Title
Public Preferences Towards Environmental Risks: The Case of Trihalomethanes

Authors
Carson, Richard T
Mitchell, Robert Cameron

Publication Date
2000-08-01

2000-21

UNIVERSITY OF CALIFORNIA, SAN DIEGO

DEPARTMENT OF ECONOMICS

PUBLIC PREFERENCES TOWARD ENVIRONMENTAL RISKS: THE CASE OF TRIHALOMETHANES

BY

RICHARD T. CARSON

AND

ROBERT CAMERON MITCHELL

DISCUSSION PAPER 2000-21
AUGUST 2000

Public Preferences Toward Environmental Risks

The Case of Trihalomethanes[1]

Richard T. Carson

Department of Economics
UC San Diego

Robert Cameron Mitchell

Graduate School of Geography
Clark University

August 2000

JEL codes: Q25, Q26, D81

Keywords: contingent valuation, statistical life

[1] We gratefully acknowledge financial support from the U.S. Environmental Protection Agency (Cooperative Agreement CR810466-01-6) in undertaking some of the work reported here. The views expressed here, however, are solely those of the authors. Earlier versions of this paper were presented at meetings of the Society for Risk Analysis and the Association for Public Policy Analysis and Management. We thank Anna Alberini, Maureen Cropper and George Parsons for helpful comments on earlier versions of this paper.

ABSTRACT

We present the results of an in-depth study in a small Southern Illinois town looking at the public's preferences with respect to reducing trihalomethanes (THMs) in their public drinking water system. THMs are an interesting environmental risk to study. First they are a low-level risk created as a byproduct (via chlorination) of reducing the much larger risk of bacterial contamination. Second, THMs are a weak carcinogen (with a scientific debate over how weak) with a long latency period. Third, small towns pose an interesting policy trade-off question with respect to THMs due to the sharply rising per capita cost of carbon filtration as population decreases. Further, filtration at the home or tap level is a viable alternative to public filtration. These issues are considered in the context of designing a survey to elicit maximum willingness to pay (WTP) for a reduction in THMs. The key survey design question involves how to communicate low-level risks of different magnitudes to respondents. Respondents were randomly assigned to different risk levels and statistical tests reject the hypothesis that WTP estimates are insensitive to the risk levels assigned. Our value of a statistical life estimates are quite low relative to most estimates in the literature. Our estimates should be low, however, if respondents discount due to the long latency period. After allowing for discounting using commonly used rates, our value of a statistical life estimates are well within the range commonly found in the literature for WTP to avoid current period fatal accidents.

Introduction

This paper presents the results of an in-depth study conducted in Herrin, a small Illinois town with a population of about ten thousand people.[2] Our study focused on the issue of the benefits of a town installing a carbon filtration system to remove trihalomethanes (THMs) from its drinking water system. The removal of THMs from drinking water has long been a controversial issue and the class of chemicals has a number of properties that make them an interesting topic for those interested in risk analysis. We examine the public's preferences toward a proposed policy that would reduce the level of THMs in the town's drinking water supply. The process of explaining the key characteristics of the risks associated with THMs and the policy decision whether to reduce that risk are explored in the context of a contingent valuation (CV) survey designed to measure willingness to pay (WTP) to implement the policy. A variety of tests are conducted to assess the properties of the WTP estimates for use in policy decisions.

THMs are a class of chemicals created during the process of chlorinating drinking water (Culp, 1984). They have been consistently shown to be carcinogenic but represent a low-level risk (Culp, 1984; Attias *et al*., 1995). In November 1979 the U.S. Environmental Protection Agency (EPA) under the 1974 Safe Drinking Water Act set an interim Maximum Contaminant Level (MCL) for total trihalomethanes (THMs) of 0.10 mg/l as an annual average (44 FR 68624). To put the risk from THMs in perspective, chlorinated water containing THMs are estimated to kill between 2 and 100 people per year in the United States, largely through increased incident of urinary tract cancer. The latency period associated with this class of carcinogens is thought to be in the 20 to 30 year range. THMs in chlorinated water present a very low level of risk to drinking water consumers. The risk from not chlorinating the water is dramatically higher and much more immediate since the chlorine kills biological contaminants. Removal of THMs from drinking water is a relatively straightforward matter that involves passing the water over some type of activated carbon filter. The major public policy discussion on this issue is whether small towns should be exempt from meeting the U.S. EPA THM standard. At present, the interim THM standard only applies to municipal water systems (surface water and/or ground water) serving at least 10,000 people that add a disinfectant to the drinking water during any part of the treatment process. Small towns pose an interesting policy trade-off question with respect to THMs due to the sharply rising per capita cost of carbon filtration as population decreases. Other drinking water contaminants such as arsenic pose a similar size of place versus per capita cost tradeoff.

We consider these issues in the context of designing a survey to elicit maximum WTP for a reduction in THMs. The key survey design question involves how to communicate low-level risks to respondents. The literature on risk perception contains many examples of the difficulty lay people have in grasping just how low many risks are (*e.g.*, Davies, Covello, and Allen, 1986; Fisher, Pavolva, and Covello, 1991; Rimer and Nevel, 1999). Placing the risk communication exercise in the context of a CV survey allows us to judge the success of the exercise by comparing our findings with a set of well-defined economic predictions.

Preliminary Risk Communication Research

We chose Herrin, and two of its neighbors where some of the survey development work was conducted, Carbondale and Marion, because their water systems had exceeded the THM standard several times in the past three years. As required by the EPA, the water systems sent their customers an offical notice reporting this event. In the course of designing the questionnaire, we conducted four focus groups followed by a series of in-depth interviews over a six month time period to gain insight into the difficulties local residents might have in understanding the risks posed by excess THMs and how we might design the survey instrument's scenario to accurately communicate these risk levels. Among our findings and design decisions were the following:

• Despite receiving a notice that their water system had exceeded the EPA standard for THMs, the focus group participants were mostly unaware of the notice and of THMs and their risks. When they considered THMs they tended to confuse them with other drinking water risks, especially those from PCB contamination that had been an issue in a nearby city several years earlier. To deal with the low level of knowledge about THMs, we included basic information in the survey instrument about THMs and the risks they pose and highlighted the distinction between THMs and PCBs in several places in the questionnaire.

• Focus group participants tended to place a higher value on collective drinking water improvements than on improvements taken by individual consumers such as the purchase of bottled water or the installation of under-the-sink purification devices. Spontaneous comments by respondents in our in-depth interviews clearly indicated that many citizens hold a strong

---

[2] For Herrin's official web page see http://www.VillageProfile.com/illinois/herrin/index.html.

preference for collective improvements because they value knowing that their fellow townspeople will also be protected.[3] In our survey we measure the value of collective drinking water improvements.

•Participants were sensitive to how local authorities treat a risk like that posed by excess THMs. On the one hand they are generally skeptical about assurances from the local authorities that the drinking water is "safe," on the other hand they believed strongly that if there really was a serious problem with their drinking water, the authorities and especially the local press would publicize the seriousness of the situation. This default assumption required us to specify how local authorities viewed the risks posed by the levels of excess THMs in our scenarios. We modeled this information on the actual assurances the relevant authorities had given at the time of the local THM excess risk notifications.[4] It also turned out to be important to set up the divergence in views between the U.S. EPA and state and local agencies as a way of justifying the need for public input on the issue. It was necessary to explain the concept of the U.S. EPA maximum contaminant level (MCL) and the fact that THM levels below the MCL were not risk free, only less risky. We described the risk reductions in terms of bringing an existing THM risk down to the MCL risk level.

•The focus group participants found it difficult to grasp the nature of mortality risks and how a contaminant can affect them. We adopted the following approach to communicate this information: The first part of our scenario described the concept of a "basic" risk everyone faces of dying which get greater as people get older. We used the example of how life insurance premiums increase with age. We then introduced the concept of what we called "extra" or "special" risks to which some people are exposed and others are not and described magnitude of the extra risk a Hollywood stuntman, a police officer and someone taking a single airline flight is exposed to. In an attempt to acquaint people with the concept of the monetary value of risk

---

[3] This suggests that purchases of household filtration devices would not be fully reflective of the full range of relevant benefits for a policy that would involve collective provision of the risk improvement.

[4] In each of the three communities we studied, the local authorities consistently downplayed the danger posed by the excess THMs. Our findings in this study are contingent on this scenario element as we believe that if our scenario had stated that the state and local authorities urged voters to support a referendum to raise water rates to cover the cost of reducing THMs, respondents might have been willing to pay significantly more money for the same risk reductions.

reductions the scenario said that someone who is a stuntman would have to pay a large extra premium for life insurance on top of his or her basic premium.

•Not surprisingly, it was hard for focus group participants to grasp how "low" a low-level risk was. We used several examples of low-level risks to help communicate this concept. One was the risk of dying in the crash of a scheduled U.S. airliner flight. Another was the risk of being killed by lightning. In the scenario we told the respondents that out of two million people who die in the U.S. from all causes, lightning kills only 116. We said people would only have to pay five cents for a $100,000 life insurance policy against being killed by lightning in any given year because this type of death is so infrequent.

