Morbidity and Mortality: How Do We Value the Risk of Illness and Death?

PROCEEDINGS OF SESSION IV: VALUING MORBIDITY AND MORTALITY: PESTICIDES AND TOXICS

A WORKSHOP SPONSORED BY THE U.S. ENVIRONMENTAL PROTECTION AGENCY'S NATIONAL CENTER FOR ENVIRONMENTAL ECONOMICS AND NATIONAL CENTER FOR ENVIRONMENTAL RESEARCH

April 10 – 12, 2006

National Transportation Safety Board
Washington, DC 20594

Prepared by Alpha-Gamma Technologies, Inc.
4700 Falls of Neuse Road, Suite 350, Raleigh, NC 27609

ACKNOWLEDGEMENTS

This report has been prepared by Alpha-Gamma Technologies, Inc. with funding from the National Center for Environmental Economics (NCEE). Alpha-Gamma wishes to thank NCEE's Maggie Miller and the Project Officer, Cheryl R. Brown, for their guidance and assistance throughout this project.

Table of Contents

Integrating Economic and Behavioral Questions into National Health Surveys

US EPA NCER/NCEE Workshop

11 April 2006

National Human and Nutrition Examination Survey (NHANES)

- Nationally representative, continuous, longitudinal study
- Collects data on demographics, health status, diet and chemical exposure
- 2-year cycle, 5000 participants per year
- Oversampling of special populations
- Value: individual data can be linked to objective, scientific health measures for hypothesis testing

Opportunity and Value-Added

- Improved estimates of the value of reduced morbidity and mortality to justify environmental health policies
- Better understanding of how people assess, perceive and respond to risk
- Improved analyses of the accountability of regulatory decision-making
- Design of targeted policies to minimize risk or support risk-averting behavior

NCER Proposal –
Costs of asthma medication

- Costs of asthma medication:
  - You have already said that you use one, or more, of these medications to control your asthma. How many months, out of the past three months, did you need to take this medication everyday or almost everyday?

NCER Proposal –
Behavioral response to poor air quality

Some people change their activities on days when air pollution is bad, while others go ahead with their activities as planned. On days in the past year when you thought or were informed air quality was bad, did you ever do anything differently, provided you had the choice, such as:

- Restrict the amount of your time outside?
- Exercise indoors instead of outside?
- Choose less strenuous activities?
- Cancel activities?
- Avoid areas with heavy traffic?
- Take medication?
- Close windows of your house?
- Stay indoors?
- Did nothing differently

National Children's Study

- Longitudinal, 21yr follow up study of 100,000 children from birth to adulthood
- Data to be collected on physical, biological and psychosocial environments, as well as exposure to chemicals
- Hypotheses frame the study survey
- Adjunct studies can be proposed
- Funding for FY07 uncertain

Next Steps

- RFA on analysis of NHANES data
- Reproductive/developmental outcomes, air pollution and CVD, drinking water contaminants and GI illness
- Other national health surveys

Use of Contingent Valuation to Elicit Willingness-to-Pay for the Benefits of Developmental Health Risk Reductions

Katherine von Stackelberg
Center for Risk Analysis, Harvard University
kvon@hsph.harvard.edu

and

James K. Hammitt
Center for Risk Analysis, Harvard University
IDEI and LERNA-INRA, Université de Toulouse
jkh@harvard.edu

March 2006

Abstract

We report several contingent valuation surveys to elicit willingness-to-pay for risk reductions associated with decreases in exposure to a chemical, PCBs, in the environment. We also develop Quality Adjusted Life Years (QALYs) from the survey using either standard gamble or time-tradeoff elicitation methods to explore the relationship between QALYs and willingness to pay (WTP), and to develop QALY weights for subtle developmental effects. The results of the contingent valuation are designed for incorporation into an integrated risk model to demonstrate the economic impact of risk reductions. Respondents showed a positive and proportional relationship between decreasing the risk of a 6-point reduction in IQ and WTP. Socioeconomic variables were not statistically significant predictors of WTP, while behavioral variables were strongly predictive and statistically significant. The range of mortality risks that respondents would accept on behalf of their (hypothetical) 10-year-old child is 2 in 10,000 to 9 in 1,000 per IQ point, and WTP per IQ point is $466 (95% confidence interval = $380, $520). QALY weights elicited via time tradeoff (reduction in life expectancy) were statistically significantly different from QALY weights elicited via a standard gamble. Respondents who answered questions about ecological endpoints first were willing to pay a small additional amount when asked about human health effects, but those respondents who answered questions about human health endpoints first were not willing to pay any additional amount when subsequently asked about ecological effects. WTP models demonstrate the importance of obtaining behavioral and cognitive information from respondents when eliciting WTP and in tests of sensitivity to scope.

1. Introduction

Potential health effects resulting from exposure to environmental chemicals can range from severe terminal illnesses such as cancer to milder, systemic illnesses. One category of effects that is receiving increased attention includes developmental and reproductive effects, such as reduced fertility, low birth weight, genetic defects, and cognitive deficits. The policy implications of these exposures have yet to be realized, in part because the relationship between exposure and effects is not well quantified, and in part because there is a dearth of data and information with which to quantify the benefits of risk reductions associated with exposure to chemicals that exert these kinds of effects. One such chemical, polychlorinated biphenyls or PCBs, contribute to the existence of fish consumption advisories in virtually every state, indicating that this exposure has important implications for public health. Other contaminants, such as mercury and lead, also pose developmental risks.

Cannon *et al*. (1996) conducted a scoping study to evaluate the literature and data available with which to quantify the value society places on avoiding potential effects from *in utero* exposures to chemicals. Their primary finding was that there are very few existing studies with which to quantify the monetary (or other valuation metric) of these effects. Cost of illness techniques can be used to quantify the impacts of some birth defects, but these would be restricted to fairly severe outcomes requiring ongoing treatment and attention. For other, more subtle effects, such as mild cognitive deficits, cost of illness and other related techniques are inadequate for capturing the range of costs and for estimating welfare measures. In addition, the authors acknowledge that existing cost of illness analyses related to the costs associated specifically with low birth weight (a very nonspecific effect in terms of the relationship between exposure and outcome) do

not reflect the total costs associated with the occurrence of these endpoints (Cannon *et al*., 1996).

Stated preference methods have been used frequently for the evaluation of risk reductions related to mortality (Hammitt and Liu, 2004; Hammitt and Graham, 1999) to obtain estimates of the value of a statistical life (Alberini, 2005), and increasingly also to value morbidity endpoints (Dickie and Gerking, 2002; Van Houtven *et al*., 2003, 2004; Krupnick, 2004). Fewer studies have evaluated potential morbidity effects for risks and exposures to children, which generally must be evaluated by parents (Dockins *et al*., 2002). While imperfect, these methods provide policy makers with information on how the general public might trade-off income against reductions in the risk of specific health effects. The results of the surveys presented here contribute to the growing literature on the relationship between WTP and reductions in risk of mild developmental delays.

2. Survey Design and Development

The surveys were designed over a one-year period and involved several informal pilot surveys, focus groups, and a pretest. From the onset, the surveys were designed to be administered over the Internet using a professional survey firm, Knowledge Networks. The research goal was to evaluate whether a CV might provide a feasible method for obtaining economic values for endpoints consistent with how they are expressed in a typical risk assessment framework (drawing from the experience of the lead author at an actual Superfund site) and explore how people respond to questions regarding potential effects to children and wildlife as a result of exposure to a specific chemical in the environment. To that end, there were numerous open-ended questions for which respondents were invited to provide comments as they progressed through the surveys.

These open-ended responses provide important insights into respondent motivations and thinking short of actually sitting with the respondent.

The primary objective of the surveys was to elicit an approximation of the monetized loss in utility consistent with economic theory experienced by respondents resulting from potential effects associated with exposure to PCBs. Another objective of the surveys was to measure WTP for risk reductions, consistent with the results that risk assessments generate. The surveys were designed so that members of the general public could follow and understand the issues, and the surveys asked various questions throughout to gauge what respondents already knew (or thought they knew) concerning chemicals in the environment and how they felt, in a general sense, about exposure to chemicals (*e.g.*, whether they thought it was a serious issue, or even feasible that the kinds of effects described in the survey could really occur). The surveys are based on a generic, non-specific site (although there are numerous actual PCB-contaminated freshwater systems across the United States and it is likely that there is at least one system in the general area in which the respondent lives); nonetheless, the surveys were designed to be plausible and the payment vehicle realistic and believable.

Respondents to the survey are first told that government officials in their State are responsible for allocating resources and are interested in individual opinions to inform potential policies. The first question asks respondents to rate the importance of several issues, including reducing crime, cleaning up the environment, improving education, reducing taxes, protecting State waterways, improving library services, reducing air pollution, and providing additional security at public events. The second question asks respondents to consider whether current State budget allocations should be reduced or

increased, keeping in mind that overall expenditures cannot be increased without an increase in revenue. Respondents are reminded that State policy makers are responsible for allocating resources, and that people may feel differently about these allocations depending on their own beliefs and knowledge. Respondents are informed that State policy makers are interested in learning how taxpayers feel about specific issues.

The survey then proceeds to set up the specific valuation question, which involves the potential effects of a specific chemical (PCBs – we ask "have you ever heard of PCBs?") in a large, unnamed freshwater system in the state in which the respondent resides. This system is contaminated, and the company or companies ostensibly responsible went out of business some years ago. Therefore, the State is contemplating setting up a special "cleanup" fund to be funded through a one-time increase in the State income tax.

We chose a payment vehicle that calls for a one-time increase in the State income tax, to be kept in a fund earmarked for a cleanup remedy for the (unnamed) freshwater system. The question states that the risk will decrease if the cleanup is conducted if the income tax is raised by the bid amount for all, not just for the respondent (Johansson-Stenman, 1998), which has been shown to generate values consistent with economic theory. However, not all States have an income tax, and this was not explicitly acknowledged. Another format might be to specify an increase in the property or local tax for those States without an income tax; however, for the sake of consistency across all respondents, we chose the income tax payment vehicle. The cleanup is described as occurring over several years, and the survey also states that even after cleanup is complete, it will still take several years for the wildlife receptors to recover. In addition,

the risks will never go to zero. Respondents are presented with an initial bid randomized from a bid vector ranging from $25 to $400. If the respondents agree to the initial bid, they are presented with a bid that is double the first bid (if they agree to $400 initially, then they are asked if they would be willing to pay at $800). If respondents do not agree to the initial bid, then they are presented with a bid that is half as much ($10 if they did not agree to $25 initially).

A particular issue that arises with double-bounded CV estimates from the literature is a failure to achieve consistency (Hanemann, 1991; Hanemann and Kaninnen, 2001; McFadden and Leonard, 1993). We used a double-bounded dichotomous choice (Hanemann, Loomis and Kanninen, 1991) which has been shown to substantially increase the statistical power of the WTP estimate, at the expense of a downward bias in the estimate because the second response is not incentive-compatible (Carson *et al*., 2003). There is evidence that in some cases, responses to the second bid are inconsistent with responses to the first bid. Some authors (*e.g.,* Alberini, 1995) have shown that pooling the responses to the first and second bids leads to some bias in the coefficient estimates, but a gain in efficiency.

The bid vector for the second part of each survey (except combined) takes as its starting point the next highest bid that was agreed to in the first part of the survey. One could randomize the bid vector, but true randomization could lead to a bid being offered for the combined valuation that would be less than what a respondent already agreed to for an individual endpoint. One could randomize the bid amount offered for the combined endpoints starting with the bid amount just above what had already been agreed to, but that isn't true randomization. Therefore, we decided to offer the next highest bid

following the one already agreed to (except in the case where a respondent said No-No to the first bid: in that case, we randomized the combined bid as well). Table 1 shows the relationship between the bid amounts for just the individual endpoints in the first part of each survey and the bid amounts for the combined total across both endpoints.

There are a series of motivation and "confidence" questions, including:

D6. *Thinking back on your responses for the tax you'd be willing to pay when thinking about the potential effects of PCBs on humans, how confident would you say you were about whether you would be for or against this referendum on a scale of 1 to 5 where 1 is "Not confident at all" and 5 is "Very confident"?*

The next set of questions asks about the confidence in responses for the endpoints individually and jointly (Conf.Human; Conf.Total). Another question asks whether respondents feel they can separate ecological and human endpoints in the valuation question. Another set of questions asks about familiarity with PCBs, concern about chemicals in the environment, and whether the respondent believes that PCBs really can cause these effects in humans and animals (risk.baby; risk.wldlf; ChemConcern; PCBConcern). Finally, respondents are asked to rate their trust on a one to five scale concerning the information they receive from a number of sources, including different web sites, print media, and television.

2.1. Endpoint Selection

Health effects resulting from environmental exposures can be acute (immediate) or chronic (longer term). Acute effects can often be ameliorated if the source of the exposure is removed (*e.g.*, asthma attacks as a result of air pollution), while chronic effects by definition tend to extend beyond the period of exposure (*e.g.*, the asthma itself, or the kinds of developmental effects explored here). In addition, with chronic effects, there can also be a latency period (*e.g.*, cancer, liver disease and other diseases that might

not reveal themselves until long after exposure has ceased). The bulk of the WTP studies found in the literature are for respiratory exposures (Van Houtven *et al*., 2003 provide a meta-analysis of 136 studies) leading to episodes of asthma or angina attacks. This study is designed to evaluate willingness to pay for a subtle effect (in humans) that occurs with a fairly large probability (20% chance if exposed) relative to typical cancer risks at Superfund sites.

The weight-of-evidence for a relationship between *in utero* polychlorinated biphenyl (PCB) exposure and developmental outcomes has been well established and continues to grow (Schantz *et al*., 2003). However, as with most epidemiological studies, discrepancies exist among measures of exposure and the strength of the relationships between the measures of exposure and developmental outcomes. Some of those discrepancies are attributable to differences in analytical methods, particularly in older studies (Longnecker *et al*., 2003) that had higher detection levels and less sophisticated quantitation techniques. Both epidemiological as well as animal studies demonstrate statistically significant increases in developmental delays and effects with increasing maternal PCB exposure (Jacobson and Jacobson, 2002b; Jacobson *et al.*, 2002; Levin *et al.*, 1988; Schantz *et al*., 1989, 1991; ATSDR, 2000). These effects can be seen in newborns as measured by the Bayley Scales of Infant Development to older children, measured either directly in terms of IQ or from other, related tests.

In terms of potential developmental effects, it is the *in utero* exposures that have been most implicated in terms of effects (Jacobson *et al.*, 1999; Jacobson and Jacobson, 2002b). Several studies have shown that although absolute doses of PCBs may be higher during breastfeeding due to mobilization of PCBs stored in maternal lipid, the protective

effects of breastfeeding itself together with other factors (*e.g*., nurturing home environment) potentially ameliorate the detrimental effects of PCBs. The children who showed the most statistically significant dramatic developmental delays were those exposed *in utero* and who were not breastfed. Breastfeeding may therefore be protective against developing these effects even if maternal body burdens are relatively high (Jacobson *et al*., 1999; Jacobson and Jacobson, 2002a).

