID variable in order to investigate starting point bias (see forthcoming discussion), titled 4-15b and 4-16b,respectively.

[57] Appendix E includes a final specification for Path C ("Appendix-2"), which includes the ED and MVG variables (they were highly significant for the Path C estimations) and an identically specified estimation for Path A ("Appendix-1"). Comparable results are found for these estimations.

[58] However, turning to Appendix E, when investigating the identically specified estimations for equations Appendix-1 and Appendix-2, the Appendix-1 estimation still has more significant explanatory variables than Appendix-2, with DISTRVR significant in Appendix-1 as well.

Third, considering both the joint nature of the good to be valued and the lack of time allowed for the respondent to reflect upon the topic and the nonmarket good at hand, the respondents in Path C faced serious cognitive obstacles, as discussed in sections III.C. and IV.C.4. These severe cognitive obstacles may induce respondents to base their WTP on the initial bid offered to them, thereby causing starting point bias. This would not be expected for Path A estimations, but would be anticipated for Path C estimations. The INITIALBID variable was included in order to determine if starting point bias is occurring. If the INTIALBID variable is significant, starting point bias is present. If INTIALBID is insignificant, no starting point bias exists. Appendix D includes estimations of equations 4-15 and 4-16, including the INITIALBID variable, titled 4-15b and 4-16b, respectively. Appendix E includes an additional Path A and Path C estimation, titled Appendix-1 and Appedix-2 respectively, both of which include the INTIALBID variable. As revealed by the sign and significance of the INITIALBID variable's coefficient in each of the four estimations, the expected result occurred – the Path C models suffer from extreme starting point bias, while the Path A models do not[59].

Fourth, it is important to analyze any symptoms of embedding. Since Path C respondents are valuing a joint good (i.e. a policy package that would reduce flooding *and* improve the ecological quality of the watershed) while Path A respondents are being asked *only* to value a flood control plan, it would be expected that Path A's mean WTP < Path C's mean WTP. However, Path C's mean WTP is *less* than the mean WTP of Path A, potentially indicating an embedding phenomena.

[59] The differences in the p-value for Path A and C are so great that they would seem alarming. However, it should be recalled the dramatic difference in the timing of the WTP elicitations for each path. Path A respondents were asked WTP after approximately 25 minutes while Path C respondents were asked virtually immediately.

In order to examine this further, the data sets for Path A and C were combined. All data that was available for Path A, but not available for Path C was not included in the combined data set. In addition, a Path C dummy variable, PATHC, was generated as discussed in section IV.D.1. Next, a model was estimated with this combined data set (see AppendixE, equation Appendix-3), including numerous PATHC interaction terms. The estimated PATHC coefficient is negative, but insignificant. Therefore it is not possible to determine conclusively whether or not embedding exists (i.e. not able to reject the null hypothesis that embedding exists), although it seems unlikely as the PATHC coefficient is negative and insignificant rather than significant and positive, which would definitively indicate embedding. Considering that 1) the Path C respondents were facing severe cognitive difficulties, 2) were likely most familiar with the flood risk maintenance aspect of the good they were valuing, and 3) were decidedly unfamiliar with the ecological aspect, they would tend to use the flood risk amenity as a reference value. Remembering also that there is conclusive evidence of starting point bias in Path C estimations, it seems likely that any embedding is being caused by Path C respondents' efforts to overcome cognitive obstacles, by either cueing off the initial bid, referencing the more familiar flood risk good, or a combination of both. In sum, Path C's unfortunately designed WTP elicitation for a joint, policy packaged good seems to have caused starting point bias and potentially embedding. These facts, when coupled with Path C's unavoidable spatial bias and more limited data set, demonstrates that Path C's estimations should be examined with great caution. In contrast, Path A estimations are not subject to these cautionary caveats.

All models were tested for heteroskedasticity using White's Test. The results of those tests, both with and without cross terms, indicated no heteroskedasticity present in the equation 4-15 model, nor in any of the estimated models. The F-statistics for equation 4-15 were .2642 and .4715 corresponding to the exclusion and inclusion of cross terms respectively.

However, evidence of multicollinearity was found in both of the estimations. The Path C estimation exhibited numerous symptoms of nulticollinearity: highly sensitive estimations, estimated coefficients with incorrect signs, relatively high standard errors, and somewhat low t-statistics for the independent variables with a high model F-statistic (8.82). Path A, also exhibited similar symptoms of multicollinearity[60]. However, despite the fact that there are symptoms of mutlicollinearity in Path A estimations, all the estimated coefficients of the explanatory variables have the expected signs and all at least marginally statistically significant (with one-tailed t-tests). Furthermore, even if mutlicollinearity does exist, it will not render the estimation biased or inconsistent.

Based upon the above results and discussion, the forthcoming interpretations will be drawn from the estimation of equation 4-15, and implications, and future research initiatives will be elucidated. The coefficients of the estimation can be interpreted as follows, *ceteris paribus*[61]:

- INC – A 1% increase in household income will result in a 1.5596% increase in WTP. This suggests that flood risk management may be luxury good as its income elasticity exceeds 1%[62].

[60] The simple correlation coefficient was found between the ACTION variable and WTP = .29 while the unadjusted R-squared of the estimation was .38.

[61] For explanatory variables in level form, their elasticities were also calculated at the mean when appropriate.