•We found that casting various risks in terms of the risk of smoking cigarettes helped our pretest subjects to grasp the relative magnitude of low-level risks. We calculated that the risk of being killed when taking a single airliner trip was equivalent to the extra risk caused by the lifetime consumption of two cigarettes, and that the extra risk of being a stuntman was equivalent to smoking 33,000 cigarettes in a lifetime. The use of cigarette equivalence for the risk from excess THM also had the effect of emphasizing its low-level carcinogenic nature and 20-30 year latency period.[5]

•As we expected, people found it difficult to judge whether a particular risk improvement was large or small when it was described solely in numerical terms, such as annual deaths per 100,000 people. They wanted information about how this risk compared with other risks. We developed several risk communication techniques to overcome this problem, the most important of which is a risk ladder. A great deal of our instrument development work was spent testing various types of risk ladders.

•The focus group participants had difficulty understanding the first type of risk ladder we tested, a Smith-Desvousges type ladder which used a logarithmic scale to locate the number of people who die annually from various causes or activities (Smith and Desvousges, 1987). Smith and Desvousges used this ladder as a visual aid to communicate the risk from hazardous wastes in a CV study. The large range of different types of risk levels on their ladder—from the annual

---

[5] Neither of us are experts in risk analysis, therefore our translation of risk levels reported in the literature to the various risk levels we use in our survey should not be viewed as authoritative in any respect. We believe the risk comparisons we present to the respondents are sufficiently accurate for our purposes and that the substitution of more accurate estimates would not substantially change our findings. We also believe there is a need for an authoritative catalog of risk levels quantified in easy-to-understand metrics such as cigarette equivalents.

risk of stuntmen dying (2,000 of 100, 000) to the .05 deaths per year due to floods—and their use of a single scale to represent the full range of risks from high to low made it difficult for them to squeeze the ladder into a single page small enough for the interviewers to use comfortably in the field. Their solution to this problem was to use a logarithmic scale and to show breaks in the ladder between the different risk intervals.[6] We concluded their risk ladder was not satisfactory for our purposes because their scale did not clearly describe the range of low level risk (below 1.0 in 100,000) comparisons we wished to show and because many focus group participants found it difficult to understand logarithmic scale.

•Tests of an alternative risk ladder in our in-depth interviewers were favorable. The new ladder continues to use annual mortality per 100,000 people to depict risk levels and to use risk examples as anchors. The most important ways it differs from the Smith-Desvousges risk ladder are: (1) it uses an equal interval rather than a logarithmic scale and (2) it employs the device of magnifying the very lowest portions of the ladder to facilitate the portrayal of the very small risk reductions we asked our respondents to value. Respondents are first shown a base ladder for the full range of risk levels between 0 annual deaths per hundred thousand to 1,000 deaths per hundred thousand people. The yearly chance of death faced by people in several age ranges is listed on the left and five examples of "extra" risks on the right. After acquainting the respondent with the information on the base ladder, the interviewer unfolds a second page that is displayed on the right, next to the base ladder. This page shows two successive expansions of the risk levels at the bottom of the base ladder. The first sub-ladder expands the 0-25 mortality rate segment of the base ladder approximately 19 times and the second expands the 0-1 mortality rate of the first sub-ladder by an equivalent amount. The annual extra risk of dying per 100,000 people is shown on the sub-ladders for 13 occupations or situations.[7]

---

[6] Smith and Desvousges augmented their ladder with pie charts to show the low-level risk changes from controlling hazardous wastes they asked their respondents to value. In our view, however, the use of separate pie charts to show low level risk changes separates them from the important context provided by the risk ladder. Smith and Desvousges also attempted the difficult task of trying to convey separate probabilities for the risk of exposure and the risk of mortality if exposed. We did not face this issue in this study (which is typical of many environmental risks) because all households were exposed to THM via their household water supply.

[7] We pretested our ladder by using it in a series of in-depth interviews using our draft instrument. Respondents appeared to have much less trouble understanding the risk comparisons displayed on our ladder than they did with the Smith-Desvousges ladder. The interviewers reported that many respondents expressed considerable interest in the risk information presented in this fashion.

•The choice of which comparative risks to display on a risk ladder obviously plays an important role in framing the THM risks for the respondents.[8] We experimented with various types of risks to find the ones respondents found most meaningful to compare with the THM risk improvements. For example, we added the risk of dying in an auto accident because people asked where this would be placed on the ladder and dropped several recreational risk examples (such as dying while hang gliding) because focus group participants regarded voluntary risks such as these irrelevant to the risks imposed on drinking water users.

Figure 1 shows a black and white version of our final risk ladder. We used a color-coded version of this ladder in the field to underscore the different risk ranges. (FIGURE 1 ABOUT HERE)

Structure of the Survey Instrument

We can best summarize the risk communication portion of our scenario by listing the relevant topics in the sequence they are presented to the respondent: (1) Explanation of the relationship of "extra" risk to "basic" risk, (2) examples of low level risks, (3) the distinction between voluntary and involuntary extra risk, (4) application of a number of cigarettes per lifetime metric to the risk examples discussed earlier in the scenario,[9] (5) explanation of our risk ladder, and (6) description of how THMs are created in drinking water, the risk they pose at the U.S. EPA maximum contaminant level, and where this risk is located on the risk ladder.

After telling respondents about THMs and how THMs could be reduced, we informed them that any reduction would require an increase in their water bill to pay for a drinking water filtration bond issue. They are told Herrin could get one of three different risk reduction programs, A-C, by installing different levels of carbon filtration technology on the town's water treatment plant, each of which would bring the town's existing THM risk down to the U.S. EPA

---

[8] That context influences risk perception is well documented in the experimental literature on risk perception (*e.g.*, Kahneman, Slovic, and Tversky, 1982). The inevitability of this phenomenon and the potential magnitude of these effects place a burden on the researcher to justify the context he or she uses. The context we used was intended to minimize potential sources of bias and to be policy relevant.

[9] Two examples are: (1) the risk of being killed when taking a single airliner trip is equivalent to extra risk caused by the lifetime consumption of two cigarettes and (2) the extra risk of being a stuntman is equivalent to smoking 33,000 cigarettes in a lifetime.

MCL risk level. The respondents are shown a risk ladder on which the size of the risk reduction offered by each program is labeled A, B and C.

At this point in the interview, respondents are asked about their willingness to pay for the reduction in risk from THMs. They are first asked whether they would potentially be prepared to vote in favor of taxing themselves to reduce THMs in their drinking water. If the answer is no, the respondent is recorded as being willing to pay $0.[10] If the answer is yes, the respondent is asked to value THM risk improvements by stating the maximum amount his or her household water bill could be increased and have the respondent still favor installing the carbon filtration system for each of three THM risk improvements where C = smallest, B = middle, and A = largest.

We varied the size and sequence of the risk improvements. First we randomly assigned respondents to one of two sets of risk improvements (A or B) with the largest risk reduction in the A set being smaller than the smallest risk reduction in the B set. Second, within treatments A and B there was a further random assignment that varied the order in which the levels were presented. Thus, there are four subsamples (A1, A2, B1, and B2). Table 1 displays the properties of the experimental design used in terms of these four subsamples. (TABLE 1 ABOUT HERE) This design allows us to test whether respondent WTP estimates are sensitive to: (1) the magnitude of risk levels they are asked to value, (2) the order in which the particular risk reduction levels are asked, and (3) the interaction between the magnitude of the risk levels and question order. Our design allows for both out-of-sample and within sample tests of the sensitivity of the WTP estimates to the magnitude of the risk reduction valued.

Card B11 shown in Figure 2 displays the visual representation used for the 2.43 x 10^{-5} risk reduction, which was the first risk reduction one of the subsamples of respondents was asked about. (FIGURE 2 ABOUT HERE) This card shows the change in initial (FROM) to subsequent levels (TO) if the program was implemented as well as the actual magnitude of the change. In order to give the respondents the most meaningful context for making their risk valuation

[10]Note that this approach makes estimated aggregate WTP more conservative since some respondents will have a strategic incentive to answer no on the basis of their perception of the expected cost of the filtration system if their actual WTP is less than that expected cost. Respondents who were asked what the cost of the program would be were told they would later have a chance to state the maximum they were prepared to vote in favor of. We considered the option asking respondents how they would vote if the cost of the program was $0, but found that this was an implausible question for many respondents.

judgments, the respondent is shown a card which describes each risk reduction in terms of: (1) the absolute change in THM levels, (2) the general risk of dying per 100,000, and (3) the cigarette equivalent consumption which in this case is a reduction from 65 cigarettes in a lifetime to11. Each respondent is shown similar cards for the other two risk reductions the respondent was asked about.

Our THM survey instrument uses an open-ended elicitation format. The open-ended format is not incentive compatible in the sense that truthful preference revelation is not always an optimal strategy (Hoehn and Randall, 1987) and, further, valuation questions concerning multiple levels of a public good where only one level of the good can be supplied are also known not to be incentive compatible (Carson, Groves, and Machina, 1999).[11] It is not possible to examine the first type of strategic behavior in the context of this study, but it is possible to look at the issue of the strategic incentives that follow from asking respondents multiple questions about different levels of the same good as long as we are prepared to assume that the valuation for the good should be a smooth continuous monotonic function of the quantity of the good. Under this condition, respondents should increase (or decrease) the value of the largest risk they are asked to value relative to the smallest, in order to encourage the risk-reduction agency to either supply more or less risk reduction. We formally test this proposition in the construct validity section below.