However, regardless of the exposure issues, there is a substantial body of evidence that show declines in various cognitive responses across both human and animal studies (summarized in EPA, IRIS, www.epa.gov/IRIS/; ATSDR, 2000), typically as a result of *in utero* exposures. Much of our understanding of the implications of slight declines in cognitive ability across a population is based on work done relative to lead exposures (Schwartz *et al*., 1985; Schwartz, 1994). The research conducted in this area shows that slight declines in IQ which are difficult to detect in individuals and which may or may not lead to noticeable adverse effects on an individual basis are significant on a population level in terms of a population shift in IQ. Other cognitive effects include other kinds of developmental delays such as declines in reading comprehension to levels below grade level, low scores on analytical tests and tests of simple math problems, and behavioral responses.

The risk reductions used in the surveys are based on the results from Jacobson *et al*. (2000) who present a linear relationship between lipid-normalized breast milk concentration of PCBs and outcomes including a 6-point reduction in IQ and a 7-month deficit in reading comprehension as evidenced by scores on the WISC-R at eleven years for the Michigan cohort..

2.3. Risk Reduction and Tests of Scope

Sensitivity to scope can take several forms. Typically, these are referred to as *regular* embedding, (part-whole bias), and *perfect* embedding, or sensitivity of WTP to the stated risk reduction. There are two "part-whole" aspects to these surveys: one is within an endpoint, and the other is across endpoints. The human health endpoint doesn't have quite the same part-whole property as the ecological version of the survey since the potential human health effects of *in utero* exposures to PCBs include a panoply of developmental effects, all or some of which may or may not occur. Indeed, as stated in the survey:

> "*Studies involving children exposed while in the womb to PCBs have shown that these children perform less well on a variety of developmental tests throughout childhood. Government officials are interested in knowing whether you would be willing to pay a tax to remove the source of the PCBs for the benefit of protecting children exposed in the womb. Children that have been exposed to PCBs have been shown to have slightly lower IQ than average children, read at slightly below grade level, and are less able to perform simple math problems. The chemical doesn't cause the exact same effects in every child, but it does cause some effect in every child.*"

However, IQ does encompass general intelligence while reading comprehension is but one component of intelligence, allowing us to explore differences and/or similarities in the way respondents consider IQ versus reading comprehension as endpoints. Reduction in IQ as an endpoint has been well-studied in the literature particularly relative to exposures to lead and mercury. However, in terms of developmental endpoints, there is enough interindividual variability in IQ that makes an endpoint such as reading comprehension, which doesn't vary as much across repeated tests of any one individual, potentially more interesting in terms of valuation.

There has been increasing discussion in the CV literature concerning the effect of the placement of a particular good or endpoint within a valuation sequence and the influence that has on respondent valuation (Carson and Mitchell, 1995; Diamond, 1996; Bateman and Willis, 2001). Different WTP estimates are obtained depending on the order in which the benefits are presented, and additionally, the summation of the individual WTP values is often not the same as the overall WTP obtained without specifying individual endpoints. This is the issue of embedding, or part-whole bias, across endpoints. We explore this by administering three different versions of the survey. Two versions ask exactly the same set of questions except in opposite order (HHFirst, Ecofirst), and one survey asks only about the combined set of potential effects and risk reductions (human and ecological) to evaluate adding-up properties.

We evaluate perfect embedding by randomizing two different risk reductions for each endpoint across respondents as shown in Table 2. That is, each respondent sees only one risk reduction per developmental and ecological endpoint, but there are two risk reductions for each endpoint randomized across each subsurvey. We focus a number of the analyses on the risk reduction coefficient across surveys and endpoints.

2.4 Questions Related to Motivation

The survey contains a number of questions related to respondents' knowledge and beliefs regarding chemicals in the environment, PCBs in the environment, potential effects of PCBs, and trust in different sources of information (*e.g.*, industry scientists, media, and academia). The survey contains several follow-up questions designed to elicit motivation for agreeing to a particular bid. One question asks respondents to rate on a scale from not important to very important the specific reasons why they might be willing to pay to reduce potential risks to unborn children. We asked this follow-up question if

the respondent answered N-Y, Y-N, or Y-Y (*e.g.*, they agreed to any offered bid). The reasons include:

B5. People have lots of different reasons for voting for the program. Please rate the importance of the following reasons why you might vote for the program:

I'm worried about the potential risk to my own unborn children

I'm worried about the potential risk to unborn babies generally

I support a cleanup no matter what the risk might be (I don't like the idea of chemicals in the environment generally)

Some other reason: please specify

Likewise, for those respondents who answered N-N and were not willing to pay any amount, we asked the following:

D4. The State is interested in knowing why you would vote against the program. There are lots of different reasons why you might vote against the program, like it just isn't worth that much money, or it would be difficult for your household to pay that much even though you support the program, or you are opposed to dredging as an alternative. Or there might be some other reason.

Isn't worth the money…….1

Difficult for my household to pay……..2

Don't believe the cleanup would work...3

Some other reason, please specify: …...4

2.5 Quality Adjusted Life Years

All respondents see a set of questions designed to elicit utility weights for mild cognitive effects using either a standard gamble or time-tradeoff question format. Utility weights are typically elicited using a QALY index derived by questioning respondents about specific health states. The QALY index is defined as the product:

$$q \cdot T \qquad (1)$$

where:

q = a numerical gauge of the quality of the health index on a scale of zero to one (typically zero is the health state equivalent to death and one is perfect health, although values less than zero are possible for "worse than death" health states)

T = duration of health state

In one set of questions, respondents are asked to assume that they have a 10-year old child with a mild cognitive deficit, and are then offered either a standard gamble (SG) or time tradeoff (TTO) question concerning the mortality risk they would accept on behalf of their child for a perfect cure. These two approaches, SG and TTO, are the two primary methods used in the literature to elicit QALY weights (Gold, 1996).

The standard gamble offers the respondent a choice of a mild cognitive deficit in the child (either the reduction in IQ or reading comprehension deficit) for the remainder of the child's life (assumed to be 60 years) in comparison to a lottery of perfect health for that duration versus death. Respondents are asked about the probability of death that would be considered equivalent to a lifetime with a mild cognitive deficit. Table 3 shows the specific probabilities which range from 2.5 in 10,000 to 40 in 10,000.

The other elicitation scheme uses time tradeoff. Under this approach, the survey asks about years of longevity in perfect health a respondent would give up on behalf of the (hypothetical) 10-year old child to avoid a mild cognitive deficit that lasts a lifetime (60 years assuming a lifetime of 70 years). To correspond to the probabilities given above, the question asks about weeks of longevity that respondents would be willing to give up on behalf of an exposed child as shown in Table 3.

The question follows the same double-bounded dichotomous choice format as for WTP. That is, respondents are shown a time-tradeoff or probability of death, and if they respond "Yes", the followup questions asks about a larger number of weeks, or higher

probability of death. If they respond "No," the number of weeks, or probability, is cut in half. Respondents are shown a visual aid for the probability based on "dots" (Corso et al., 2001). The QALY weight that is assigned is equal to 1 – mortality risk interval agreed to by an individual respondent. The relationship between WTP and QALYs is given as:

$$WTP = \beta_0 * (\Delta q * \Delta t)^{\beta_1} + \varepsilon \qquad (2)$$

where:

Δq = change in health related quality of life

Δt = specific time period applicable to the quality weight

In this survey, respondents are asked to assume they have a 10-year-old child with the cognitive deficit, and what risk would they be willing to assume for this hypothetical child for a perfect cure. In the analysis, we assume that the child would live to be 70 years, so the duration of this health state is 60 years. In theory, WTP should increase proportionally relative to the gain in QALYs, which is testable under the hypothesis that $\beta_1 = 1$.

As with the WTP interval, the mortality risk that any given respondent agrees to is observed as an interval rather than the single value. Therefore, it was necessary to determine a single (conditional mean) mortality risk (or QALY weight, equal to 1-mortality risk) for each respondent. This was done as follows. First, we assume that the mortality risk interval for each respondent based on the two questions represents a single risk distribution. For each individual respondent *j*, there exists an upper and lower bound on the value, call these U*j* and L*j*, where L*j* is the minimum risk agreed to (which could be zero) and U*j* is the maximum risk the respondent accepted. The likelihood for this respondent is [F(U*j*) - F(L*j*)], where F is the cumulative distribution function (CDF) for the assumed distribution, which depends on a small number of parameters (*e.g.*, mean

and variance for normal). The likelihood for the full sample is just the product over *j* of the individual contributions to the likelihood, which depends on the parameters of the distribution function. To maximize it, we calculated the first derivatives with respect to the parameters and set them equal to zero.

2.6 Survey Administration

A professional survey firm, Knowledge Networks (KN), administered the survey to a panel representative of the US general population via a web-based survey mechanism during Spring 2005. The statistical foundation of the research panel stems from the application of probability-based sample selection methodologies to recruit panel members. The KN web-enabled panel is the only available method for conducting Internet-based survey research with a nationally representative probability sample (Couper, 2001; Krotki and Dennis, 2001).

The Knowledge Networks Panel, recruited randomly through Random Digit Dialing, represents the broad diversity and key demographic dimensions of the U.S. population. The web-enabled panel tracks closely the U.S. population on age, race, ethnicity, geographical region, employment status, and other demographic elements. The differences that do exist are small and are corrected statistically in survey data (i.e., by non-response adjustments). The web-enabled panel is comprised of both Internet and non-Internet households, all of which are provided the same equipment for participation in Internet surveys. Internet-based surveys are increasingly showing favorable comparisons to mail and telephone survey methods (Berrens *et al*., 2003).

There are four main factors responsible for the representativeness of the web-enabled research panel. First, the panel sample is selected using list-assisted random digit dialing telephone methodology, providing a probability-based starting sample of U.S.

telephone households. Second, the panel sample weights are adjusted to U.S. Census demographic benchmarks to reduce error due to non-coverage of non-telephone households and to reduce bias due to nonresponse and other non-sampling errors. Third, samples selected from the panel for individual studies are selected using probability methods. Appropriate sample design weights for each study are calculated based on specific design parameters. Fourth, nonresponse and poststratification weighting adjustments are applied to the final survey data to reduce the effects of non-sampling error (variance and bias).

The endpoint selection, specific risk reduction, and follow up human health questions are all randomized across the respondents. There are two human health endpoints, two risk reductions, two ecological endpoints and associated risk reductions, and two quality adjusted life year questions randomized across respondents. Each respondent faces only one human health endpoint and associated risk reduction, one ecological endpoint and associated risk reduction, and one QALY mortality risk (either SG or TTO).

In the next section, we report the results of the surveys and discuss the implications of the results.

3. Model Framework and Survey Results

Economic theory postulates that society is comprised of individuals who make tradeoffs in order to satisfy their preferences, or, put another way, to maximize their utility.

The statistical model for CV responses must satisfy both statistical and economic criteria (Hanemann and Kaninnen, 2001). CV responses can be modeled as discrete dependent variables with binary responses since respondents can either state “yes” or

"no" to a particular bid value. An equivalent but alternative modeling form takes the bid interval agreed to by an individual respondent as the dependent variable. In economic terms, the statistical model for CV responses must be consistent with the theory of utility maximization inherent in economic models. This assumes individuals show preferences for market commodities (x) and nonmarket amenities (q) as represented by a utility function $U(x,q)$ which is continuous and non-decreasing (Hanemann, 2001). Individuals face budget constraints based on income (y) and prices of the market commodities (p). Individuals are assumed to be utility-maximizers given a budget constraint (*e.g.*, disposable income). Willingness to pay, or the compensating variation (C) is the maximum an individual is willing to pay to secure an increase to the nonmarket amenity. In this case, the nonmarket amenity is expressed as a risk (r); therefore, a decrease in the risk increases utility $U(x, r)$.

Each respondent has an indirect utility function for which one can plot the tradeoff between risk and income while maintaining utility as given by the slope of that curve.

The economic measure of value is given as:

$$v(p, r_1, y\text{-}C) = v(p, r_0, y) \quad (3)$$

where C = the amount of money at which the individual is indifferent between a lower probability of risk and higher income, and r_0 and r_1 are different levels of:

- Risk of a 6-point reduction in IQ to an unborn child given maternal exposure (IQ)
- Risk of a 7-month deficit in reading comprehension given maternal exposure (RC)

The assumption is that a smaller risk relative to baseline leads improves well-being so compensating variation, or WTP, is positive. Expected utility is roughly

proportional to risk; consequently WTP should be approximately proportional to risk, and we test for this. As individuals spend more money, the utility loss increases. However, WTP is likely small with respect to income and so an income effect is also likely to be negligible.

All analyses are conducted using S-Plus 6.2 (Insightful Corporation, 2004) and Microsoft Excel.

3.1 Descriptive Statistics

Table 4 presents the frequencies of response to the bid vectors across the surveys. The proportion of yes responses decreases as the offered bid increases.

Table 5 provides a summary of the demographic characteristics of the sample, and for comparison purposes, data from the 2000 census. This table shows that the sample is representative of the US population. The median income differs, but this is primarily attributable to the fact that income was provided in terms of ranges, and the median income was estimated from the midpoint of the range provided for each individual. If one compares the income distribution (shown in the table below the median and mean income), it shows that survey samples are statistically indistinguishable from the demographics of the US population.

The sample also shows a lower proportion of individuals with less than a high school education as compared to the general public, and a higher proportion of individuals with at least an associates degree. However, it is not clear that more traditional survey methods (e.g., direct mail and/or telephone) would have reached a higher proportion of this fraction of the population.

Table 6 provides the means for model covariates.

3.2 Statistical Models

The double-bounded dichotomous choice elicitation format used here is analogous to interval-censored survival data in medical and engineering settings which model time to illness or failure of a component. In this case, we know the interval within which WTP for any individual respondent lies; for example, for the yes-yes response, it is known that the interval lies somewhere between the highest amount the respondent agreed to and infinity. Table 1 shows the intervals for each bid vector based on the initial bids for each survey, and Table 4 shows the proportion of respondents for each bid interval.

The WTP model takes the form:

$$LNWTP_i = \beta_0 + \beta_1 LN(\Delta Risk) + \beta_2 LNIncome + \beta_x X + \varepsilon \qquad (4)$$

where

WTP for the i^{th} individual in the interval given in Table 1

ΔRisk – is the risk reduction (0.1 or 0.15)

Income – respondent household income

X – vector of respondent-specific attributes as given in Table 6

ε – error term

The log likelihood function can be maximized assuming a particular parametric distribution (*e.g*, lognormal) or by using the Turnbull nonparametric modification of the Kaplan-Meier estimator, which makes no assumptions about the shape of the underlying WTP distribution (Carson *et al*., 2003; Hanemann and Kanninen, 2001). We evaluated several parametric forms (*e.g*., lognormal, weibull) and found the lognormal to provide the best fit based on a Likelihood Ratio test. In addition, properties of the lognormal distribution facilitate interpretation of the results. Figure 1 presents the visual goodness-of-fit plots across distribution types.

Parameter estimation is accomplished through maximum likelihood methods to obtain the values of unknown statistical parameters that are most likely to have generated the observed data. Figure 2 shows the WTP function for reading comprehension (IQ=0) for two risk reductions (0 = small risk reduction, 1 = large risk reduction) and for IQ (IQ=1).