[62] It is possible to use a one-tailed t-test of significance to determine if the coefficient of lnINC (β_1) is significantly greater than 1. The computed t-ratio = 1.4878. At the 10% level of significance, the null hypothesis that $\beta_1 = 1$ is rejected and it can be concluded that β_1 is statistically significantly >1 and flood risk is a luxury good.

- INC – Evaluated at the mean income, a $10,000 increase in income will result in a $5.03 increase in WTP.
- DISTRVR – For every 1000 meter increase in a household's distance from the river, WTP decreases by .555%.
- DISTRVR – For every 1% increase in a household's distance from the river, WTP decreases by .4399%.
- DISTRVR – Evaluated at the mean distance, a 1000 meter increase in a household's distance from the river will decrease WTP by $5.42.
- PROTEST – When protest bidders are included in the estimation, they negatively effect WTP.
- AGE – For every year older a respondent is, WTP decreases by .0319%.
- AGE – A 1% increase in respondent age will result in a 1.679% decrease in WTP.
- MARRY – If a respondent is married, WTP decreases by .5063%.[63]
- TENURE – For every additional year lived in a dwelling, WTP increases by .0429%.
- TENURE – For every 1% increase in the length of time lived in a residence, WTP increases by .6697%.
- POPDENSITY – For every 1 unit increase in population density WTP increases by .0000447%.[63]
- POPDENSITY – For every 1% increase in population density, WTP increases by .2860%.[63]
- POLITICS – The more politically conservative a respondent considers himself, WTP decreases by 2.0485%.[63]
- ACTION – The more strongly that a respondent feel that he/she would be willing to lend support to a community organization concerned with flood risk management, WTP increases by 1.290%.

In sum, the Path C equation evidenced spatial and starting point biases, as well as potential embedding. In contrast, the Path A estimation was free from any signs of spatial and starting point biases. The Path A estimation also showed no evidence of heteroskedasticity and few symptoms of multicollinearity. The estimation displayed sufficient reliability and validity, providing useful and revealing coefficient interpretations. Furthermore, the mean WTP values for Paths A and C as well as the results from the combined estimation seem to indicate the presence of embedding, potentially caused directly by the cognitive challenges faced by Path C respondents. Although it was not possible to obtain benefit estimates for the maintenance of current

[63] A marginally significant variable, included because theory indicates it should be important.

levels of flood risk due to the limitations imposed by the FRM variable, the insights gained in this estimation will allow a more robust WTP function to later be estimated.

VI. Conclusions and Implications

The purpose of this analysis was to test the potential of the CVM WTP survey method as a flood control benefit assessment technique. This study also attempted to determine the relative importance of demographic, attitudinal, and community characteristics in a household's reported WTP. Finally, this study attempts to identify potential research extensions of this survey data.

While it was not possible to obtain benefit estimates, this study did confirm the hypothesis that a household's WTP for the maintenance of current levels of flood risk is positively related to flood risk as proxied by the household's distance from the river. More generally, this study supports the more general hypothesis that WTP is positively related to the provision of flood protection. In addition, this study demonstrated that cognitive obstacles are substantial, particularly when asking respondents to value more than one nonmarketed good jointly (as a policy package) and when respondents are not given adequate time to reflect about the nonmarket good to be valued. This study further demonstrated that these cognitive difficulties can prove to be so great as to create severe starting point biases, confirming CVM literature. These starting point biases may, in fact, be of such a magnitude that the embedding phenomena may appear to be present. In addition, the income elasticity of flood risk management was estimated to be 1.5596%, indicating that flood risk management may be a luxury good[64]. This study also demonstrated that attitudinal and behavioral variables are critically important in the estimation of WTP, confirming CVM literature.

These conclusions' implications for the larger research project of which this work was a part of are clear. First, all households must be geocoded so to avoid any initial

spatial bias. Second, the digitization of the flood plains must be completed. This will allow a FRM to be calculated for all households, thereby making benefit estimation and cost-benefit policy analysis possible. Furthermore, when spatial bias problems are overcome and when FRMs are available, models should be re-specified, investigated all available and theoretically viable variables. Third, the starting point bias clearly apparent in Path C estimations, but absent in Path A estimations, should be investigated further. This research program has the unique opportunity to explore two estimations, valuing the similar goods, one which does not display starting point bias and one that strongly exhibits the bias. Relatedly and fourth, when spatial bias problems are overcome, the embedding phenomena should be reinvestigated, paying special attention to a potential interrelationship between embedding and cognitive difficulties (and thus starting point bias). Fifth, in future surveys, respondents, particularly those valuing a joint good in a policy package context, should be given ample time to reflect about the issues at hand. This can be accomplished by asking would-be Path C respondents attitudinal and behavioral questions similar to those asked of Path A respondents. Sixth, considering the bounded, iterative bidding format of the WTP elicitations, different estimation methods should be explored. Finally, when the sample is extended with the completion of the geocoding process and the recalculation of the FRM, the identification and treatment of protest bidders must be reassessed.

In the context of the greater CVM research community, this study confirms two basic, yet fundamental CVM concepts. First, this study has shown that the CVM, in particular this WTP survey, can potentially be used to generate viable benefit estimates useful for policymakers. Second, this study has shown that the survey instrument is a

---

[64] Although this was confirmed by the t-test.

delicate one and that the difficulty and nuances behind “the art of asking questions” (Payne, 1951) must never be underestimated, particularly those which may induce starting point biases.