Another component of our scenario design is the use of a water utility bill as the payment vehicle. We find respondents accepted this vehicle as realistic because of its close tie to the actual problem. We ask for willingness to pay in annual payments that has the desirable property of conveying the idea that the filtration system, once installed, would continue to be operated.

After the valuation questions are asked, the survey instrument probes respondent motives for why they answered the valuation questions in the way they did. These are followed by questions measuring attitudinal and behavioral information and demographic characteristics. Respondents are also given the opportunity to revise their WTP amounts.

---

[11] Other possible elicitation formats such as multinomial choice or a sequence of binary choices provide different, and often, harder to analyze incentives for strategic behavior than do open-ended type formats in the context of multiple levels of the same public good. Those elicitation formats further require one to specify the cost structure of the different risk reduction levels, which should influence the respondent's optimal response strategies. The ideal elicitation framework, asking each respondent a single binary discrete choice question about a single risk reduction was well beyond the budget constraint for this project given the need for lengthy in-person interviews.

Sensitivity of WTP Estimates to Risk Reduction Magnitude and Question Order

The issue of the sensitivity of WTP estimates to the magnitude of the good being valued has been the subject of considerable debate in the literature (Mitchell and Carson, 1989; Kahneman and Knetsch, 1992; Arrow *et al.*, 1993; Carson, 1997). Mitchell and Carson (1989) originally referred to this issue as “metric bias” and note that it occurs in the context of risk reductions when respondents treat the risk reductions asked about in an ordinal manner rather than considering the actual magnitude of the risk reductions asked about. Contingent valuation estimates of WTP for risk reductions have in several instances been shown to be insensitive to the magnitude of the risk reduction being valued (Beattie *et al.*, 1998; Graham and Hammitt, 1999). Such results have the obvious and troubling implication that WTP estimates obtained by a contingent valuation survey are not reliable enough to be used by policy makers.

Several reasons have been advanced to explain why WTP estimates for a given level of risk reduction vary from study to study (Kahneman and Knetsch, 1992; Baron and Greene, 1996; Fetherstonhaugh, *et al.,* 1997). One is that the initial risk communication exercise was improperly designed. Since the WTP questions depend crucially on the success of the initial risk communication exercise, its failure should translate into CV responses that do not have the expected economic properties. In this regard, low-level risk, and particularly, differences in low-level long-term environmental risks are well known for being very difficult to effectively communicate (Fischhoff, 1990).[12] Further, our examination of the existing literature suggests that most of the troublesome cases appear in surveys where the risk reduction policy is briefly explained and the risk reduction is presented in quantitative terms with little additional context. Problems seem to be concentrated in telephone surveys where it is not possible to use visual aids. To help overcome these difficulties, we used in-person interviews with an extensive oral presentation coupled with visual aids describing both low-level risks in general and the specific THM risks that were the main focus of the CV exercise.

The major competing explanation emphasizes the inherent cognitive difficulty of understanding low-level risks and of placing monetary values on environmental amenities. The psychological heuristics and biases literature that predicts a lack of sensitivity to the quantity of

---

[12] Risk communication failures are, of course, not limited to surveys. People often do not behave in the (economically) expected manner in actual market choices involving risk.

the good being valued emphasizes the role framing effects have on human risk judgments. One of these is that the order in which risk-reduction goods are valued will affect how respondents value the goods (Moore, 1999) even when respondents are warned that they will be asked to value several risk improvements. In contrast, the standard economic framework does not predict an order effect if the entire sequence is known in advance of asking the first question. The experimental design adopted above allows us to investigate both of these predictions from the psychology literature.

Survey Administration

The sampling frame for the survey was defined as all households within the Herrin city limits. Households are chosen using a two-stage process. At the first stage, 250 household addresses are chosen at random from the phonebook. At the second stage, the dwelling unit located two dwelling units to the right of the initially chosen dwelling unit is designated to be interviewed.[13] This procedure ensured that houses without telephones have a chance of being included in the sample. At the household level, the interviewer enumerates all household members over 18 who have financial responsibility in the household in the sense that they "have or share responsibility for deciding the household budget and for paying housing, food and other expenses." Where multiple household members meet the financial responsibility criteria, the person to be interviewed is chosen according to a selection table. Interviewers received two days of training during which they conducted several practice interviews. There are 237 completed interviews out of 286 attempted interviews for a response rate of 83%. The survey instrument took a little over 30 minutes on average to administer.

Empirical Results

The basic empirical results from our study are presented in Table 2. (TABLE 2 ABOUT HERE). We provide five summary statistics for each of the six risk levels valued: (a) the percent giving a zero response, (b) median WTP, (c) mean WTP, (d) the 5% α-trimmed mean[14] (labeled

[13] In a few instances, our locational shift resulted in more than one dwelling unit meeting the shift qualification. This increased the number of sampled dwelling units from 250 to 286. Further details on the sampling plan and its execution can be found in (Mitchell and Carson, 1986).

"α[mean]" in Table 2), and (e) a "corrected" mean WTP (labeled "C[mean]" in Table 2) derived after dropping the 11 cases where the interviewer's evaluation clearly indicated that the respondent did not understand the scenario and/or it was clear that the respondent had given the same large WTP ($60 or greater) response to two or more of the risk reductions. Almost all of these cases are among those effectively dropped when we apply the α-trimmed mean procedure to the data.[15] All WTP amounts are in 1985 dollars.

A casual examination of the results in Table 2 suggests that the WTP estimates for the two risk reduction subsamples—A and B—generally increase as the magnitude of the risk reductions increases, although not monotonically so. The WTP amount for the largest risk reduction valued (1.33) by those receiving Treatment A is greater than the smallest risk valued by those receiving Treatment B (2.43). We consider this issue at more length in the construct validity section below.

A simple comparison of the lowest Treatment A and Treatment B risk values rejects the null hypothesis of no difference in WTP at the $p < .01$ level for both the corrected and uncorrected datasets. The same result occurs when the values of the middle and largest risks for Versions A and B are compared. In each case the lower absolute value for the risk improvement receives a lower WTP value at the .01 level. These results reject the metric bias hypothesis.

The test can be made more rigorous by controlling for the order in which the WTP questions were posed and for a possible interaction between treatment and order effects. Table 3 shows the results of the relevant analysis of variance (ANOVA) tests using the corrected data. These results show that there continues to be a consistent version effect (Treatment A versus Treatment B) across all three of the amounts asked, thereby rejecting the metric bias hypothesis at $p < .001$ in each case. In contrast, none of the order effect tests even begin to approach traditional levels of statistical significance. The tests involving an interaction effect between treatment version and order are never significant at the $p=.10$ level. Similar conclusions are drawn from ANOVA estimates based upon the uncorrected dataset.

---

[14] A 5% α-trimmed mean is calculated by first dropping the lowest and highest 5% of the observations and then calculating mean WTP based upon the remaining 90% of the observations. The median is a 50% trimmed mean.

[15] Seven percent of the sample provided the same positive WTP response for two or more of the risk levels they were asked about. Most of the respondents dropped in the corrected sample are elderly or have low educational levels.

Further, evidence in favor of rejecting the hypothesis of a lack of sensitivity to the absolute magnitude of the risk can be seen in the percentage of zero WTP responses received for the smallest and largest risks valued in the two treatments. For the smallest risk, 87% of respondents gave a zero WTP amount for the Treatment A risk of .04 annual deaths per thousand while 58% of respondents gave a zero WTP amount for the smallest Treatment B risk of 2.43 annual deaths per thousand, with the difference being significant at the $p < .01$ level. For the largest risk, 42% of the Treatment A respondents gave a zero WTP to 1.33 annual deaths per thousand while only 20% of the Treatment B respondents gave a zero WTP amount for the largest risk of 8.93 annual deaths per thousand, with the difference also being significant at the $p < .01$ level.

Construct Validity

We looked at construct validity in two different ways. The first compares our results with theoretical predictions; the second evaluates whether our WTP estimates are systematically related to variables that *a priori* one would expect to be predictive of the magnitude of the WTP amounts given.

The economic literature on risk (*e.g.*, Jones-Lee, 1974; 1979) poses two straightforward predictions.[16] The first is that WTP should increase with increases in the magnitude of the risk reduction ($\partial WTP/\partial\delta > 0$). The second is the rate of increase in WTP should be declining with increases in risk reductions ($\partial^2 WTP/\partial\delta^2 < 0$). Both of these predictions can be examined by examining the following regression models that were fit to the α-trimmed mean or corrected mean WTP amounts from Table 2.[17] There are six observations in these regression equations,

[16] There are two other standard predictions. The first of these has to do with WTP increasing with increases in income. We look at this prediction later in the paper. The second is that the WTP for a fixed risk reduction should increase with the level of baseline risk. Smith and Desvousges (1987) look at this prediction in the context of a CV survey and fail to find support for the hypothesis with respondents appearing to ignore the randomly assigned baseline risks. This may be the rational response from a respondent's perspective, and as such, this hypothesis may be difficult to test in a survey context.