Table 7 presents the results for several models based on the single endpoint valuation results of the HHFirst survey only. Models 1 and 2, stratified by endpoint (reading comprehension and IQ, respectively), include all covariates, while models 3 and 4 present the results for the reduced models. As shown in this table, the human health risk reduction coefficient is positively related to WTP, and approaches statistical significance for the IQ endpoint (p=0.14), but not for the reading comprehension endpoint. The only significant predictors in the full models include behavioral and motivational variables, including concern about PCBs in the environment (highly statistically significant across all four models), and the response to the QALY question (used in the model as change in QALY). As shown in Model 2, information received from scientists is positively associated with WTP (p<0.1). WTP is proportional with respect to risk reduction (coefficient = 1.0) for the IQ endpoint. Models with various interaction terms were not significant and are omitted from the table.

Table 8 presents the results from a set of models using the EcoFirst survey results for total WTP, which asks whether respondents would be willing to pay more into the cleanup fund when considering human health endpoints in addition to ecological endpoints. Models 1 and 2 are stratified by developmental endpoint for the total bid amount. Under this model, there is a difference between the risk reduction coefficient

(HHLNRR) for IQ as compared to reading comprehension as outcomes. For IQ, Table 8 shows the coefficient is 1.0 and approaches significance at $p<0.18$. For those respondents who were asked about reading comprehension as an endpoint, the risk reduction coefficient is statistically significant at –1.6 ($p<0.03$), indicating that respondents showed a negative relationship between risk reduction and WTP for this endpoint.

Models 3 and 4 in Table 8 show the results for the full models including all covariates for total WTP in the EcoFirst survey. For model 3, with reading comprehension as the endpoint, statistically significant covariates include the risk reduction coefficient, being female, concern about chemicals in the environment, whether or not the respondent believes that PCBs can cause developmental delays as a result of *in utero* exposures, and the QALY weight. All of these covariates are positively associated with WTP, except for the risk reduction coefficient. Model 4, by contrast, stratified by IQ as the endpoint, shows statistically significant covariates for the risk reduction variable, concern about PCBs in the environment, whether or not the respondent believes that PCBs can cause developmental delays as a result of *in utero* exposures, and the degree of confidence in information received from industry scientists. The risk reduction coefficient is positive, and only slightly more than proportional with respect to WTP, and statistically significant, unlike for the reading comprehension subset. Concern about PCBs in the environment generally and believing that PCBs can cause developmental delays are both positively associated with WTP for the IQ subset of respondents.

The magnitude of the risk reduction coefficient is very similar across both the HHfirst and Ecofirst surveys. Economic theory predicts that WTP should be

approximately proportional with respect to risk reduction, and this hypothesis cannot be rejected across these two datasets.

3.2.1 WTP per IQ Point

Cognitive ability, in addition to having an impact on later health status, also influences productivity through an impact on earning potential as well as through years of schooling and probability of employment. This relationship has been explored in the literature through the relationship between childhood lead exposures and loss of lifetime earnings by Grosse *et al.* (2002) and Salkever (1995). Grosse *et al.* (2002) evaluated three different linear relationships between earnings and IQ, ranging from 1.76% to 2.37% percentage earnings loss per IQ point. Based on this relationship, and the present value of earnings of a two-year-old in 2000 dollars, results in values of a one point decrease in IQ ranging from $12,700 to $17,200.

Estimates of WTP using these survey results represent WTP for a probability of a 6-point reduction in IQ, thus, WTP for a 100% probability of a 1-point reduction is estimated by dividing WTP by 6 and dividing again by the risk reduction. This assumes that WTP is linear in the probability of a reduction in IQ as a result of exposure and the number of IQ points at risk. We evaluated WTP per IQ point using both the single endpoint results from the HHFirst survey and the difference between the total valuation and single endpoint valuation from the EcoFirst survey. The result for the HHFirst survey is $466 (95% confidence interval = $380, $520) per IQ point.

3.2.2 WTP and QALYs

Table 9 shows the results of the models across surveys. The dependent variable for the first model is the interval-censored WTP for the first set of questions from the HHFirst survey, while the second model dependent variable is the total interval-censored

bid amount from the EcoFirst survey. In both cases, covariates include whether the endpoint was IQ (1) or reading comprehension (0), and a code for whether the elicitation method for the QALY weight was standard gamble (0) or time-tradeoff (1). Finally, the change in QALY for each respondent was calculated as described in section 3.4 (LNQALY). The resulting coefficients are very similar across the datasets, except for IQ. For the HHFirst survey, there is no appreciable difference in the relationship between change in QALY and WTP by developmental endpoint. But for the EcoFirst survey, the IQ coefficient is negative and statistically significant. Respondents to that survey had a 33% lower WTP when asked about IQ as compared to reading comprehension.

The individual QALY weights (1 – mortality risk) range from 0.948 to 0.99975 for a 6-point reduction in IQ. This translates to a range of mortality risks that respondents would accept on behalf of their (hypothetical) 10-year-old child of 2 in 10,000 to 9 in 1,000 per IQ point. Table 10 shows the mean, standard deviation, and number of respondents by endpoint (IQ or reading comprehension) and elicitation method (standard gamble or time tradeoff). There is no statistical difference by endpoint ($\chi 2 = 0.6$, df=1, p=0.4), while there is a statistically significant difference by elicitation method ($\chi 2 =$ 10.1, df=1, p=0.001).

We estimated WTP per QALY by dividing WTP by the expected change in QALYs, where the change in QALY accounts for the probability of having the cognitive deficit. The mean WTP per QALY is $109,000 (95% confidence interval = ($70,000, $148,000). WTP per QALY has been proposed as a potential criterion for evaluating efficacy of social programs (Baker *et al.,* 2004; Gyrd-Hansen, 2003; Krupnick, 2004; Van Houtven *et al*., 2003) based on cost-effectiveness. King et al. (2005) discuss

standards for evaluating WTP/QALY ratios, and find that this ratio varies considerably depending on the valuation methodology. In 2003 dollars, the median ratio from eight CV studies based on (personal) safety was $184,200. In contrast, revealed preference studies, based on safety, have a median value of $106,700. The results of this study are consistent with these literature values.

4. Discussion

The importance of obtaining behavioral and motivational answers from respondents in CV surveys has been shown (Heberlein *et al.,* 2005; Nunes and Schokkaert, 2003; Dubourg *et al.,* 1997). In this case, concern about PCBs in the environment and the respondent-specific QALY weighting are important, highly statistically significant predictors of WTP. The QALY weighting indirectly addresses perceived risk in that it elicits from respondents an indication of the perception the parent has about the quality of life for the child if s/he has the cognitive deficit. It addresses the issue more directly by asking about *your* hypothetical child, as opposed to how significant do you think the risks are in general (*e.g.*, risk.baby, PCBChild).

Interestingly, in responses to open ended questions, a number of respondents indicated that because there were fish consumption advisories in place in their particular State (indeed, most States), they felt the risks were lower than what had been portrayed in the survey, although the survey does indicate that the risks are only to those individuals who consume fish.

The risk reduction coefficients for IQ are both positive and approaching statistical significance based on the responses to the single endpoint in the HHFirst survey (1.0, p=0.14) and the EcoFirst total endpoint (1.1, p=0.14), providing greater confidence that the surveys have captured the relationship between risk reduction and WTP for IQ. In a

reduced model using just risk reduction as a predictor based on the single endpoint in the HHfirst survey, the coefficient is 1.0 (p=0.15), a proportional result approaching significance. The results for reading comprehension as an endpoint are not as robust. These results suggest that survey takers were able to think about IQ as a developmental endpoint and were indeed willing to pay for risk reductions, while this is not the case for reading comprehension.

It is true that these risks are not experienced directly by the respondents themselves. Women of childbearing age who are pregnant or thinking of becoming pregnant and that consume freshwater fish are the only ones who would actually be exposed, and even in that case, they do not experience the risk directly. The risk is to the unborn child. This is the most immediate that the risk can be, but the proportion of respondents who are pregnant (this question was not asked – the only information we have is the number of women of child-bearing age and the number of children by age group in the household) is itself likely a relatively small proportion of the overall respondent population.

Respondents were willing to increase their stated bids between the single ecological endpoint in the EcoFirst survey when asked about a total bid. This was not the case in the HHFirst survey (respondents were not willing to increase their stated bids when asked about ecological effects after they had already responded to human health endpoints).

The estimated WTP values per IQ point from these surveys are orders of magnitude lower than estimates based on future earnings. The estimates obtained here are approximately $500 while the estimates from the earnings literature are in the $10,000 to

$20,000 range. The results presented here represent the average WTP per IQ point from a representative sample of the American general public. It is possible that respondents do not realize (or do not think about) the implications of IQ on future earnings and so underestimated the potential value of the loss. Another possibility is that respondents recognize the effect on future earnings, but use higher discount rates in evaluating these benefits than the rates used to calculate the estimates from the literature (consistent with the idea that people discount the future too much).

The policy implications of these WTP values, however they are expressed, comes in the context of a particular decision. One of the goals of this survey was to demonstrate how stated preference methods might be used to develop economic values for risk reductions within a particular regulatory framework. In a companion paper (von Stackelberg, 2006), we develop an application based on the Hudson River Superfund site to show how this might be done.

The survey results suggest that IQ represented a more meaningful endpoint for respondents than reading comprehension. However, it is known that people have difficulty evaluating and responding to numerical differences in the magnitude of risk reduction, particularly for small risks or small effects (Hammitt and Graham, 1999; Corso *et al*., 2001; Schwartz *et al.,* 1997). Further, in this case, exposures are experienced by one cohort while effects are experienced by another who also happen to be children and therefore unable to make risk-based decisions for themselves. Women of childbearing age who are pregnant or thinking of becoming pregnant and that consume freshwater fish are the only ones who would actually be exposed, and even in that case, they do not experience the risk directly. The risk is to the unborn child. This is the most immediate

that the risk can be, but the proportion of respondents who are pregnant (this question was not asked – the only information we have is the number of women of child-bearing age and the number of children by age group in the household) is itself likely a relatively small proportion of the overall respondent population. However, this is an issue that is likely to arise time and again with significant policy implications given the increasing evidence of *in utero* environmental exposures leading to significant and potentially lasting health effects later in life. It is, after all, children who presumably still have most of their lives in front of them and will be the ones who directly experience the repercussions of decisions made today, ostensibly on their behalf.

References

(ATSDR) Agency for Toxic Substances and Disease Registry. 2000. Toxicological profile for polychlorinated biphenyls (PCBs). US Department of Health and Human Services, Public Health Service. Prepared by Syracuse Research Corporation, November.

Alberini, A. 1995. Efficiency vs bias of willingness-to-pay estimates: Bivariate and interval data models. *Journal of Environmental Economics and Management* 29:169-180.

Alberini, A. 2005. What is a life worth? Robustness of VSL values from contingent valuation surveys. *Risk Analysis* 25(4):783-800.

Baker, R., S. Chilton, C. Donaldson, M. Jones-Lee, H. Metcalf, P. Shackley and M. Ryan. 2004. Determining the societal value of a QALY by surveying the public in England and Wales: A research protocol. Submission to the National Institute for Clinical Excellence and National Co-ordinating Centre for Research Methodology, School of Population and Health Sciences, University of Newcastle upon Tyne.

Bateman, I.J. and K.G. Willis, Eds. 1999. Valuing environmental preferences: Theory and practice of the contingent valuation method in the US, EU, and developing countries. New York: Oxford University Press. 645 pp.

Berrens, R., A. Bohara, H. Jenkins-Smith, C. Sliva and D. Weimer. 2003. The advent of Internet surveys for political research: A comparison of telephone and Internet samples. *Political Analysis*, 11(1).

Cannon, M., R. Kopp and A. Krupnick. 1996. *Valuation of Developmental and Reproductive Effects: A Scoping Study.* Resources for the Future. October, 1996.

Carson, R.T. and R.C. Mitchell. 1995. Sequencing and nesting in contingent valuation surveys. *Journal of Environmental Economics and Management* 28:155-173.

Carson, R.T., R.C. Mitchell, M. Hanemann, R.J. Kopp, S. Presser and P.A. Ruud. 2003. Contingent valuation and lost passive use: Damages from the Exxon Valdez oil spill. *Environmental and Resource Economics* 25:257-286.

Corso, P.S., J.K. Hammitt and J.D. Graham. 2001. Valuing mortality-risk reduction: Using visual aids to improve the validity of contingent valuation. *Journal of Risk and Uncertainty* 23:165-184.

Couper, M.P., M.W. Traugott and M.J. Lamias. 2001. Web survey design and administration. *Public Opinion Quarterly 65* (2) 230 –253.

Diamond, P. 1996. Testing the internal consistency of contingent valuation surveys. *Journal of Environmental Economics and Management* 30:337-347.

Dickie, M. and S. Gerking. 2002. Willingness to pay for reduced morbidity. Manuscript for presentation at the workshop, "Economic Valuation of Health for Environmental Policy: Assessing Alternative Approaches," March 18-19, 2002, Orlando, FL.

Dockins, C., R.R. Jenkins, N. Owens, N.B. Simon and L.B Wiggins. 2002. Valuation of childhood risk reduction: The importance of age, risk preferences, and perspective. *Risk Analysis* 22(2):335-346.

Dubourg, W.R., M.W. Jones-Lee and G. Loomes. 1997. Imprecise preferences and survey design in contingent valuation. *Economica* 64:681-702.

Gold, M.R., J.E. Siegel, L.B. Russell and M.C. Weinstein. 1996. Cost-effectiveness in health and medicine. New York: Oxford University Press.

Grosse, S.D., T.D. Matte, J. Schwartz and R.J. Jackson. 2002. Economic gains resulting from the reduction in children's exposure to lead in the United States. *Environmental Health Perspectives* 110(6):563-569.

Gyrd-Hansen, D. 2003. Willingness to pay for a QALY. *Health Economics* 12:1049-1060.

Hammitt, J.K. and J.D. Graham. 1999. Willingness to pay for health protection: inadequate sensitivity to probability? *Journal of Risk and Uncertainty* 18:33-62.

Hammitt, J.K. and J-T. Liu. 2004. Effects of disease type and latency on the value of mortality risk. *Journal of Risk and Uncertainty* 28(1):73-95.

Hanemann, W.M. 1991. Willingness to pay and willingness to accept: How much can they differ? *American Economic Review* 81(3): 635-647.

Hanemann, W.M., J.B. Loomis and B.J. Kanninen. 1991. Statistical efficiency of double-bounded dichotomous choice contingent valuation, *Am J Ag Ec* 73:(4)1255-1263.

Hanemann, W.M. and B. Kanninen. 2001. Statistical analysis of discrete-response CV data. In *Valuing Environmental Preferences*, Bateman, I.J. and K.G. Willis, Eds. New York: Oxford University Press.

Heberlein, T.A., M.A. Wilson, R.C. Bishop and N.C. Shaeffer. 2005. Rethinking the scope test as a criterion for validity in contingent valuation. *Journal of Environmental Economics and Management* 50:1-22.

Insightful Corporation, Inc. 2004. Users Manual for S-Plus.

Jacobson, S.W., L.M. Chiodo and J.L. Jacobson. 1999. Breastfeeding effects on intelligence quotient in 4- and 11-year-old children. *Pediatrics* 103:71-76.

Jacobson, J.L and S.W. Jacobson. 2002. Breast-feeding and gender as moderators of teratogenic effects on cognitive development. *Neurotoxicology and Teratology* 24:349-358.