[17] Many of the respondents dropped by these two approaches whose pattern of responses suggests a failure to distinguish between risk levels are elderly and less well educated. The δ-trimmed mean effectively drops almost all of the respondents who give the same positive WTP amount for all three risk levels as well as a subset of those providing $0 for all three risk levels. The process of adjusting the mean estimate by removing a small number of cases where the interviewers indicated substantial problems with understanding the question picks up a subset of these cases.

one from each treatment, following the common biometrics practice of fitting the dose response models to the relevant summary statistic from each treatment group.[18]

Starting with the α-trimmed mean estimates, which tend to drop out a group of respondents who gave the same WTP response, either low ($0) or high, for all three risk levels, one estimates the model:

$$\text{(1)}\quad \log(WTP) = \underset{(16.98)}{2.3520} + \underset{(8.08)}{.6298}*\log(\delta)\ ,$$

where t-statistics are in the parentheses and the adjusted R^2 is .928. This model fits significantly better than a linear model and avoids the problem in any linear model with a positive constant term of suggesting a positive WTP for a zero reduction in risk.

Because the installation of a particular filtration level will only provide one risk improvement, economic theory suggests strategic behavior may be optimal in which WTP for the largest risk will be over or understated depending upon the respondent's ideal risk-cost combination. We construct a position variable POS, which equals –1 for the smallest risk, 0 for the second largest risk, and 1 for the largest risk that a respondent was asked to value. This variable allows for symmetric deviations in the valuation model estimated in (1) based upon the relative magnitude of the particular risk in the set of risks the agent was asked to value. The estimated regression model including this variable is:

$$\text{(2)}\qquad \log(WTP) = \underset{(31.87)}{2.364} + \underset{(10.74)}{.5380}*\log(\delta) + \underset{(3.31)}{.3600}*POS\ ,$$

[18] Results from fitting models to the individual data produce qualitatively similar results but with, as should be expected, much lower R-squares. Some difficulties arise with the use of individual data in that the log of zero is undefined and usual correction approaches of taking the log of zero plus a small positive amount conceptually go against the possibility that some agents are at a corner solution. A standard Tobit model has the undesirable property of implying that some agents have negative latent WTP values. Some type of spike mixture model along the lines suggested by Werner (1999) which took account of the correlation structure between the three responses given by each agent would likely be appropriate if our interest was in fitting the individual data including covariates. Here, we are principally concerned with the aggregate value of a statistical life function and our approach is more transparent.

where the adjusted R^2 is now .979 and the estimate of σ from the regression is 0.181. This model represents a clear improvement over the model presented in equation (1) suggesting that agent WTP for a particular risk was influenced not only by the actual magnitude of the risk, δ, but also by its relative position in the set of three risk.

It is possible to allow for a non-symmetric effect with respect to POS by allowing the lowest risk valued and the highest risk valued to have different coefficients. The F-test for the sum of these two coefficients being zero is 0.668 (p=0.499). This test suggests that the hypothesis of symmetric response to POS cannot be rejected.

Turning now to the corrected mean WTP data from Table 2, the estimated model is:

$$\text{(3)} \quad \log(WTP) = \underset{(19.60)}{2.5508} + \underset{(6.27)}{.4593} * \log(\delta) \ ,$$

where the adjusted R^2 is .885. Again the log-log model fits substantially better than does a linear model.

The regression model fit for the corrected mean WTP data with the POS variable is:

$$\text{(4)} \quad \log(WTP) = \underset{(120.23)}{2.5635} + \underset{(25.23)}{.3630} * \log(\delta) + \underset{(12.10)}{.3778} * POS \ ,$$

where the adjusted R^2 is now .998 and the estimate of σ is 0.052. Again this model represents a clear improvement over its counterpart (3) without the POS variable. The F-test for allowing POS having a different low and high effect is 0.040 (p=.860).

We use equations (2) and (4) with POS set equal to zero in order to derive our WTP estimates and in turn to make our estimates of the statistical value of life (SVL). In using these equations it is necessary to add back in a function of σ to obtain consistent estimates under the assumption of normality of the error term (Goldberger, 1968). For equation (4), the appropriate formula is given by,

$$\text{(5)} \quad E(WTP) = EXP[2.5635 + .3778 * \log(\delta) + (.052)^2/2],$$

where 2.5635 and .3778 were the regression model coefficients and 0.052 was the estimate of σ. Consistent estimates from equation (2) can be obtained in a similar fashion. Figure 3 displays estimated WTP from equation (2) and (4) as a function of the magnitude of the risk reduction valued.

Our empirical results suggest a significant premium for the largest risk reduction the respondent is offered. After correcting for this premium associated with the largest offered risk reduction, it is not possible to reject that the hypothesis that the WTP amounts for all six of the risk reductions are drawn from a single smooth continuous underlying risk-WTP function where the log of WTP increases linearly in terms of the log of risk reduction. The log-log functional form is consistent with economic theory underlying the valuation of risk reductions. It is also interesting to note that the log-log functional form is commonly used in dose-response experiments. Further we note that a dose response relationship between WTP and the magnitude of risk reduction is a good analogy in terms of summarizing the results. This is done in a simple figure that traces out how the percent willing to pay a particular increase in their household water bill decreases as the size of the water bill increases.

The results from both equations (1) and (2) are consistent with the two predictions of economic theory that $\partial WTP/\partial\delta > 0$ and $\partial^2 WTP/\partial\delta^2 < 0$. These results are robust to using either the raw mean WTP estimates or the α-trimmed mean estimates as the dependent variable. Support for the $\partial WTP/\partial\delta > 0$ proposition is not sensitive to the functional form we used, while support for the $\partial^2 WTP/\partial\delta^2 < 0$ proposition, which concerns the curvature of the function, comes from comparing the fit of a linear specification with that of the log-log specification in equations (1) and (2). The linear specification clearly provides an inferior fit to the data. A quadratic model in δ provides a similar fit to the log-log model and also suggests acceptance of the declining marginal utility ($\partial^2 WTP/\partial\delta^2 < 0$) hypothesis. This declining marginal utility for risk reductions may have strong policy implications. We discuss the implications of these findings in a later section of this paper.

Our empirical results are all consistent with several other tests that economic theory suggests concerning how WTP amounts should change with changes in risk levels. Construct validity is examined by regressing log(WTP) or the probability of a non-zero WTP response on several covariates that *a priori* should have particular signs. After controlling for the size of the risk reduction, we find that household size is significantly related to the probability of a non-zero

WTP response. Conditional on giving a positive WTP response, we find that the log of income is a highly significant ($p < .001$) predictor of log(WTP) as was a rating (prior to the main CV scenario) of the harm done from chemical contaminants in Herrin's drinking water. Respondents over 55 are willing to pay substantially less than those under 55, where the 55-year-old threshold was chosen in reference to the 20-30 year latency period of cancer from THM exposure. Household size had a significant positive effect on the probability of providing a non-zero WTP amount.

Value of a Statistical Life

The expectation of the statistical value of life (SVL) can be calculated by $E(SVL[\delta^*])=(100{,}000/\delta^*)(WTP[\delta^*]/2.86)$, where δ^* is the risk reduction of interest, $1/\delta^*$ provides the appropriate scale factor aggregation factor when multiplied by 100,000, $WTP[\delta^*]$ is the predicted household WTP for δ^*, and 2.86 is the estimate of the average household size from our survey. In Table 4 we present our estimates for the six risk reductions for both the α-trimmed mean estimates from equation (2) and the corrected mean equation (4).[19] Figure 4 is a graph of the estimated SVL functions using both equations (2) and (4) over the entire range of risk reductions considered, 0.04 [in 100,000] to 8.93. Figure 5 graphs the function after dropping the values of the smallest and largest risk reductions. This makes it possible to observe more details about the shape of the WTP function over its central range

From Table 4, someone familiar with the SVL literature (*e.g.*, Cropper and Freeman, 1991) will immediately note that our estimates are on the low side of the range of SVL estimates in the literature. Almost all of the other estimates, however, are obtained by looking at current accident rates. That our estimates should be lower than most of the other SVL estimates in the literature is consistent with economic theory's predictions for respondents who discount their WTP for the risk reduction due to the risk's 20 to 30 year latency period.

After we discount our statistical life value estimates using a range of (exponential) discount rates typically used in the literature, our statistical life value estimates are well within the range commonly found in the literature for WTP to avoid current period fatal accidents. Figure 6

---

[19] Note that these estimates are conservative in the sense that they set POS=0. Using POS=1 might be a better alternative in terms of making a correction for strategic behavior if one believes that open-ended CV WTP estimates are biased downward.

displays the implied value for a current statistical life for a 25-year latency period and discount rates ranging from 0% to the common consumer credit card rate of 18%.

Discussion

THMs represent the quintessential type of environmental risk that confront government regulators. They involve a very low-level risk with a long latency period that is imposed upon the public as a side effect of a government action to provide a public good, in this case the reduction of harmful biological contaminants by chlorinating drinking water. Any MCL short of totally eliminating THMs from drinking water is not completely safe. There are clear economies of scale in installing equipment to bring THMs to the MCL that lead to pressure to exempt small producers. A non-federal agency must implement the technical remedy. In this particular case, our results suggest that installing a public provided activated carbon filtration systems whose only benefit is to reduce THM levels to the standard is not likely to be welfare improving for Herrin.

Our CV survey appears to have worked well in this situation in the sense that: (a) most people were able to answer the WTP questions, (b) the amounts provided generally seem reasonable, if not on the low side, given the existing literature on SVL, (c) all predictions from economic theory with respect to risk (which our study was designed to test) were confirmed, and (d) the WTP amounts are systematically related to other factors such as income, age, and household size in the direction that one would expect.

Problematic with both the open-ended format used in this study and the asking of multiple questions are the incentives for strategic behavior. While the first cannot be tested for here, our empirical results, with respect to the very significant effect of the position variable (POS) in equations (2) and (4), are consistent with such behavior. We have attempted to undo the strategic behavior by setting the POS variable to equal zero in calculating WTP estimates for different risks levels. This is the conservative choice relative to setting POS equal to one, which might be justified if one thought that all of the open-ended WTP responses are biased downward, which often appears to be the case. Interestingly, while we find strong support for the presence of strategic behavior, which is predicted by economic theory, we find no evidence in support of the order effects predicted by some of the psychological literature on framing effects.

Even though we have strongly rejected the null hypothesis put forth that agents in CV studies are insensitive to the quantitative magnitudes of the good that they are asked to value, an issue that naturally arises is whether the magnitude of the differences in our WTP estimates are correct. Over small enough changes in risk, one would expect the change in WTP to increase at almost a linear rate in terms of increases in risk reductions (Machina, 1995). According to expected utility theory, if we assume that income effects are sufficiently small and available income sufficiently large, we would expect this approximate linearity to hold over a fairly large range (Weinstein, Shepard, and Pliskin, 1980; Hammitt and Graham, 1999).

Compared with most studies of risk values, our set of risk reductions both cover a much wider range—almost three orders of magnitude (0.04×10^{-5} to 8.93×10^{-5}) and value smaller risks than do most studies. For the larger risk reductions we examine our results show an approximately linear relationship between WTP and the size of the risk reductions. For the smaller risk reductions, however, marginal WTP *declines* considerably over the range of risk reductions we examined.[20] There is an important policy implication of this finding—for the same number of statistical lives saved, the benefits of government actions to reduce very small risks over very large numbers of people are considerably greater than the benefits of actions that reduce large risks over a much smaller number of people.

There is another factor, however, which works in a different direction. As the size of the risk reduction being valued decreases, the number of respondents who give the corner solution of a zero WTP increases. This can be formally examined by stacking the observations in Table 2 and then fitting a logistic regression model to the probability of providing a zero WTP response as a function of the risk reduction asked about. Again, as we saw earlier using log(δ) provides a substantially better fit than δ. The fit is also improved by adding the POS variable, which suggests that some of the strategic behavior is coming through saying "no" to risk reduction levels other than the highest level offered. The results from this model are:

$$\text{(7)}\quad \text{logit(prob. of \$0 WTP)} = .2047 - .4358^{*}\log(\delta) - .4717^{*}\text{POS},$$
$$\qquad\qquad (2.38)\quad (-7.05)\qquad\quad (-4.13)$$

---

[20] We are not the first to observe this effect. Blomquist (1982) appears to have been the first researcher to note that the studies (mostly using hedonic pricing) valuing smaller risk changes tend to have higher SVL estimates than

where the (McFadden) pseudo R-square is 0.179. Figure 7 shows a graph of this relationship and suggests that a THM risk reduction would have to be larger than 1 in 100,000 in order to gain majority approval. This suggests the possibility of considerable divergence between political support for a program to reduce very low-level risks and the economic value of such a plan. The lower the risk level, the lower the percentage of people who would vote for a further reduction and the higher the expected value of each saved life.

Concluding Remarks

There is abundant evidence that people have difficulties with making decisions about low-level risks in their daily lives. The task of the CV survey designer would be much easier if the objective were simply to predict how the public would vote on a policy issue if offered the opportunity. In that case, the survey designer would only need to incorporate into the wording of the survey instrument the likely information set the voters will have available to them when they vote. The key difficulty in using CV to estimate WTP for risk reductions is not describing the reduction in the CV instrument but ensuring that respondents actually comprehend the size of the risk reduction by an effective method of communicating this information. The results of this paper suggests that it is possible to effectively communicate risk levels to respondents even at low risk levels but replicating our success will likely be neither quick nor inexpensive. A careful program of pretesting would be necessary along with use of the expensive in-person survey mode.

There has been surprisingly little work on the properties of different communication devices in the context of a CV survey (*e.g.*, Loomis and duVair, 1993).[21] Much of the focus of this paper has been on the development of a risk communication device for a low-level environmental risk with a long latency period. Our risk ladder with its cigarette equivalence scale combines properties of a Smith and Desvousges (1987) type risk ladder with risk communications devices that put the risk reduction in the context of a meaningful risk scale. It would be interesting to see

---

those valuing larger risk changes. This empirical regularity has continued to hold in both revealed preference and CV studies.

[21] There has been some work in the risk communications literature (*e.g.*, Sandman, Weinstein, and Miller, 1994), on trying to find risk communication devices that adequately conveyed changes in small risks, although again less than we would have expected.

how our hybrid risk ladder fares compared to other risk communication devices given its apparent success here. One encouraging recent paper by Curso, Hammitt and Graham (2000) compares three different risk communication devices in the context of a CV survey using risks that are larger and fall in a more narrow range than those used in this study. With the first risk communication device, WTP estimates were effectively insensitive to the magnitude of the risk. The WTP estimates for the second device showed some sensitivity, and the WTP estimates for the third increased almost linearly as the magnitude of the risk reduction increased. This third device should be tested with lower level risks like the ones used in this study.

Additional research is also needed to determine how people discount risks over different time horizons as a wide range of estimates is obtained by assuming a wide range of different plausible discount rates. Horowitz and Carson (1990) provide a method for examining this issue. Cropper and Portney (1990) use it look at a number of important policy issues involving discounting and find indications of non-exponential discounting over very long time horizons. In the specific context of this study it would be useful to know whether exponential discounting is a reasonable approximation over the 20-30 year latency period used and what a reasonable approximation to the discount rate for long term cancer risks is.

We suspect the reality is that people value different types of risks differently (Horowitz and Carson, 1991), that risk reductions of different magnitudes for the same risk suggests different SVL numbers, and that people do not use simple exponential discounting when they value risks with long latency periods. While such results would not be inconsistent with economic theory, they substantially complicate the situation for a policy analyst who is trying to determine whether the net benefits of a government action to help reduce risks are positive. The always-harried analyst would like to have a single SVL number and discount rate to use in an exponential manner. The empirical question is how much such a simple view of the world diverges from how people actually value risks. If divergences are large, can they be quantified in such a way that some easy to use function can be develop to replace a single SVL number and discount rate?

References

Arrow, K., R. Solow, P.R. Portney, E.E. Leamer, R. Radner, and H. Schuman (1993), "Report of the NOAA Panel on Contingent Valuation." *Federal Register*, vol. 58, pp. 4601-4614.

Attias, L. A. Contu, A. Loizzo, and M. Massiglia (1995), “Trihalomethanes in Drinking Water and Cancer—Risk Assessment and Integrated Evaluation of Available Data in Animals and Humans,” *Science of the Total Environment*, vol. 171, pp. 61-68.

Baron, Jonathan and Joshu Greene (1996), “Determinants of Insensitivity to Quantity in Valuation of Public Goods: Contribution, Warm Glow, Budget Constraints, Availability, and Prominence,” *Journal of Experimental Psychology: Applied*, vol. 2, pp. 107-125.

Beattie, J., J. Covey, P. Dolan, L. Hopkins, M. Jones-Lee, N. Pidgeon, A. Robinson, and A. Spencer (1998), “On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part I - Caveat Investigator,” *Journal of Risk and Uncertainty*, vol. 17, pp. 5-25.

Blomquist, G. (1982), “Estimating the Value of Life and Safety: Recent Developments,” in *The Value of Life and Safety*, edited by M.W. Jones-Lee, ed., Amsterdam, North-Holland.

Carson, R.T. (1997), "Contingent Valuation Surveys and Tests of Insensitivity to Scope." in *Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods*, edited by R.J. Kopp, W. Pommerhene, and N. Schwartz. Boston, Kluwer.

Carson, R.T., T. Groves and M. Machina (1999), “Incentive and Informational Properties of Preferences Questions,” Plenary Address, European Association of Environmental and Resource Economists, Compatibility Issues in Stated Preference Surveys,” Oslo Norway, June.

Culp, Gordon, ed. (1984), *Trihalomethane Reduction in Drinking Water*. Park Ridge, NJ: Noyes Publications.

Corso, P.S., J.K. Hammitt, and J.D. Graham (2000), "Evaluating the Effects of Visual Aids on Willingness to Pay for Reductions in Mortality Risks," paper presented at the annual meeting of the Association of Environmental and Resource Economists, January.