Jacobson, J.L. and S.W. Jacobson. 2002. Association of prenatal exposure to an environmental contaminant with intellectual function in childhood. *Clinical Toxicology* 40(4):467-475.

Jacobson, J.L., J. Janisse, M. Banerjee, J. Jester, S.W. Jacobson and J.W. Ager. 2002. A benchmark dose analysis of prenatal exposure to polychlorinated biphenyls. *Environ Health Perspect* 110(4):393-398.

King, J.T., J.Tsevat, J.R. Lave and M.S. Roberts. 2005. Willingness to pay for a quality-adjusted life year: Implications for societal health care resource allocation. *Medical Decision Making* 25(6):667-677.

Krotki, K. and M.J. Dennis. 2001. Probability-based Survey Research on the Internet. Presented at the August 2001 Conference of the International Statistical Institute.

Krupnick, A.J. 2004. Valuing health outcomes: Policy choices and technical issues. Washington, D.C.: Resources for the Future, 104 pp.

Levin, E.D., S.L. Schantz and R.E. Bowman. 1988. Delayed spatial alternation deficits resulting from perinatal PCB exposure in monkeys. *Arch Toxicol* 62: 267-273.

Longnecker, M.P., Wolff, M.S., Gladen, B.C., Brock, J.W., Grandjean, P., Jacobson, J.L., Korrick, S.A., Rogan, W.J., Weisglas-Kuperus, N., Hertz-Picciotto, I., Ayotte, P., Stewart, P., Winneke, G., Charles, M.J., Jacobson, S.W., Dewailly, E., Boersma, E.R., Altshul, L.M., Heinzow, B., Pagano, J.J. and A.A. Jensen. 2003. Comparison of polychlorinated biphenyl levels across studies of human neurodevelopment. *Environ Health Perspect* 111(1):65-70.

McFadden, D. and G. Leonard. 1993. Issues in the contingent valuation of environmental goods: methodologies for data collection and analysis, In Hausman, J.A. Ed. *Contingent Valuation: A Critical Assessment*. New York: North-Holland, 165-215.

Salkever, D.S. 1995. Updated estimates of earnings benefits from reduced exposure of children to environmental lead. *Environmental Research* 70:1-6.

Schantz, S.L., E.D. Levin, R.E. Bowman. 1989. Effects of perinatal PCB exposure on discrimination-reversal learning in monkeys. *Neurotoxicol Teratol* 11: 243-250.

Schantz, S.L., J.J. Widholm and D.C. Rice. 2003. Effects of PCB exposure on neuropsychological function in children. *Environ Health Perspect* 111(3):357-376.

Schwartz, J. 1994. Societal benefits of reducing lead exposure. *Environmental Research* 66:105-124.

Schwartz, J., H. Pitcher, R. Levin, B. Ostro and A.L. Nicholas. 1985. *Cost and Benefits of Reducing Lead in Gasoline: Final Regulatory Impact Analysis*. Washington, D.C.: Office of Policy Analysis, Environmental Protection Agency.

Schwartz, L.M., S. Woloshin, W.C. Black and H.G. Welch. 1997. The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine* 127:966-972.

Van Houtven, G., M. Rousu, J-C. Yang, C. Pringle, W. Wagstaff and J. DePlatchett. 2003. Valuation of morbidity losses: Meta-analysis of willingness-to-pay and health status measures. Final Report, Prepared for the Food and Drug Administration by Research Triangle Institute.

Van Houtven, G., Powers, J., Jessup, A. and J-C. Yang. 2004. Valuing avoided morbidity using meta-analysis: What can health status measures and QALYs tell us about WTP? Research Triangle Institute Working Paper 04-01, www.rti.org/enrepaper/.

TABLE 1: Initial Bid Vectors and Followup Bids for the CV Surveys

	Bid vectors based on final response in first section and are given as initial bid, upper, lower:			
Initial Bid	Y-Y[1]	Y-N[1]	N-Y[1]	N-N
$25	C ($100, $200, $50)	B ($50, $100, $25)	A ($25, $50, $10)	random
$50	D ($200, $400, $100)	C ($100, $200, $50)	B ($50, $100, $25)	random
$100	E ($400, $800, $200)	D ($200, $400, $100)	C ($100, $200, $50)	random
$200	F ($800, $1000, $400)	E ($400, $800, $200)	D ($200, $400, $100)	random
$400	G ($1000, $1500, $800)	F ($800, $1000, $400)	E ($400, $800, $200)	random
$800	H ($2000, $1500, $800)	G ($1000, $1500, $800)	F ($800, $1000, $400)	random

Notes:
1 – It is possible, in the followup, to respond "no" to a value for the total that had already been agreed to in the previous section. In that case, respondents are shown the following prompt: "You already agreed you'd be willing to pay this amount for human health benefits alone. Now we're asking about the total you'd be willing to pay"

TABLE 2: Risk Reductions in the Surveys

Endpoint	Context	Small Risk Reduction	Large Risk Reduction
Eagle	Probability of reproductive impairment significant enough to affect viability of the population	0.1	0.15
Species Sensitivity Distribution (SSD)	Probability of reproductive significant reproductive effects to 20% of all avian species in a freshwater ecosystem	0.25	0.4
Reading Comprehension	Probability of reading at approximately 7 months below grade level	0.1	0.15
IQ	Probability of a 6-point reduction in IQ	0.1	0.15

TABLE 3: Mortality Risk and Longevity Reduction Questions to Determine QALYs

Initial Probability of Death versus Successful Treatment	Followup Probability if "yes"	Followup Probability if "No"	Initial Reduction in Longevity (days)	Followup Reduction if "yes" (days)	Followup Reduction if "No" (days)
5 in 10,000	10 in 10,000	2.5 in 10,000	11	22	5
10 in 10,000	20 in 10,000	5 in 10,000	22	44	11
20 in 10,000	40 in 10,000	10 in 10,000	44	88	22

QALYcode = 1 if life expectancy reduction, 0 if mortality risk

TABLE 4: Proportion of Respondents in Each Bid Interval for HHFirst (Single Endpoint) and Ecofirst (Total Across Endpoints)

HHFIRST -- Single Endpoint	**IQ (n=208)**					**RC (n=196)**				
Bid Amount	n	Y-Y	Y-N	N-Y	N-N	n	Y-Y	Y-N	N-Y	N-N
A ($25, $50, $10)	35	11%	3%	0%	3%	35	12%	2%	1%	4%
B ($50, $100, $25)	36	8%	4%	1%	5%	32	7%	5%	4%	2%
C ($100, $200, $50)	27	3%	3%	2%	5%	21	3%	1%	2%	3%
D ($200, $400, $100)	30	4%	3%	2%	4%	33	4%	4%	2%	7%
E ($400, $800, $200)	41	2%	5%	4%	8%	40	4%	5%	2%	10%
F ($800, $1000, $400)	33	4%	1%	1%	9%	32	3%	4%	2%	7%

ECOFIRST -- Total Bid for Both Endpoints	**IQ (n=194)**					**RC (n=208)**				
Bid Amount	n	Y-Y	Y-N	N-Y	N-N	n	Y-Y	Y-N	N-Y	N-N
A ($25, $50, $10)	11	0%	2%	2%	2%	14	2%	0%	1%	3%
B ($50, $100, $25)	16	2%	3%	2%	3%	18	1%	3%	2%	2%
C ($100, $200, $50)	37	11%	5%	6%	3%	47	10%	5%	5%	3%
D ($200, $400, $100)	47	6%	8%	7%	4%	39	3%	8%	6%	2%
E ($400, $800, $200)	30	0%	7%	5%	4%	31	3%	2%	4%	5%
F ($800, $1000, $400	32	3%	1%	6%	8%	32	5%	1%	6%	3%
G ($1000, $1500, $800)	10	2%	2%	2%	0%	11	1%	1%	3%	0%
H ($1500, $2000, $1000)	5	2%	0%	1%	0%	9	3%	0%	1%	0%

TABLE 4, continued: Proportion of Respondents in Each Bid Interval for the Combined Survey

COMBINED	**Combined (n=204)**				
Bid Amount	n	Y-Y	Y-N	N-Y	N-N
A ($25, $50, $10)	37	11%	4%	0%	3%
B ($50, $100, $25)	41	9%	6%	0%	5%
C ($100, $200, $50)	23	4%	2%	1%	4%
D ($200, $400, $100)	34	5%	4%	2%	5%
E ($400, $800, $200)	35	2%	5%	1%	9%
F ($800, $1000, $400)	29	3%	3%	0%	8%

TABLE 5: Demographics for each Subsurvey and the US Census

	ECOFIRST		HUMANFIRST		COMBINED	
Demographic	Eagle (n=193)	SSD (n=210)	RC (n=196)	IQ (n=208)	Combined (n=204)	US Census Data[1]
Some high school, no diploma	7%	8%	19%	11%	16%	20%
High school	29%	30%	29%	35%	32%	29%
Some college, no degree	23%	20%	21%	24%	21%	21%
Associate degree (AA, AS)	15%	12%	7%	5%	6%	6%
Bachelor's degree	17%	19%	16%	19%	14%	16%
Master's degree	4%	7%	7%	5%	9%	6%
Other	5%	4%	2%	2%	1%	3%
Black, Non-Hispanic	10%	12%	12%	15%	12%	12%
Hispanic	9%	15%	17%	9%	11%	13%
Other, Non-Hispanic	5%	5%	4%	4%	5%	0%
White, Non-Hispanic	76%	68%	67%	72%	72%	75%
Female	57%	50%	48%	51%	52%	51%
Male	43%	50%	52%	49%	48%	49%
Income						
Less than $10,000	12%	10%	12%	13%	13%	10%
$10,000 to $14,999	11%	5%	9%	8%	4%	6%
$15,000 to $19,999	5%	4%	5%	4%	8%	6%
$20,000 to $24,999	8%	10%	6%	8%	5%	7%
$25,000 to $29,999	8%	7%	10%	6%	5%	6%
$30,000 to $34,999	7%	7%	5%	4%	8%	6%
$35,000 to $39,999	4%	10%	10%	10%	9%	6%
$40,000 to $49,999	9%	11%	10%	6%	15%	11%
$50,000 to $59,999	10%	9%	7%	13%	7%	9%
$60,000 to $74,999	10%	9%	8%	12%	12%	10%
$75,000 to $99,999	11%	9%	12%	6%	7%	10%
$100,000 to $124,999	2%	3%	5%	5%	5%	5%
$125,000 to $149,999	1%	2%	1%	1%	2%	3%
$150,000 to $174,999	1%	1%	1%	0%	2%	2%
$175,000 or more	2%	2%	2%	0%	1%	2%
Divorced	12%	15%	13%	20%	14%	10%
Married	52%	50%	48%	46%	52%	54%
Separated	2%	2%	3%	4%	1%	2%
Single (never married)	26%	28%	28%	26%	29%	27%
Widowed	7%	5%	7%	4%	3%	7%

1: Data provided for males and females combined (except gender); therefore, percentages may not equal 100 due to combining. Data from: factfinder.census.gov, 2000 Census

TABLE 6: Means for the Covariates Across Subsurveys

		Eagle (n=193)		SSD (n=210)		IQ (n=208)		RC (n=196)		Combined (n=204)	
		ECOFIRST				HHFIRST				COMBINED	
Parameter	Parameter Name	Mean	Stdev	Mean	Stdev	Mean	Stdev	Mean	Stdev	Mean	Stdev
Education (1 for college and above, 0 otherwise)	EDUCAT	0.53		0.61		0.55		0.53		0.50	
White (1 for yes, 0 otherwise)	WHITE	0.76		0.68		0.72		0.67		0.71	
Black (1 for yes, 0 otherwise)	BLACK	0.09		0.12		0.15		0.12		0.22	
Hispanic (1 for yes, 0 otherwise)	HISPANIC	0.09		0.15		0.09		0.17		0.14	
Gender (1 if Female, 0 if Male)	MALE	0.57		0.50		0.52		0.48		0.52	
Natural log of income	LNInc	10.36	0.86	10.46	0.83	10.41	0.86	10.41	0.89	10.38	0.89
Married (1 if yes, 0 otherwise)	MARRIED	0.52		0.50		0.46		0.48		0.52	
Live in a metropolitan area (1 if yes, 0 if no)	METRO	0.83		0.82		0.83		0.84		0.79	
Natural log of ecological risk reduction	LNEcoRR	-2.09	0.20	-1.17	0.23	-1.67	0.49	-1.60	0.52	-2.11	0.21
Natural log of human health risk reduction	HHLNRR					-2.09	0.20	-2.09	0.20	-2.09	0.20
Have you ever heard of PCBs (1 if yes, 0 otherwise)	PCBs	0.48		0.50		0.45		0.43		0.41	
Confidence in response to single endpoint valuation (scale of 1 to 5 where 1 is not confident and 5 is very confident)	ConfWildlife	4.39		4.16	1.64	3.70	1.15	3.62	1.16	na	
Confidence in total	ConfTotal	4.55	1.19	4.06	1.71	3.67	1.11	3.60	1.15	3.31	1.39
QALY code (0 if standard gamble, 1 if time tradeoff	QALYcode										
Are you able to think about ecological endpoints separately from human (1 if yes, 0 if no)	eco.sep	0.78		0.72		0.71		0.77		na	
Are you able to think about ecological benefits separately from human health benefits? (1 if yes, 0 otherwise)	eco.ben.sep	0.62		0.63		0.62		0.64		na	
Concerned about chemicals in the environment (1 if yes, 0 otherwise)	ChemConcern	3.12		2.96		3.04		2.89		3.03	

TABLE 6: Means for the Covariates Across Subsurveys

Concerned about PCBs in the environment (1 if yes, 0 otherwise)	PCBConcern	2.96		2.77		2.69		2.62		2.87	
Do you believe PCBs can cause reproductive effects in wildlife? (1 if yes, 0 otherwise)	PCBWildlife	0.66		0.59		0.59		0.60		0.60	
Do you believe PCBs can cause developmental effects in children exposed *in utero* ? (1 if yes, 0 otherwise)	PCBChild	0.61		0.54		0.65		0.60		0.59	
Rate the risks facing eagles in this state (0 = not sure, 1 = not serious, 2 = somewhat serious, 3 = very serious, 4 = extremely serious)	risk.wldlf	2.14	1.17	2.04	1.20	1.94	1.17	1.94	1.19	2.08	1.13
Rate the risks facing unborn babies in this state (0 = not sure, 1 = not serious, 2 = somewhat serious, 3 = very serious, 4 = extremely serious)	risk.baby	2.22	1.27	2.01	1.28	2.17	1.25	2.11	1.30	2.16	1.29
How often do you watch programs on television about wildlife (1 = never, 2 = rarely, 3 = sometimes, 4 = often)	tv.wldlf	2.99	0.88	2.91	0.97	2.75	0.94	3.03	0.93	2.90	0.90
Do you live near freshwater (1 = yes, 0 = no)	live.fw	0.69		0.64		0.60		0.60		0.66	
How much time do you spend on a river, lake, or stream? (1 = never, 2 = rarely, 3 = sometimes, 4 = often)	time.fw	2.60	1.03	2.65	1.02	2.49	0.97	2.61	1.03	2.62	0.99
How often do you eat recreationally caught fish (0 = never, 1 = a few times a year, 2 = a few times a month, 3 = a few times a week)	eat.fish	2.50	0.81	2.53	0.85	2.51	0.80	2.47	0.85	2.57	0.83
How much confidence do you have in information you receive from government sources (1 = none, 2 = some, 3 = a lot)	conf.gov	1.85	0.56	1.78	0.49	1.93	0.53	1.85	0.30	1.85	0.51
How much confidence do you have in information you receive from industry scientists (1 = none, 2 = some, 3 = a lot)	conf.sci.ind	1.88	0.58	1.82	0.54	1.85	0.62	1.81	0.60	1.86	0.58