Cropper, Maureen L. and A. Myrick Freeman (1991), "Environmental Health Effects," in *Measuring the Demand for Environmental Quality* edited by John B. Braden and Charles D. Kolstad, Amsterdam, North-Holland.

Cropper, Maureen L. and Paul R. Portney (1990), "Discounting and the Evaluation of Lifesaving Programs," *Journal of Risk and Uncertainty*, vol. 3, pp. 369-379.

Davies, Clarence, Vincent T. Covello, and Frederick W. Allen, eds. (1986), *Risk Communication: Proceedings of the National Conference on Risk Communications*, Washington, DC, Conservation Foundation.

Fetherstonhaugh, D. P. Slovic, S.M. Johnson, and J. Friedrich (1997), "Insensitivity to the Value of Human Life: A Study of Psychophysical Numbing," *Journal of Risk and Uncertainty*, vol. 14, pp. 283-300.

Fischhoff, Baruch (1990), "Understanding Long Term Environmental Risks," *Journal of Risk and Uncertainty*, vol. 3, pp. 315-30.

Fisher, Ann, Maria Pavolva, and Vincent Covello, eds. (1991), *Evaluation and Effective Risk Communications Workshop Proceedings*, Washington, DC, U.S. Environmental Protection Agency.

Goldberger, A.A. (1968), "The Interpretation and Estimation of Cobb-Douglas Functions," *Econometrica*, vol. 36, pp. 464-472.

Hammitt, J.K. and J.D. Graham (1999), “Willingness to Pay for Health Protection: Inadequate Sensitivity to Probability?” *Journal of Risk and Uncertainty,* vol. 18, pp. 33-62.

Hoehn, John P. and Alan Randall (1987), “A Satisfactory Benefit Cost Indicator from Contingent Valuation,” *Journal of Environmental Economics and Management*, vol. 14, pp. 226-247.

Horowitz, J.K. and R.T. Carson (1991), “A Classification Tree for Predicting Consumer Preferences for Risk Reduction,” *American Journal of Agricultural Economics*, vol. 73, pp. 1416-1421.

Horowitz, J.K and R.T. Carson (1990), “Discounting Statistical Lives,” *Journal of Risk and Uncertainty*, vol. 3, pp. 403-413.

Jones-Lee, Michael W. (1976) *The Value of Life: An Economic Analysis*. Chicago, University of Chicago Press.

Kahneman, D. and J.L. Knetsch (1992), "Valuing Public Goods: The Purchase of Moral Satisfaction." *Journal of Environmental Economics and Management,* vol. 22, pp. 57-70.

Loomis, J.B. and P.H. duVair (1993), “Evaluating the Effect of Alternative Risk Communication Devices on Willingness to Pay: Results from a Dichotomous Choice Contingent Valuation Experiment,” *Land Economics*, vol. 69, pp. 287-298.

Machina, M.J. (1995), “Non-Expected Utility and the Robustness of the Classical Insurance Paradigm,” *Geneva Papers on Risk and Insurance Theory*, vol. 20, pp. 9-50.

Mitchell, Robert Cameron and Richard T. Carson (1986), "Valuing Drinking Water Risk Reductions Using the Contingent Valuation Method: A Methodological Study of Risks From THM and Giardia," report to the U.S. Environmental Protection Agency, May.

Mitchell, R.C. and R.T. Carson (1989), *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Baltimore: John Hopkins University Press.

Moore, Don A. (1999), “Order Effects in Preference Judgments: Evidence for Context Dependence in the Generation of Preferences,” *Organizational Behavior and Human Decision Processes*, vol. 78, pp. 146-165.

Rimer, Barbara and Paul Van Nevel, eds. (1999), *Cancer Risk Communications: What We Know and What We Need to Learn*, Bethesda, MD, National Cancer Institute.

Smith, V. Kerry and William H. Desvousges (1987), “An Empirical Analysis of the Economic Value of Risk Changes,” *Journal of Political Economy*, vol. 95, pp. 89-114.

Weinstein, M.C, D.S. Shepard, and J.S. Pliskin (1980), “The Economic Value of Changing Mortality Probabilities: A Decision-Theoretic Approach,” *Quarterly Journal of Economics*, vol. 44, pp. 373-393.

Werner, M. (1999) “Allowing for Zeros in Dichotomous-Choice Contingent-Valuation Models,” *Journal of Business and Economic Statistics*, vol. 17, pp. 479-486.

FIGURE 1a: Risk Ladder (Upper Part)

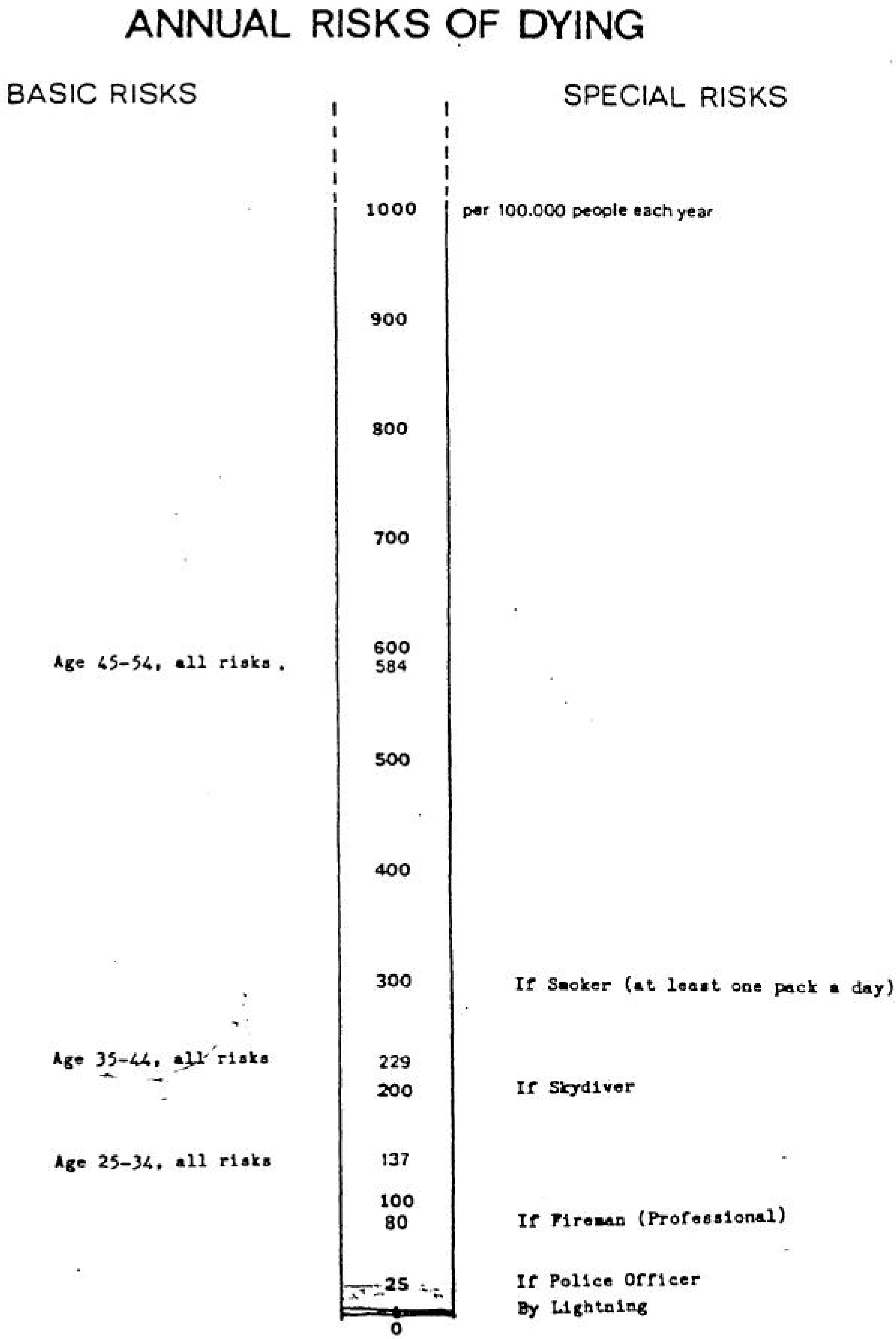

FIGURE 1b: Risk Ladder (Lower Part)