TABLE 6: Means for the Covariates Across Subsurveys

How much confidence do you have in information you receive from university scientists (1 = none, 2 = some, 3 = a lot)	conf.sci.univ	2.25	0.59	2.27	0.60	2.21	0.60	2.20	0.59	2.31	0.56
How much confidence do you have in information you receive from television sources (1 = none, 2 = some, 3 = a lot)	conf.tv	1.70	0.58	1.68	0.54	1.72	0.55	1.70	0.56	1.71	0.56
How much confidence do you have in information you receive from government web sites (1 = none, 2 = some, 3 = a lot)	conf.gov.web	1.87	0.50	1.78	0.53	1.87	0.55	1.83	0.54	1.81	0.50
How much confidence do you have in information you receive from commercial web sites (1 = none, 2 = some, 3 = a lot)	conf.comm.web	1.69	0.52	1.62	0.52	1.61	0.55	1.59	0.52	1.65	0.52
How much confidence do you have in information you receive from nonprofit web sites (1 = none, 2 = some, 3 = a lot)	conf.np.web	2.10	0.62	2.09	0.58	2.04	0.58	2.02	0.60	2.05	0.59
How much confidence do you have in information you receive from university web sites (1 = none, 2 = some, 3 = a lot)	conf.uni.web	2.21	0.59	2.20	0.54	2.12	0.64	2.06	0.62	2.15	0.58
How much confidence do you have in information you receive from print media (1 = none, 2 = some, 3 = a lot)	conf.print	1.86	0.56	1.88	0.40	1.84	0.53	1.81	0.51	1.88	0.54

TABLE 7: Model Results for HHFirst Model for Developmental Endpoints

	Model 1 RC only	Model 2 IQ only	Model 3 across endpoints	Model 3 RC only	Model 4 IQ only
Intercept	1.3 (2.4)	1.6 (2.9)	4.6 (1.1)****	3.5 (1.5)**	5.9 (1.6)****
Risk Reduction	0.1 (0.7)	0.5 (0.7)	0.7 (0.5)	0.4 (0.7)	1.0 (0.7)
Age	-0.01 (0.009)	0.002 (0.009)			
Education	0.2 (0.3)	0.6 (0.3)*			
Race (Ref = White)					
Other	1.2 (0.7)	0.6 (0.9)			
Black	0.2 (0.5)	0.1 (0.4)			
Hispanic	0.07 (0.4)	0.2 (0.6)			
Male	0.2 (0.3)	-0.02 (0.3)			
Income	-0.08 (0.2)	-0.01 (0.2)			
Married	0.09 (0.3)	-0.1 (0.3)			
Metro	0.9 (0.4)**	0.1 (0.4)			
PCBConcern	0.9 (0.2)****	0.4 (0.2)***	0.9 (0.1)****	1.1 (0.1)****	0.8 (0.2)****
QALY	0.2 (0.1)**	0.3 (0.1)****	0.3 (0.1)****	0.2 (0.1)***	0.3 (0.1)***
risk.baby	0.08 (0.1)	0.3 (0.1)**			
live.fw	0.3 (0.3)	-0.02 (0.2)			
eat.fish	0.2 (0.2)	0.1 (0.2)			
confgov	0.3 (0.3)	0.2 (0.3)			
conf.sci.ind	-0.2 (0.3)	0.5 (0.3)*			
conf.sci.uni	0.4 (0.3)	0.8 (0.3)**			
conf.tv	0.03 (0.3)	-0.5 (0.4)			
conf.print	0.3 (0.4)	0.2 (0.4)			
-2*Log-Likelihood	423	460	942	444	492
	n=192	n=206	n=398	n=192	n=206

* p<0.10, ** p<0.05, *** p<0.01, **** p<0.001

TABLE 8: Model Results for EcoFirst Model for
Total WTP Based on Developmental Endpoints

	Model 1 RC only	Model 2 IQ only	Model 3 RC only	Model 4 IQ only
Intercept	2.2 (1.5)	7.3 (1.3)****	-0.008 (2.6)	4.4 (2.1)**
Risk Reduction	-1.6 (0.7)**	1.0 (0.6)	-1.3 (0.6)**	1.1 (0.6)**
Eagle			0.2 (0.3)	-0.4 (0.2)
Education			-0.2 (0.3)	0.06 (0.3)
Race (Ref = White)				
Other			0.4 (0.9)	-0.3 (0.4)
Black			0.4 (0.5)	0.4 (0.4)
Hispanic			0.8 (0.4)	0.2 (0.4)
Male			0.5 (0.3)**	0.2 (0.2)
Age			-0.003 (0.009)	-0.001 (0.008)
Income			-0.04 (0.2)	-0.02 (0.2)
Married			0.3 (0.3)	-0.1 (0.3)
Metro			0.05 (0.4)	-0.1 (0.3)
PCBConcern			0.6 (0.2)****	0.4 (0.2)**
risk.baby			0.3 (0.1)***	0.2 (0.1)*
live.fw			0.1 (0.3)	-0.3 (0.3)
QALY			0.2 (0.06)***	0.06 (0.05)
eat.fish			-0.2 (0.2)	0.2 (0.2)
confgov			0.4 (0.3)	0.3 (0.3)
conf.sci.ind			-0.3 (0.3)	-0.1 (0.2)
conf.sci.uni			0.2 (0.4)	0.5 (0.2)**
conf.tv			-0.01 (0.3)	-0.1 (0.3)
conf.print			0.1 (0.3)	0.4 (0.3)
-2*Log-Likelihood	704	658	635	569
	n=208	n=194	n=205	n=188

* p<0.10, ** p<0.05, *** p<0.01, **** p<0.001

TABLE 9: WTP versus QALY Across Surveys

	HHFirst Single Endpoint	Ecofirst Total Endpoint
Intercept	5.6 (0.3)****	5.8 (0.2)****
IQ	0.08 (0.2)	-0.4 (0.2)**
QALYcode	0.1 (0.2)	0.1 (0.2)
LNQALY	0.3 (0.07)****	0.1 (0.04)***
-2*Log-Likelihood	1034	1345
	n=398	n=397
* p<0.10, ** p<0.05, *** p<0.01, **** p<0.001		

TABLE 10: Mean (Standard Deviation) QALY Weights by Endpoint and Elicitation Method

Elicitation Method	IQ (1)	n	RC (0)	n
Standard Gamble (Mortality Risk) (0)	0.993 (0.016)	192	0.993 (0.016)	215
Time Tradeoff (Decrease in Longevity) (1)	0.987 (0.021)	204	0.989 (0.019)	183

QALYweight by QALYcode, Kruskal-Wallis $\chi 2=10.3$, $p=0.001$

QALYweight by Endpoint, Kruskal-Wallis $\chi 2=0.6$, $p=0.4$

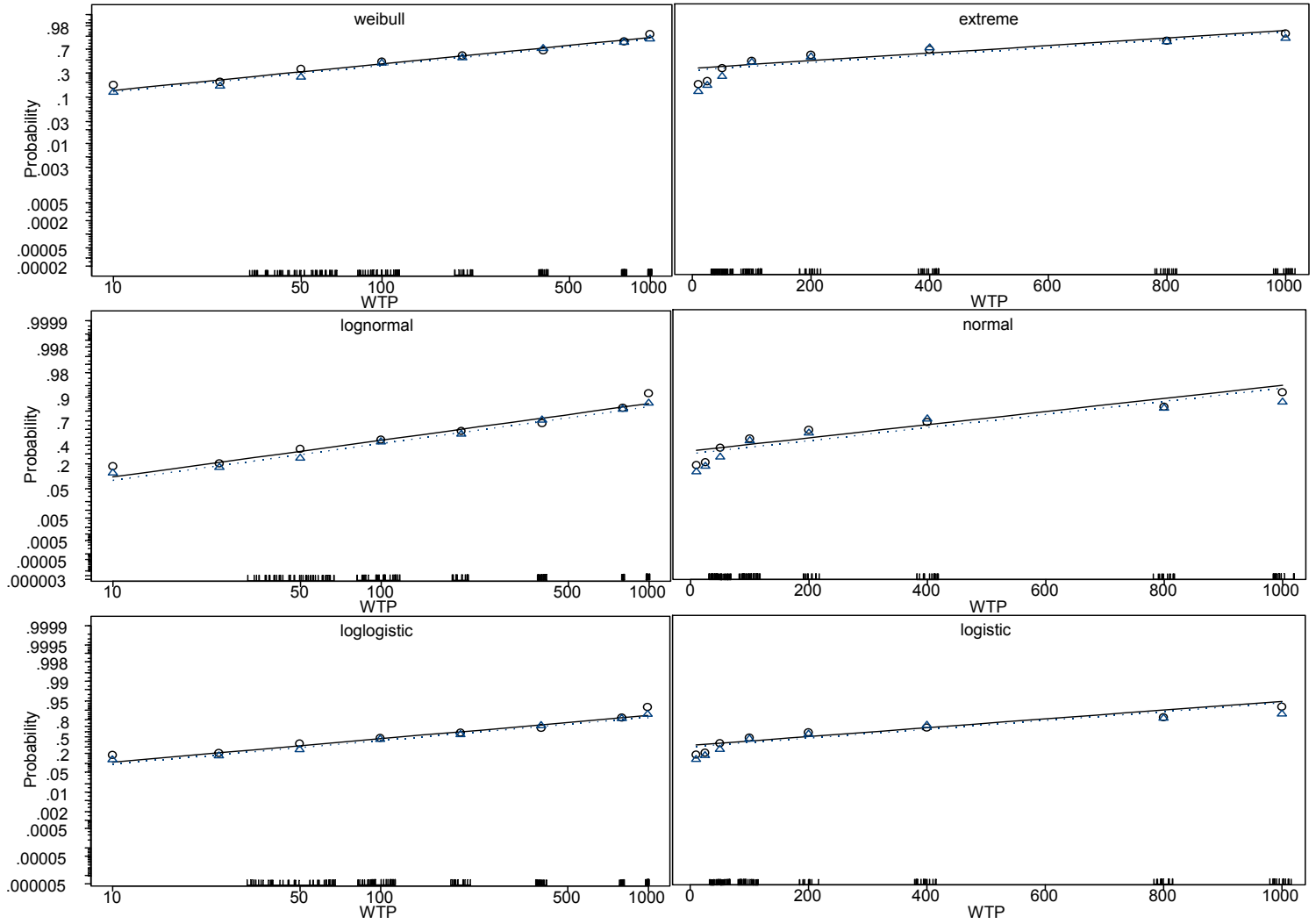


FIGURE 1: Probability Plots for the HHFirst Single Endpoint

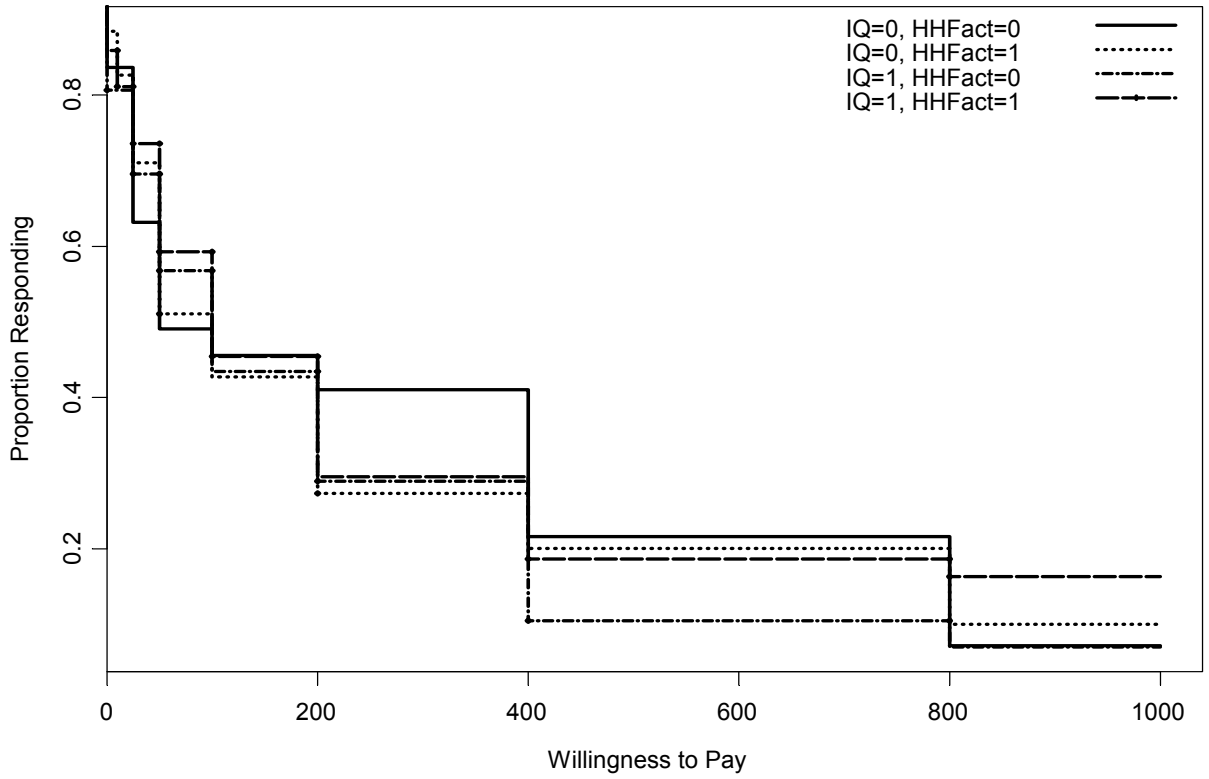


FIGURE 2: Willingness to Pay Across Risk Reductions for Human Health Endpoints

Parental Decision-Making and Children's Health

Ann Bostrom, Sandra Hoffmann, Alan Krupnick and Wictor Adamowicz[1]
With Robin Goldman, Michael McWilliams, and Jeremy Varner [2]

INTRODUCTION

Recent interest in valuation of children's health has raised many questions for how stated preference studies are conducted (EPA 2000, 2003, OECD 2006). One of the most pressing methodological questions is how parental willingness to pay (WTP) should be elicited in a stated preference survey. Typically, stated preference surveys randomly sample households and then either randomly sample adults within the household or, where pre-screened panels are used, rely on the person in the panel. The responding adult is asked to report household willingness to pay. These study designs assume that stated household WTP is invariant to who reports it or, at least, that there is no systematic bias between respondents on the basis of gender or other observable demographic characteristics.

This approach is consistent with a unitary model of household decisionmaking, which assumes that the household acts as a single decisionmaking unit, with a single set of fixed preferences and a single budget constraint (Samuelson 1956, Becker 1974). Since the 1970s, this view has been augmented by the view that household level consumption and labor supply decisions are the outcome of a bargaining process between adult decision makers in the household (Ashworth and Ulph 1981, Manser and Brown 1980, McElroy and Horney 1981). The empirical literature on alternative household models has focused on identification of departures from the unitary model using secondary household level data (Browning et al. 1994, Lundberg et al. 1997, Browning and Chiappori 1998).