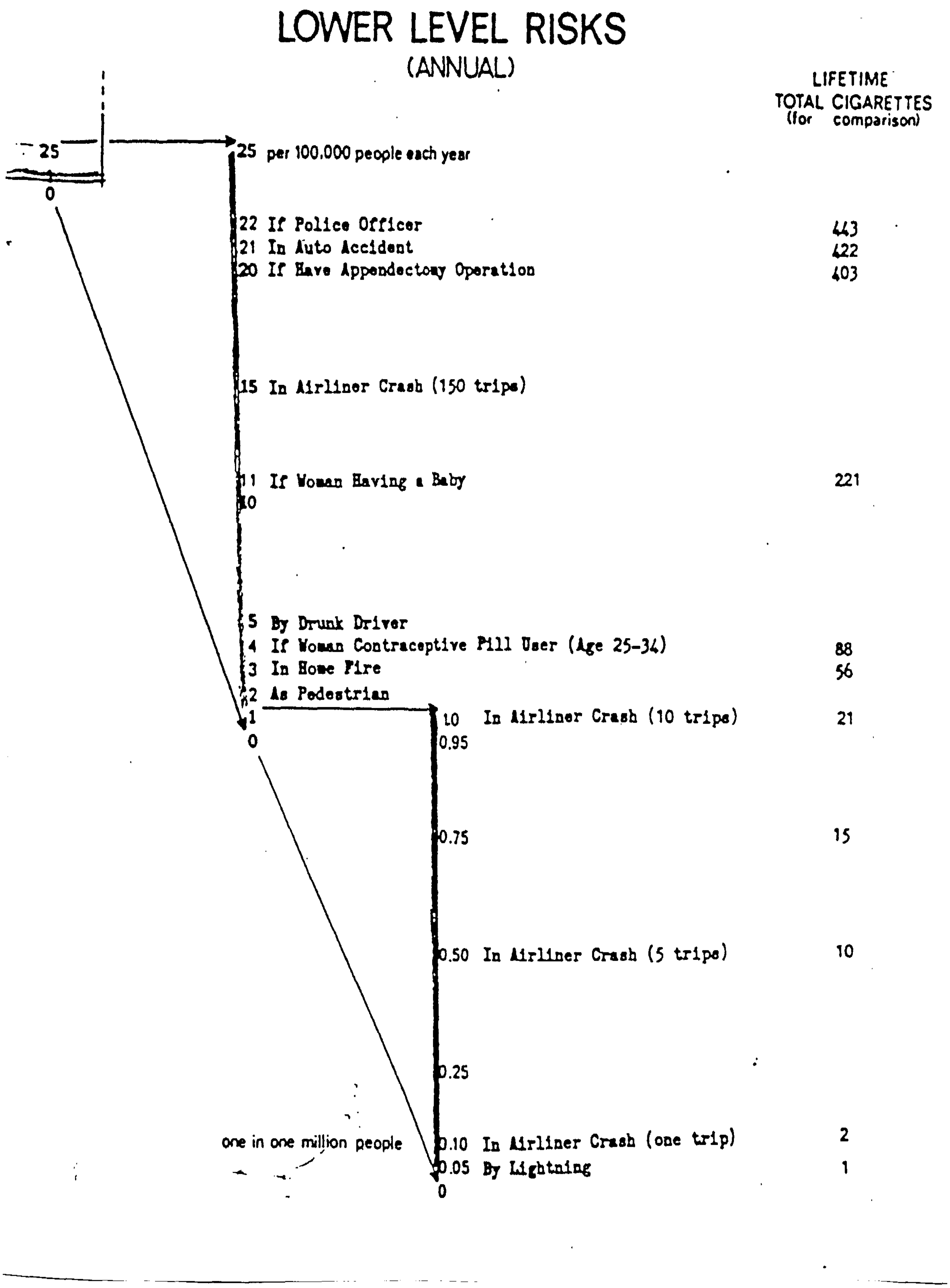

FIGURE 2: Card B-11

B to A

REFERENDUM PASSES	**From**		**To**	**CHANGE IF**
Level of THMs ppm	.55	→	.10	.45 parts per million
General Risk of Dying per 100,000	3.0	→	.57	2.43 per 100,000
General Risk Equivalent in Total Lifetime Cigarettes	65	→	11	54

Conditions:

1. THMs only source of contamination.
2. Reduced only to EPA standard.
3. Authorities say risk level not high enough to worry about.

TABLE 1: Experimental Design For Tests of Metric Bias and Question Order Effects

Valuation Question Order	Smaller Risk Changes	Larger Risk Changes
1-2-3	Version A1 .04/ .43/ 1.33	Version B1 2.43/ 4.43/ 8.93
2-1-3	Version A2 .04/ .43/ 1.33	Version B2 4.43/ 2.43/ 8.93

TABLE 2: Household WTP Higher Water Bills For THM Risk Reductions

THM Risk Improvement Reductions (ppm)	Annual Deaths per 100,000	Version A (N=121)				
Version/From/To/Change		Percent Zero	Median	Mean	α[Mean]	C[Mean]
A1: .11/.10/.01	(.04)	87%	$0	$3.78	$1.13	$2.86
				(±$2.76)	(±$1.41)*	(±$1.82)
A2: .18/.10/.08	(.43)	66	0	11.37	8.30	9.19
				(±4.33)	(±3.72)	(±3.37)
A3: .33/.10/.23	(1.33)	42	17	23.73	18.99	20.49
				(±7.37)	(±6.35)	(±5.20)
				Version B (N=117)		
B1: .55/.10/.45	(2.43)	58%	0	15.23	12.70	11.79
				(±4.64)	(±4.25)	(±3.38)
B2: .90/.10/.80	(4.43)	39	20	26.25	23.08	23.51
				(±8.99)	(±5.78)	(±5.39)
B3: 1.65/.10/1.55	(8.93)	20	36	44.27	42.32	42.68
				(±7.22	(±7.98)	(±7.32)

*Ninety-five percent confidence intervals in parentheses.

TABLE 3: Tests for Metric Bias and Question Order Effects

Metric Test:

Ho: Ordinal Ranking—Amount A_i = Amount B_i , i = 1,2,3

H1: Cardinal Ranking/Scope Sensitivity—Amount A_i < Amount B_i , i = 1,2,3

Order Effect Test:

Ho: No Question Order Bias—Order 1,2,3 = Order 2,1,3

H1: Question Order Bias—Order 1,2,3 ≠ Order 2,1,3

Interaction Test: Ho: No Interaction Effect

A_i - B_i for Order 1,2,3 = A_i - B_i for Order 2,1,3

H1: Interaction Effect

A_i - B_i for Order 1,2,3 = A_i - B_i for Order 2,1,3

<u>Analysis of Variance Results (N=230)</u>

Amount 1	VERAB	F= 21.93	P < .0001
	ORDER	F = 0.09	P =-.7641
	ORDER*VERAB	F = 2.29	P = .1319
Amount 2	VERAB	F =19.90	P < .0001
	ORDER	F = 0.41	P = .5243
	ORDER*VERAB	F = 2.05	P = .1537
Amount 3	VERAB	F =20.74	P < .0001
	ORDER	F = 0.04	P = .8478
	ORDER*VERAB	F = 0.48	P = .4886

VERAB = Treatments A and B

ORDER = Treatments 1 (1,2,3) and 2 (2,1,3)

ORDER*VERAB = Interaction between treatments A and B and 1 and 2

FIGURE 3: Household WTP
For .04 to 8.93 in a 100,000 Risk Reductions

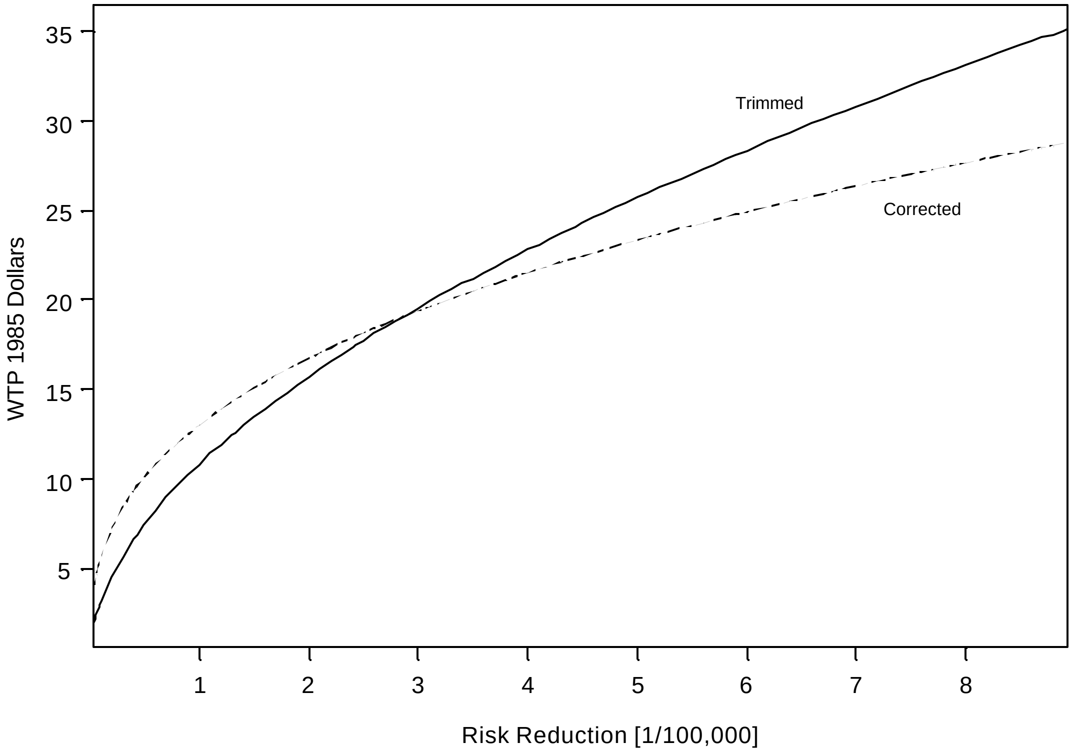

TABLE 4: Value of a Statistical Life in Millions of 1985 Dollars

Risk Reduction (δ)	Equation (2) (α-Trimmed Means)	Equation (4) Corrected Mean
0.04	1.672	3.527
0.43	0.558	0.777
1.33	0.331	0.378
2.43	0.251	0.258
4.43	0.190	0.176
8.93	0.137	0.113

FIGURE 4: Value of a Statistical Life
For .04 to 8.93 in a 100,000 Risk Reductions

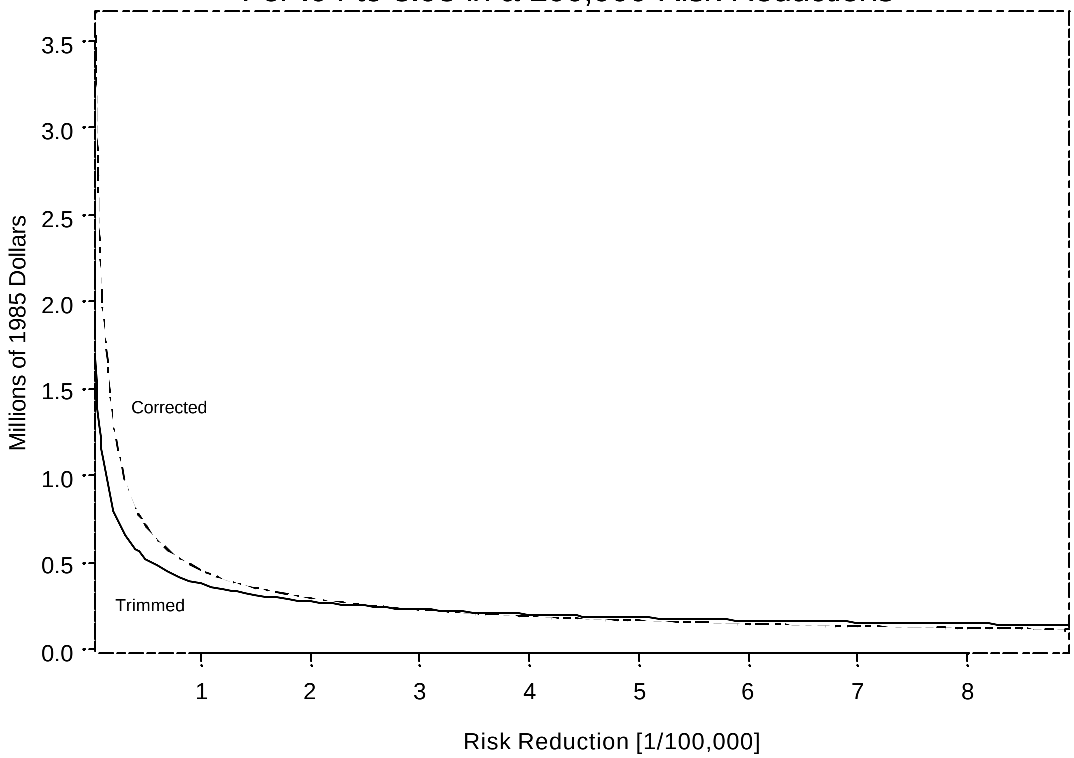

FIGURE 5: Value of a Statistical Life
For .43 to 4.43 in a 100,000 Risk Reductions

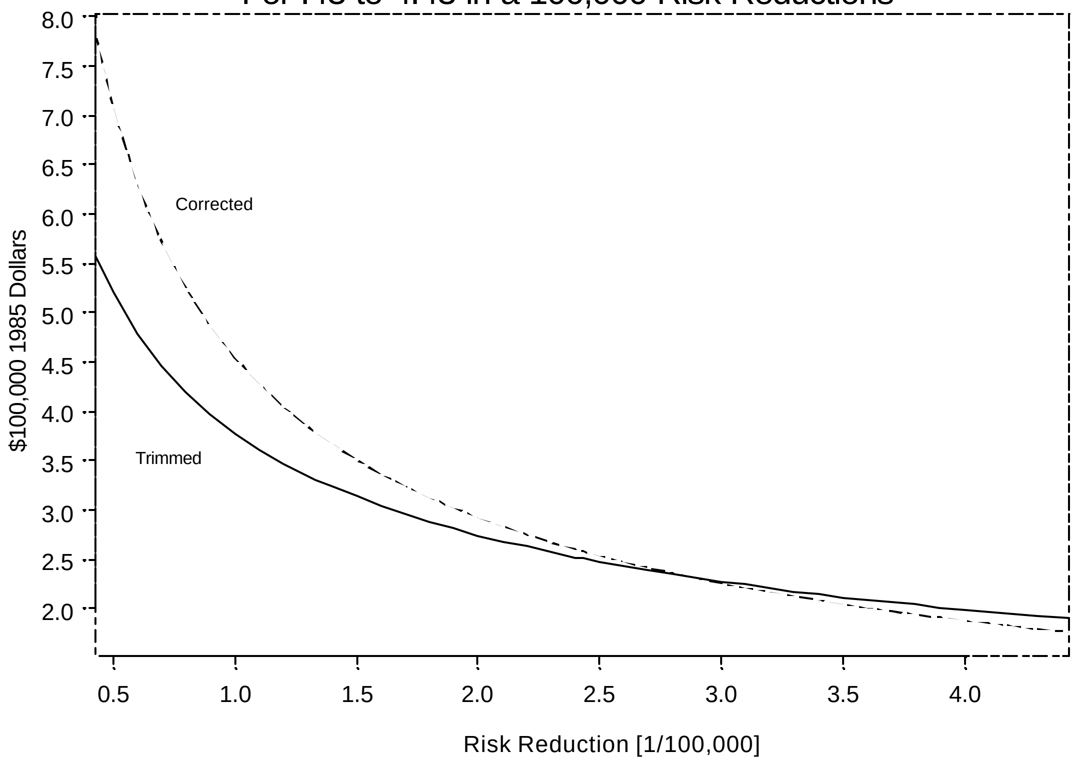

FIGURE 6: Implied Values of a Current Statistical Life
Assuming a 25 Year Latency and a 1 in 100,000 Risk Reduction

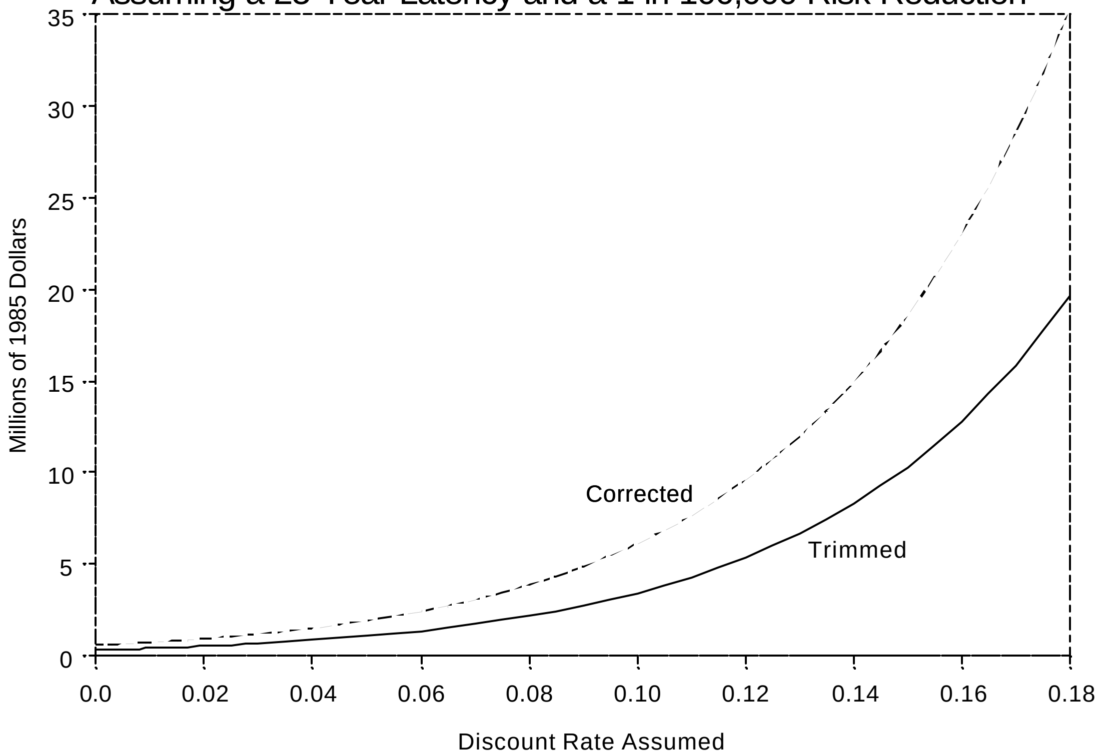

FIGURE 7: Opposition to Program by Magnitude of Risk Reduction