Stated preference surveys, by their nature, collect individual level data. As a result, it is critical to understand the relationship between individual statements and household level choice. For example, Bateman (2005) shows that unless adults in a multi-adult household fully pool income, the standard approach of asking one adult to provide household WTP will not give an accurate estimate of household WTP. It is unclear at present whether respondents are providing their own preferences or their appraisal of the outcome of a household decision process, whether unitary or bargained. This problem may be particularly important in valuing children's health outcomes. Differences between parents' risk perceptions, risk attitudes, knowledge about and responsibility for children's health and care,

[1] The authors are respectively from the Georgia Institute of Technology, Resources for the Future, Resources for the Future, and University of Alberta. This project is funded by a grant from the EPA STAR program. We would like to thank Will Wheeler and the ORD staff, other staff at the USEPA and reviewers for the EPA STAR program for their support of this research. We would also like to thank James Bason and his staff at the University of Georgia's Survey Research Center for their work in survey administration.
[2] Research Assistants, RFF, RFF and Georgia Institute of Technology, respectively.

and control over household budget could affect individual parents' responses about their WTP to reduce their children's health risks.

These concerns suggest that elicitation of household WTP in a stated preference study, in particular eliciting parents' WTP for reductions in children's health risks, may be more complex than typically assumed in stated preference studies. To begin to sort out this complexity and ultimately help design a WTP survey of households, we conducted a study examining parental decision-making about a variety of decisions, including reducing children's health risks in the context of lead paint exposure. This paper reports on some of the findings of this study, focusing on family decision processes, leaving to another paper analysis of how parents perceive and react to decisions about reducing lead paint risks to their children.

Section one of this paper provides a review of the economics literature on household decisionmaking and of the "mental models" literature related to eliciting decision models. In section two we set out the methodology used in this study. In section three, we present results. The implications of these results for design of stated preference surveys is discussed in section four.

1. LITERATURE

Household Economics Literature

A fundamental problem for economics in studying family decision making is that modern microeconomics has a subjective, individualistic theory of value, but data are typically collected at the household level (Vermeulen 2004). As a result, even though households are micro-societies, it is difficult to infer the role that individuals play within the household. Effectively, what modern household economics attempts to do is to infer the relationship between preferences of individuals within the household and the decisions reached by the household from household level revealed preference data.

Early models aggregate individual utility into a unitary household-level social welfare function. Samuelson (1956) does this by assuming the family acts *as if* it were maximizing a weakly separable household welfare function that is increasing in individual household members' utility, $W_h = W(u_1(\boldsymbol{q}_1), u_2(\boldsymbol{q}_2), u_3(\boldsymbol{q}_3), \ldots)$. The family is assumed to allocate income across family members by consensus, $Y = y_1 + y_2 + \ldots$. Samuelson (1956) shows that if one can assume that income is distributed within the family "so as to keep each member's dollar expenditure of equal ethical worth, the family can be said to *act as if* it maximizes such a group preference function." Becker (1974) assumes the household acts *as if* a benevolent family dictator were allocating total household purchasing power among family members to maximize a weakly separable and increasing in the household head's own consumption and other family members' utility, $W_h = U_1(\boldsymbol{q}_1, u_2(\boldsymbol{q}_2), u_3(\boldsymbol{q}_3), \ldots)$. Unitary models imply an income pooling hypothesis, namely that only aggregate household income and not individuals' income affects resource allocation within the household.

An alternative approach to modeling household decisions uses game-theoretic models that explicitly take the behavior of individual household members into account. One major class of models assumes non-cooperative bargaining (Leuthold 1968, Ashworth and Ulph 1981, Browning 2000). Household members maximize their own utility taking other household members' behavior as given. The resulting intra-household allocations may not be Pareto-efficient. These models imply restrictions on observable household behavior that are not implied by the unitary household models. The second major class of models assumes cooperative bargaining (Manser and Brown 1980, McElroy and Horney 1981). Household members bargain over division of the gains of cooperation that accrue from living as a family. The bargaining power of household members and assumptions regarding which bargaining strategy is used determines the specific intrahousehold allocation resulting from the bargaining (McElroy and Horney 1981, Manser and Brown 1980). These bargaining models allow for the possibility that the source of non-labor income affects allocation of household resources, i.e., that income is not pooled. Using this modeling framework, empirical studies have shown that children's health and welfare outcomes can differ depending on whether mothers or fathers are given transfers of income (Lundberg et al. 1997, Phipps and Burton 1996, Hoddinott and Haddad 1995, Doss 1996, Strauss et al 2000). This finding is not explainable by unitary models.

A major criticism of the cooperative and non-cooperative bargaining models has been that it is not possible empirically to tell whether the household is rejecting a particular choice and or whether the assumed bargaining structure does not fit the data (Vermeulen 2004).[3] More recently, an alternative class of models that avoids this problem, called collective household models, has gained acceptance (Bourguignon and Chiappori 1992, Browning and Chiappori 1998). As in other bargaining models, individuals in collective household models maximize their own utility. But unlike cooperative and non-cooperative bargaining models, collective models assume only that the outcomes of the bargaining process are Pareto-efficient. In general, one individual in the household maximizes their own utility from the household allocation of consumption, leisure and a public good subject to similarly defined utility of other household members being greater than or equal to their reservation utility. Household allocation of resources is also assumed to be influenced by the reservation utility (Apps and Rees 1997). Reservation utility is usually suppressed in formal presentation of collective models because they are a function of wage and unearned income and are unobservable (Apps and Rees 1997). Similarly, factors in addition to price, wage and non-labor income, that are recognized to affect individual utility are generally suppressed in these models formal notation (Apps and Rees 1997). Utility is maximized subject to a pooled budget constraint with income including both labor and non-labor income. Assuming that individual utility functions are concave and the budget constraint is convex, the household's problem can be characterized as maximization of a weighted Utilitarian social welfare function subject to a unified full income budget constraint:

[3] This discussion of collective household models draws heavily on Vermeulen's (2004) review of collective household models.

$$W(\boldsymbol{p},\boldsymbol{w},\boldsymbol{y}) = \max_{\boldsymbol{q}^a,\,\boldsymbol{q}^b,l^a,\,l^b,\boldsymbol{Q}} \delta(\boldsymbol{p},w,\boldsymbol{y})u^a\left(\boldsymbol{q}^a,\,\boldsymbol{q}^b,l^a,\,l^b,\boldsymbol{Q}\right)+\left(1-\delta(\boldsymbol{p},w,\boldsymbol{y})\right)u^b\left(\boldsymbol{q}^a,\,\boldsymbol{q}^b,l^a,\,l^b,\boldsymbol{Q}\right)$$

s.t.

$$\boldsymbol{p}'\boldsymbol{q} + w^a l^a + w^b l^b \le y^a + y^b + y^H + \left(w^a + w^b\right)T$$

where, in the two adult household case, $i = (a, b, H)$ indexes individuals *a*, *b* and the household respectively, $\boldsymbol{p}$ is a vector of prices, $\boldsymbol{w}$ is a vector of wages, $\mathbf{y}$ is a vector of individual and household level non-earned income, δ are social welfare weights in the household, u^i is individual *i*'s utility, l^i is leisure, $\boldsymbol{q}^i$ are private goods, $\boldsymbol{Q}$ is a household public good, $\boldsymbol{q} = (\boldsymbol{q}^a, \boldsymbol{q}^b, \boldsymbol{Q})'$ and T is time.

In collective models, the social welfare weights are interpreted as reflecting the bargaining power or influence of an individual in household decisions. Changes in relative wages or prices could result in a change in household consumption patterns not only due to direct wage and price effects, but also because such effects could change the relative weight of individuals' preferences in the household decision. Empirically, exogenous factors that affect the standing of the individual in the labor market or in marriage, such as education or employment history, individual non-labor income, or changes in divorce or marital property law, are also hypothesized to influence bargaining power within the household (Browning and Chiappori 1998, Phipps and Burton 1998, McElroy 1990). The fact that bargaining weights may change in response to exogenous factors implies that at the household level preferences can no longer be seen as fixed as they are in the older unitary models.

One attractive feature of the collective household model is that it includes the unitary model as a special case. There are several ways this can arise. First, the welfare weight on one individual could be fixed at one. Fixed weights are unaffected by changes in exogenous changes in relative wage, employment history, or unearned wealth. Depending on the structure of the individual's utility function with a fixed weight of one, the resulting model could look like Samuelson's consensus model, Becker's benevolent dictator model, or could take other, less benign, forms. Another possibility is that the welfare weights are fixed between 0 and 1 and the utility functions take particular forms. For example, if there is no household public good and no consumption or leisure externality, then the household social welfare function will be strongly separable in individual utilities, which is again akin to Samuelson's (1956) model (Vermeulen 2004). Finally, if individual preferences are identical, the collective model collapses to a unitary model.

Apps and Rees (1996, 1997) and Chiappori (1997) extend the basic collective model to cases where there is household production of nonmarket goods. These extensions are informative in the context of children's health, which in effect is "produced" in a household as the result of family decisions. Bourguignon (1999) extends the basic modeling framework to include children as a public consumption good to adult household members. In some cases, the good produced by the household is not marketable or has poor substitutes in the market. In these cases, the shadow price of the good and the shadow price of individuals' labor in household production are determined endogenously. In these models, relative wages and prices generally can no longer be used to identify the bargaining weights or by implication the correct household model (Apps and Rees 1996, Chiapporri 1997). This, in part explains the focus on unearned

income in empirical studies using revealed preference data to test alternative household models (Lundberg et al. 1997; Doss 1996). Stated preference studies may be able to directly estimate relationships implied by alternative household models that are unobservable in revealed preference data.

Stated preference surveys, by their nature, collect individual level data. As a result, it is critical to understand the relationship between individual statements and household level choice. For example, Bateman (2005) shows that unless adults in a multi-adult household fully pool income, the standard approach of asking one adult to provide household WTP will not give an accurate estimate of household WTP. It is unclear at present whether respondents are providing their own preferences or their appraisal of the outcome of a household decision process, whether unitary or bargained. This problem may be particularly important in valuing children's health outcomes.

The above models provide a framework for deciding when it is appropriate to ask one member of the household a stated preference question or when both members need to be asked, as well as how to ask these questions and what types of supplementary information to request.

If there are differences in preferences, then it may not be adequate to survey a single household member. One likely way in which preferences may differ between spouses regards their preferences over health risks. Many studies show gender differences in risk perceptions (e.g., Finucane et al. 2000) and some in risk taking (Byrnes et al. 1999). In most cases males are found to have lower concerns about risk, or perceive risks as being smaller, than females (Davidson and Freudenburg 1996; Flynn, Slovic and Mertz 1994). In some cases differences male-female risk attitudes depend on the type of risk being examined or on more complex relationships between the risk and the individual (Finucane et al. 2000). Nevertheless, differences between men and women in their risk attitudes appear to be robust findings across various risk categories and analytical methods. Since risk attitudes affect the form of individual's utility function, gender differences in risk attitudes could lead to different responses to questions about WTP to reduce risk to children's health.

Division of responsibility for household production activities may also affect household decisions affecting children's health. There are several ways in which this could result in individual preferences mattering in a WTP study, whether the model is a collective model or a Samuelson type unitary model. Responsibility for a certain class of activity may influence the weight placed on a person's utility. One possibility might be a domain-specific dictator, as in the traditional case where, "my wife makes all the decorating decisions." It may also result in greater weight being placed on the utility of the person with responsibility for a particular activity in decisions related to that activity, perhaps because the person has gained greater knowledge about that domain. Finally, in a model with household production of a non-market good, differences in responsibility for provision of that good would lead to individuals' time constraints being affected differently. For example, it is possible that if one person has primary or sole responsibility for home repairs, that the tradeoffs that person is willing to made on removal of lead paint might differ from those of a partner who has little responsibility for home repair. Existing collective household models have assumed that bargaining weights are

invariant to the decision domain. We hypothesize here that weights vary by decision domain. Thus the bargaining weight is actually a vector of weights. The value of scalars in this vector may change with the type of decision being made. This allows for the possibility of specialization within the household. We also hypothesize that knowledge or skill in household tasks affects these domain-specific bargaining weights.

The household models also suggest that supplementary information, includes good measures of income and variables to estimate weights, needs to be collected. Welfare weights are a function of exogenous variables affecting the standing of individuals' preferences in family decisions including: information on individuals' unearned income, relative wages, education and other variables affecting individuals' prospects in the labor market. Because stated preference surveys rely on individuals' subjective evaluations and because bargaining power depends on both party's evaluation of their own and the other party's position, it is also important to know whether individuals differ in their subjective estimates of these variables.

Mental Models Literature

Choice decisions involving multiple parties, like those in a family, are more complex than individual decisions and may involve hierarchies of choices. Mental models research offers a systematic way to investigate the structure of individual and group decisions and can provide a sounder scientific basis on which to design a valuation survey.

Two decades of work in cognitive and decision science has begun to show how people represent knowledge about their decision environment in mental models (Gentner and Stevens 1983, Langan-Fox 2000). Craik (1943, p 61) described mental models as "small-scale model[s] of external reality" that people invoke and 'run' in their heads to see how to understand and explain the world. These models are associations that exist within long-term or short-term memory and strongly influence how information is retained, recalled and used in decision settings (Bainbridge 1991).

Recent studies have examined how mental models of decisionmaking in a team setting differ from those of individuals (Orasanu and Salas 1993, Adelman et al. 1986). A marriage can be seen as a team, with differentiated roles and responsibilities. Team mental models research provides a methodological foundation for eliciting mental models of joint decisionmaking from couples (Rouse, Cannon-Bowers, and Salas 1992, Daniels, de Chernatony and Johnson 1995). Researchers have long assumed that teams work better if members share mental models of team tasks and processes, and that members' mental models of both task and team process become more similar – that is, more shared - over time. Levesque et al (2001) found instead that mental models of team tasks and processes diverged over time, as team members specialized. Literature on group decisionmaking indicates that individuals in groups often defer decisionmaking power to those perceived to have more knowledge or experience in the decision context (Sorkin et al 2001).

Langan-Fox et al. (2000) found cognitive interviewing techniques, including open-ended questions followed with prompts asking respondents to elaborate, and visual card sorting, to be useful in eliciting mental models of team decisionmaking. These same methods have been used successfully to elicit mental models of individual decisions to engage in risky activities, like

smoking in adolescents (Lynch 1995) and lay mental models of indoor radon risk and risk mitigation (Bostrom et al. 1992).

In this study we elicit both individuals' and couple's mental models of lead hazards and of the couple's (dyadic) decision-making process**.** We elicited task-specific knowledge (i.e., about lead hazards), task-related knowledge (i.e., about the couple's risk decision-making), individuals' risk-related attitudes and beliefs, and knowledge of their partner's risk-related attitudes (cf. Cannon-Bowers and Salas 2001). The approach extends mental models research used in other risk domains (e.g., Bostrom et al. 1992, Morgan et al. 2001) by building on team mental models research (Levesque Wilson and Wholey 2001, Mohammed and Dumville 2001).

2. METHODOLOGY

We conducted in-person interviews with thirty-five couples (70 individuals). Samples of this size have been found adequate to capture much of the conceptual variability in a substantive domain (Morgan et al. 1992). Each spouse was first interviewed individually (all couples in the sample happened to be married); spouses were then brought together and interviewed as a couple. Finally spouses were again separated and asked to complete a written questionnaire, which characterized their decisionmaking styles, took sociodemographic information, asked numerous questions about their relationship and attitudes towards risks in general and lead paint exposure, in particular.

This survey included three strategies to assess parental decisionmaking: characterization of direct statements by parents of how they make decisions; analysis of responses to closed-ended questions about decisionmaking and factors hypothesized to affect decisionmaking in the literature, and finally; examination of hypothetical decisionmaking about lead paint mitigation.

The study drew from the population of two-parent households in Atlanta, Georgia, with children under the age of 7, living in housing built before 1979. We limited the population to owner-occupied housing. Including rental housing would increase the heterogeneity of the sample by raising additional issues of control over abatement interventions, and by changing the relevance of control options. Given a small sample size, a decision was made to control for family structure to reduce heterogeneity. U.S. Census of Housing data was used to identify neighborhoods in the Atlanta, Georgia Metropolitan Statistical Area with housing stock built before 1979. Households were sampled from phone number lists by the Survey Research Center at the University of Georgia and screened for appropriate characteristics in initial phone contacts. The first fifteen interviews were conducted by research assistants at a central location at Georgia Institute of Technology. Because of difficulty in recruiting couples to travel to the interviews, the final fifteen interviews were conducted in couples' homes, by the Survey Research Center interview staff.

A semi-structured interview protocol was used to investigate parental decision-making behaviors and their mental models of lead paint risks. Prior to the interviews, each spouse was asked to write down three recent major children's health decisions. Interviewers selected the highest-ranking jointly mentioned decision as a focus for the first part of the individual interviews. The

individual interviews began with open-ended questions exploring the decisionmaking process involved in the couples' most recent major children's health decision. Follow-up prompts were used to assure that issues such as what the problem was, what decision was reached, who identified the problem, who was involved in the decision, whether prior discussion took place, who initiated the discussion, what factors were considered, how the respondent felt about the decision, how their spouse felt about it, and whether this was a typical decision. Open-ended questions were used to ask about differences between this decision and more routine purchase or home repair decisions. The next section of the interview dealt with children's environmental health problems and focused on parental awareness and level of concern about lead paint hazards compared to other environmental hazards. Finally, each spouse was presented with a hypothetical lead paint decision scenario and asked to talk through what they thought their family would do. Follow-up prompts were used to assure that information on the information desired, factors considered and role of cost in the family decision was collected. After a break, spouses were interviewed as a couple.

The couple's interview followed much the same protocol as the individual interviews, except that in the hypothetical lead paint decision, instead of eliciting possible health effects and mitigation options, the couple was given a list of specific effects and options. They were asked to sort these by seriousness of concern, effectiveness and likelihood that a mitigation option would be selected.

Finally, the spouses were again separated and asked to fill out a written questionnaire (see Appendix I). This questionnaire included questions about household decisionmaking styles in various domains (e.g., home decorating and home repair), basic demographic information, homeownership, education, employment and commitment to the labor market, income, household financial management, time spent in various household production activities, division of responsibility for specific types of family decisions, beliefs and attitudes about children's environmental health risks, and knowledge about impacts of lead on children's health.

The written questionnaire also included a set of questions on marital adjustment, the Dyadic Adjustment Scale (DAS). The DAS is a 32-question instrument developed by Spanier (1976) to assess the quality of the relationship perceived by married or cohabiting couples. The DAS remains the most frequently used instrument with different groups of participants and cultures for assessing the quality of married life (Casas & Ortiz, 1985; Crane, Allgood, Larson & Griffin, 1990; Shek, 1994). The items for the DAS were those chosen out of an initial pool of 100 that (a) were normally distributed; (b) discriminated between married and divorced people; and (c) loaded highly on one of four factors (Dyadic Consensus; Dyadic Cohesion; Dyadic Satisfaction; and Affectional Expression). Response scales differ across the questionnaire, with the consensus items including verbally anchored response scales that represent the extent of agreement or disagreement between the spouse and his or her partner for each item (from always agree, to always disagree; or from all of the time to never). The total score is the sum of scores on all items, ranging from 0 to 151. The scale scores have been found to have good content and construct validity (Spanier, 1976). Spouses with scores below 98 are classified as discordant (Eddy et al., 1991; Jacobson et al., 1984).

Spanier built the DAS on four subscales, one of which, the dyadic consensus subscale, is particularly relevant to decisionmaking. The dyadic consensus subscale consists of thirteen items assessing spousal agreement on issues ranging from, for example, handling family finances, household tasks and amount of time spent together, through friends, ways of dealing with parents or in-laws, religious matters, major decisions, and philosophy of life. Two of these items (on major and career decisions) are sometimes used as an alternative consensus subscale (Busby et al., 1995).

3. RESULTS

In this section, we provide both qualitative and quantitative results concerning decisionmaking processes across the couples. The former are drawn from the open-ended oral parts of the interviews; the latter from the written survey. The former as of this writing cover 19 couples. The latter cover 35 couples.

Qualitative Results

To initiate our personal interviews with parents, we asked each parent to list three recent major child health decisions, or family decision affecting their child. We then selected the most important of these listed by both parents independently. The individual and couple interviews each opened with a request that the parents describe this decision: "Could you tell me about [the most recent major child health] decision that your family made?" The health decisions discussed by the nineteen couples included vaccination decisions (4 couples), toothache, earache or ear surgery (4 couples), accidents (skiing, falling through a window), illnesses (asthma, fever and cold, food poisoning), what to do about a bleeding birthmark, and choices about summer camp, high school, and speech therapy. A fourth of the couples had made the decision in question within the previous six months, another fourth within the previous year.

Several features of their responses are of interest, including how they structured the decision, and what kinds of factors they took into account in making it. To learn more about how they structured the decision, we asked whether they had discussed the decision at the time it was made and/or prior to that time, and who had initiated those discussions. All couples said that they discussed the decision, and most had also discussed it previously, for an hour or less. In almost all cases, couples reported that the mother had initiated the discussion that led to the decision. The two exceptions were a sole father-initiated discussion, and one couple who initiated the discussion mutually. In the couple interviews, the couples also reported that the mother usually initiated such discussions. All of the couples reported having agreed with the decision.

When asked what factors they took into account in making a major health care decision, in this case regarding a severe, acute onset ear ache, one couple [4C] responded as follows:

> *Mother:* "just wanted to be sure that she was, that we took care of it. We wanted to be sure that… She could not go in pain. We had to do something. We had a fear of long term effects of all these burst eardrums."

Mother: "That's about it"
Interviewer: Are there any other factors?
Father: "I am sure we thought about the financial part of it"

Another couple described their child's seizures and epilepsy, and the difficult decision they had to make whether or not to give her medicine to control them. When asked by the interviewer to talk about how the decision was made, the father responded first:

Father: "Well, she [mother] discussed it with me, she did the research on the internet.
Found out exactly what the medicine could do and how it would help her [daughter]… so"
Mother: "That's after the, a, the neurologist, you know, discussed it with me. I went home and looked it up, you know… the internet is a great thing!"

As these conversations illustrate, parents' reports of these decisions emphasize the urgency of many child health decisions, the empathy parents feel with their children when they are in pain, that information is usually incomplete, but both the internet and a variety of experts and friends can be called on to fill in gaps. However, the data suggest that the majority of couples chose the plan of action that was initially considered or most common.

In terms of learning about the viability of a WTP survey of parents about their children's health, we were concerned that cost would not be a factor for a significant share of spouses. The following comment from one mother illustrates our concern:

A couple discussed a decision to take their child to see a specialist about their child's persistent cough, which was not clearing up. When the interviewer asked the couple "What were the factors considered in discussing this decision?" the mother [21C] replied: "When it comes to your kids, there aren't any factors. Their health is the most important thing. Cost, nothing, that doesn't matter to me."

Yet, the majority of couples said that they considered the quality and effectiveness of the decision alternatives, for example, the quality of the hospital to which they could take their child, as well as cost. Eleven couples mentioned costs in their unprompted description of the decision process they nominated in the beginning of the survey and one mentioned it after being prompted. However, no couple reported having considered borrowing ability.

Later in the survey, spouses were asked whether cost would play a role in what to do about lead paint, assuming they found high levels of lead dust in their house. Most who responded to this question answered affirmatively – 15 wives and 15 husbands said yes, 2 wives and 2 husbands said no. Further, when asked if there were conditions under which they would choose a cheaper and less effective option, 9 of the 15 wives and 10 of the 13 husbands answering said yes, suggesting that a majority but not all of the spouses are willing to think about tradeoffs.

Descriptive statistics

For each question in the written questionnaire, there are six sets of statistics: the husband answering for himself, the husband answering about his wife, the wife answering for herself, the wife answering about her husband, the answers of the husband for himself and the wife for herself averaged across the couple, and a variety of statistics at the couple, rather than the individual level (Table 1). These latter statistics permit us to look at the degree of agreement in answers across the spouses. Disagreement in responses from spouses of some form is a necessary condition for it to matter which spouse responds in a stated preference survey. Many disagreements are what might be termed "mild." The husband says his wife does most of an activity (like helping the child with homework), the wife says she does all of it. Other disagreements are more substantial, for example, if the wife were to maintain that she does all the helping and the husband were to say he does all of it. For factual questions at the spouse level, we assume the husband's (wife's) answers for himself (herself) are true or reliable. For questions at the couple level, we will use the average answers in further analyses.

Demographics. The sample is younger and more educated than the general population. The average age of men respondents was 36; the average for women was 35. Respondents' ages ranged from 26-45 years old (table 1). On average respondents had 16 years of education. African Americans, but not other minorities are well represented. About two-thirds of the couples were white and the rest were black. Only six percent of respondents were previously married. Most (57%) have two children, with up to six children (in one family). Because the interview protocol required families to have at least one child 7 and under, 73% of children in the study fit this criterion. All were homeowners, in homes built 1979 or earlier, consistent with the sampling protocol, with average tenure 6 years. Six percent of the couples had been married previously. In general, there were minor disagreements among couples on virtually every demographic question except having been divorced. Most of these are of a level that would qualify as measurement error, but it is interesting to see that this kind of error is present even on basic factual information about the families.

Employment and Income. Employment and income patterns can affect the weight of individual preferences in family decision. Ninety-one percent of husbands and 63% of wives in our sample were employed (table 1). Most husbands (64%) said they worked more than 40 hours a week. Most wives who worked, reported working 20-39 hour per week range. Not surprisingly, the husbands' contribution to family income was far higher than the wives: 73% vs. 27%, although seven wives (of 34 answering) contributed over 50% of household income. Median pre-tax, household income (in 2004) was between $60,000 and $74,000. However, the wives thought mean family income (in 2004, before taxes) was a bit lower than the husbands did: $79,860 vs. $83,290.

Spouse's perceptions of their own and their spouse's relative contribution to family income are also theorized to affect household decisions. Husbands and wives were each asked what percent of household income they and their spouse contributed. We see from the table 2 and figure 1 that on average, husbands and wives have the same perception of the amount of income the wife is contributing. This happy *average* state of affairs masks significant

differences in perception. From the husband's perspective, the worst cases in this study are a husband who thinks his wife is contributing 40% less than she thinks she is and a husband who thinks his wife is contributing 20% more than the wife thinks she is. From the wife's perspective, one wife thinks her husband contributes 75% less than he thinks he does and another wife thinks her husband contributes 55% more than he thinks he does.

Some of these disagreements may simply be lack of knowledge about what total household income is. On average, husbands think household income is $3,400 greater than wives do. But again, at the extremes, one husband thinks their combined income is $45,000 less than the wife thinks at the other extreme, one husband thinks total household income is $30,000 greater than what the wife thinks it is.

Because attachment to the labor market figures heavily in the empirical literature on household bargaining, four additional questions were commonly asked to gauge degree of desire to working outside the home: whether the spouse would prefer to stay at home with the children, whether the respondent would prefer that their spouse stay home with the children, whether the spouse's career is more important than the respondent's, and whether the spouse feels he or she should be the breadwinner in the family. These questions evoke very different responses in husbands and wives, while the husband and wife generally agree with one another's assessments. In general, wives want to stay home with their children and do not want their husbands to do so. Husbands want their wives to stay home with the children, but have a range of feelings about themselves, *not* strongly skewed against staying home. Both wives and husbands generally agree that the husband should be the breadwinner. However, there is close to indifference about whose career is most important, with an edge to the husband's, given by both the husbands and wives.

Decisionmaking. This study focuses on financial and health decisions because these are relevant to children's health valuation. Couples in the study exhibit three general approaches to household financial management: joint management, separate management, and allocated or assigned management. In allocated or assigned management, one spouse has a housekeeping or personal spending allowance and the other spouse manages the rest of the household money. Within couples' there is general agreement about which model fits. Most (73%-79%) of the couples managing their money jointly. Most of the rest are in the assignment mode (table 3). Later in the survey, respondents were asked to make a general characterization of who makes decisions in the household and then were ask about division of decisionmaking responsibility about in specific decision contexts. Self-reporting on decision style may lead to an over-reporting of "joint" decisionmaking because people may want to view themselves as conforming to a norm that family decisions should be made jointly.

In their general characterization of who makes decisions in the household, both 88% of wives and 88% of husbands said that decisions were made jointly, although there was some disagreement at the couple level, as discussed below. Only 9% of the husbands said they made more of the decisions. Once the context was made specific, these percentages sometimes changed. Most couples make financial decisions jointly but some wives (23%) and husbands (18%) said that the husbands make more of these decisions. No men said their

wives made more of the financial decisions. For decisions involving major purchases, which is another way of describing financial decisions, 79% of wives and 82% of husbands say these decisions are made jointly. The remaining respondents are split equally in saying the wives or themselves make more (or all) of these decisions.

In their characterization of specific decision domains, there was evidence of specialization. This was particularly prominent in the context of children's health, where only 26% of wives and 32% of husbands say that decisions are jointly made. 71% of wives say they make the decisions about doctor visits for their kids, for instance, with only one saying her husband makes more of these decisions. As a group, the husbands generally agree with their wives on this issue.

Couples differ in the extent to which they agree about how they make specific decisions or manage finances. To aid this discussion, we define the following terms: Joint (spouses agree that they make decisions jointly), Agree (spouses agree that one or the other makes the decision), Disjoint (where one thinks they make decisions jointly and the other thinks the situation is different), and Disagree (one thinks one makes the decision and the other thinks the other makes the decision). Considering the general decisionmaking question first, 26 couples agreed that decisions are made jointly. None agreed that one spouse or the other makes all, most or more of the household's decisions. The rest of the responses can all be classified as disjoint. Childcare is one of the domains with the most disagreements: 13 couples agree that they make joint decisions, and 7 couples agree that the wife makes more decisions. The remaining 14 couples that answered this question are disjoint.

Allocation of time. The amount of time different individuals spend on certain activities may affect help explain patterns of decisionmaking. It is clear (table 4) that financial tasks are shared fairly equally in nearly all households, while husbands dominate only home repair and renovation in terms of the time spent on these tasks. For all other tasks, wives spend more time than their husbands do and the spouses generally agree on this. In particular, wives spend more time than their husbands caring for children. One interesting area of disagreement (or disjointedness) between couples concerns time spent helping children with their homework. Husbands think they do more of this activity than their wives think they do. There is also disagreement between spouses about who takes the kids to the doctor. While 22 couples agree that the wife spends more time taking children to the doctor, 13 couples are disjoint. A similar situation with is found for homework, cleaning the house, spring-cleaning, decorating, major purchases and financial management.

Marital Adjustment. To test if marital adjustment affects decision making in couples, we use the 32-item DAS, as described above. Scoring rules differ by question (see Appendix I). Unhappy couples have been normed to be those with a score of 98 or less. In this study both spouses completed the DAS questionnaire. Husbands' scores range from 61 to 138 with an average of 113 (table 5, figure 2). 12% of the husbands rate their marriage with a 98 or lower. Wives' scores have a wider range (59-143), but the average is the same as the husbands at 113. Only 9% of the wives scored their relationship 98 or less. In some couples, spouses have different scores on this 32-item scale. In eight couples, the wife's marital adjustment score is 10 points or more above her husband's and for another six couples the

husband's score is ten or more points greater than the wife's. All told there is only one couple with both spouses rating their relationship at or below the cutoff score of 98.

There is a subset of 13 questions in the DAS in which respondents are asked how often they agree or disagree on specific decision areas such as handling family finances, religious matters, recreation, dealing with parents or in-laws, etc. Among these thirteen decision areas, there is almost no area in which one spouse says that they always or almost always agree and the other spouse says they always or almost always disagree. We do find some serious for which both spouses acknowledge disagreements: in-laws, the amount of time the couple spends together, leisure interests and activities, and career decisions.

In this series of questions, there are also more factual questions, such as "How often do you and your spouse quarrel?" Differences in spouses' responses on these questions could be problematic because they indicate different perceptions about the quality of the marriage. For these questions, serious differences in couples responses are defined as two or more points of difference on the five-point scale in which 0 indicates poor marital adjustment and 5 indicates high. The questions "How often do you engage in a stimulating exchange of ideas?", "how often do you calmly discuss something?", and "how often do you work on a project together" provoked some serious differences between spouses' responses. There were also serious differences in responses on yes/no questions including the question about whether being too tired for sex has caused problems. Twelve couples had one spouse say Yes and the other say No. Six couples both said Yes and fourteen couples both said No.

Attitudes Towards Risk. Another factor that could influence decisionmaking is attitudes towards risk. To gauge such attitudes about lead exposure, in the oral section of the survey we asked spouses whether they were worried about lead paint. Fewer husbands (6 yes, 13 no) said they had worried about lead paint than wives (11 yes, 8 no).

In the written survey, we placed asked respondents to rank eight health risks, including lead paint, according to various dimensions of qualitative and quantitative risks. These other health risks included air pollution, climate change, radon, small pox, small pox vaccine, anthrax and influenza.

The results are voluminous, but the main ones are: (i) flu and air pollution are viewed as the most risky with lead in the middle of the group, assessed equally by the wives and husbands and viewed by both parents as a bigger risk to children than to the overall population. Climate change was the most "unknown" risk, anthrax the most "serious," and climate change had the longest lead time. Air pollution is viewed as the risk causing the most exposure. For lead, wives think exposures are more widespread than husbands do, as we saw in oral responses.

In table 6, we show detailed results for the qualitative risk dimension "controllability" across the eight risk categories for husbands and wives. Here, we supply the percentage of each gender ranking each risk as most controllable down to the least controllable. We find that lead paint is seen as the most controllable risk (not surprising, given the alternatives) and that there is more disagreement about this across the husbands than the wives.

Knowledge About Lead. As noted above, knowledge about a topic of concern may help explain patterns of specialization in decisionmaking. It may also be an indicator of ability to or interest in searching out or absorb information relevant to decisionmaking. This ability may help explain patterns of specialization. We looked at this issue both in the oral and written parts of the survey. In the oral section, we asked spouses separately and the couple together to consider a hypothetical decision concerning lead paint mitigation: "How much do you know about the health risks from lead paint (own knowledge); how much do you think your spouse knows?"

Interestingly, both husbands and wives thought their spouses knew more about health effects than they did, on average. On this question, 7 couples agreed on how much the wife knows. The average score of women on their own knowledge was 2.8, with the husbands giving their wives an average score of 3.3. Their responses were positively correlated, $r = 0.62$. There was somewhat less agreement on how much husbands know, with only 5 couples giving the same the estimates for the husband's level of knowledge, with husbands rating their own knowledge at 2.5, and wives rating their husbands' knowledge at 3.2 on average, (one-tailed paired *t*-test, $p < 0.05$), $r = 0.11$.

In the written part of the survey, we asked thirteen true-false-no opinion questions to test for knowledge about lead and its effects. Overall, the wives as a group are more often right than the husbands (if we simply sum up right answers over all 13 questions) (table 7, figure 3). On average, the wives got 10 questions right, the husbands nine. The questions most frequently missed by both groups are whether lead absorption is greater when a person has iron deficiency (TRUE), and whether lead exposure can lead to hypertension (TRUE).

It is plausible that when one spouse has more knowledge about a problem than another, that spouse might take a greater role in the decision over what to do about that problem. We therefore tallied up the number of times a husband had a different answer to the lead knowledge questions than the wife did, and in which direction. We found that for six couples the wife outperformed the husband, being correct on four or more questions her husband missed. Correspondingly, we only found two couples where the husband outperformed the wife on four or more (four) questions.

Regression Analysis

Below, two types of regression analyses are presented. The first explains couples' decisionmaking in each of five decision domains specifically relevant to children's health valuation: child doctor visits, childcare, paying bills, family income management and household purchase decisions. The second pools responses to all ten decisionmaking domains examined in the survey to explain spouses perceived decisionmaking. Both analyses feature observations at the spouse level (rather than the couple) because husbands and wives may disagree. The data set includes responses from 35 couples, or 70 spouses. Table 8 presents definitions of the dependent and explanatory variables.

Hypotheses. As discussed above, recent developments in the economic theory of household decisions focus on the weight that different household members' preferences have in household decisions, the determinants of that weighting, and whether decisions outcomes are the result of bargaining. The fundamental feature of these models is that factors that affect spouses' options outside the marriage influence their bargaining power in household decisions. On this basis, we would expect that relatively exogenous factors related to employment decisions such as: relative income or wage rates, relative education levels, the level of commitment to work, unemployment spells and the extent to which both spouses work full time will influence household decisions. We would also expect that factors that reflect the quality of communication in the marriage, here measured by the DAS, could affect household decisions. Finally, literature on group decisions suggests that relative levels of knowledge about a problem affects who influences group, i.e., family, decisions.

A fundamental empirical problem for the household literature is that the primitives of this model are unobservable. Most commonly, empirical work testing for the appropriateness of alternative models has taken a revealed preference approach relying on household level consumption and labor supply outcomes as measures of the outcomes of household decisions (Phipps and Burton 1998, Lundberg et al. 1997, Strauss 2000). Several studies have used purchases that benefit only specific members of the household, like women's or children's clothing purchases, as a measure of the influence of those individuals' preferences in family purchase decisions. Dosman and Adamowicz (*forthcoming*) elicit individual and couples choices in a conjoint stated preference survey and use this to estimate implied household welfare weights. In the study presented here, the observable outcome is who plays a role in household decisionmaking. We define the dependent variable as whether a decision is made jointly or by one of the spouses alone. To explain variation in this dependent variable, we use data collected on a wide range of independent variables that household economic models and mental models suggest could influence the role of individual preferences in family decisions.

Assuming that an individual's preferences play a greater role in the household decision if they are involved in the decision, we can use the distinction between joint and individual decisionmaking as an indicator of whose preferences carry weight in household decisions. Obviously this conclusion might not hold if altruism plays a strong role in the way families make decisions. For example, Becker's family dictator takes the utility of other household members into account. This conclusion also may not hold if the couple agrees to specialize in decisions over various domains. As a result, this work should be viewed as a means of getting at stylized facts about household decisions that will be tested more rigorously in our planned stated preference survey research.

We have specific hypotheses about how some of our independent variables affect the likelihood that decisions are made jointly or by an individual spouse. For instance, we expect that where income contributions of the spouses are more equal we are more likely to see joint decisionmaking. For other variables, we do not have hypotheses about the direction of an effect but do expect that an effect could be present, for example, for race, income, education or age. Note also that some of the explanatory variables, such as time allocation, are themselves endogenous. At this point in the analysis, we have not attempted to estimate

more complex models to account for this. We also recognize that, ideally, we should use multinomial logit techniques to analyze these data, as decisionmaking could be joint, the husband's lead or the wife's lead. However, there are not enough observations about all three options for any decision variable to justify this more complex approach.

The fundamental question we seek to address is whether it matters for stated preference survey research which adult in the household is interviewed. Even if a unitary model properly describes household behavior, it would matter who researchers interview in a stated preference study if there is specialization of responsibility for and knowledge about particular household decisions. Responsibility could vary by domain. This would suggest that the less the difference in the amount of time spouses spend on a household task, the more likely it is that decisions about that domain would be made jointly. We hypothesize that the larger the number of children, the more likely it is that spouses will specialize between home and market labor and the less likely it will be that decisions about children will be made jointly. Another way in which such specialization might arise is if one of the spouses specializes in information gathering (Sorkin et al. 2001). As a proxy for knowledge levels we use correct responses to a set of knowledge questions about lead paint hazards as an indication of information gathering performance. The less the difference in this variable, the more likely that decisions will be made jointly.

Regression results. As noted, due to small sample size we construct a bivariate dependent variable from the multivariate variables on who makes decisions. We restrict decision outcomes to a dummy variable taking a value of zero for joint decisions and one for "___ makes more of the decisions." For many decision domains there is a strong gender bias in decision responsibility across the 35 couples. So for example, for childcare decisions no respondents said husbands made this decision, while 28 said the wife made the decision. Forty respondents said childcare decisions were made jointly. In this case we dropped the observations for "husband makes most child care decisions" and constructed a binary variable with 0 for joint decisions and one for "wife makes more of the decisions." A similar pattern was followed for decision domains where few wives made more of the decisions. For decision domains that did not exhibit strong gender bias, we addressed the small sample problem by constructing a dummy in which zero indicates a joint decision and one indicates that one spouse or the other makes more decisions. In this case no observations are dropped. For income pooling, the dependent variable is defined as 0 if the couples manage their financial accounts jointly and 1 if they do not.

Table 9 provides the results from logit regression on two child-related decisions and three financially-related decisions: taking children to the doctor (1 = wife; 4 husband decision makers dropped); childcare decisions (1= wives make decisions; none dropped), paying bills (1 = either husband or wife makes decision; none dropped), financial decisions (1 = husband; 2 wife decision makers dropped), and household income management (1 = either husband or wife makes decision; none dropped).

Household income and whether one or both spouses spend time on a task are significant in explaining both child-related and financial decisions. The higher household income, the less likely it is that decisions about taking children to the doctor, childcare and finances are made

jointly. For decisions about taking children to the doctor, paying bills, and managing finances, decision are more likely to be specialized when one or the other spouse spends most of the time on this task. Years of education is significant in explaining some of the finance decisions, but not the child-related decisions. The more years of education, the less likely it is that couples pay bills jointly and the more likely it is that they pool their income. Oddly, the greater the difference in education, the more likely childcare decisions will be made jointly. The more children there are in the family the more likely it is that decisions on whether to take children to the doctor and bill paying are made jointly. Number of children is not significant for any other decisions.

There are a cluster of independent variables related to employment and income. Total household income affects both child-related decisions as well as general financial management decisions. The higher the income, the more likely it is that couples will not make child-related decisions and general financial decisions jointly. The wife being employed is associated with joint childcare decisions. However, given that the wife is employed, the greater the wife's share of the household income, the *less* likely it is that childcare decisions will be made jointly. Also the greater the wife's share of the household's income, the more likely it is that financial decisions will be made jointly. The lower the index of commitment to working in the labor market, the less likely it is that decisions to take children to the doctor will be made jointly. The implication is that where wives are at home and happy about it, they are more likely to specialize in making child medical decisions.

There are a number of results that are difficult to explain. It is not clear why a high score on the marital consensus subscale would be associated with joint decisions about taking children to the doctor, but would not be significant for other decisions. We use knowledge about lead gained through earlier parts of the survey (or from prior knowledge) as an indicator of an interest in and ability to obtain knowledge in general. It is not clear why this would be associated with joint decisions on taking children to the doctor and paying bills, but not other decisions. The index of beliefs about controllability of risks to children is associated with non-joint decisions about paying bills. It is conceivable that there is some correlation between a sense of being able to control risks with a willingness not to have as tight joint control over bill paying, but this is speculation on our part.

Table 10 presents results of logit regressions on the pooled set of decisions in all ten decisions domains. As noted above, much of the household decisionmaking literature assumes that decisionmaking models are invariant to the decisionmaking context or domain. With our data we can test this proposition by lumping together all the decision-making responses (10 domains per survey) to create a 700-observation dataset. As before, dummy variables were created to indicate whether the decisionmaking model was classified by the respondent as joint or other.

First, decision domain matters. All dummies for domains (but one) are significant (against the childcare default dummy and the show significant differences in some instances with one another, clustering in two groups, one where joint decisionmaking is more likely and the other where either spouse specializing is more likely. Once domain is controlled for, then the effect of gender on decisionmaking style (which we see in chi-square tests) is eliminated.

A number of patterns identified in the individual decision analysis become even clearer in the pooled analysis, controlling for domain. As in the individual decision domains, higher income is associated with specialization, however, the effect could only be detected when income is included as a categorical variable (above or below median income) and not when it is included as a continuous variable. More years of education are also associated with specialization. Age is associated with a higher likelihood of joint decisions, but only once we use the income dummy variable. The more children a couple has, the more likely they specialize. Similarly, holding constant household income, age, number of children, education and marital consensus, households in which wives are employed also specialize. Finally, all else constant, the higher the DAS subscale score for marital consensus, the more likely it is that decisions are made jointly.

4. CONCLUSIONS AND IMPLICATIONS

The major concern of this paper is whether asking WTP questions of one parent in a two-parent household will lead to an accurate representation of household willingness to pay. This paper does not directly address this issue, in the sense that we do not ask different parents their household WTP and compare their responses. That is our next step. This paper addresses a prior question, although one that has implications for a WTP survey, which is whether decisions in a variety of domains are made jointly or there is specialization by spouses and what factors drive this difference. We infer that if spouses in the household specialize in decisionmaking then asking different spouses a WTP question is more likely to lead to different answers.

This investigation is informed by both the economics literature on household behavior and the mental models literature on group decisionmaking. It looks to the economics literature for variables that are expected to influence the relative role of different spouses in household decisions. It looks to the mental models literature both for factors that influence the role of individual's in-group decisions and for methodology to systematically study couples' decisionmaking processes.

In general, decisionmaking style varies by domain and is affected by variables that are expected from theory to contribute to power (welfare weights) in the relationship, such as income share, wife employment status, work commitment, and differences in education. From the literature on mental models, as well as from an interpretation of the household behavior literature, we also expect and find evidence for effects of domain knowledge, time spent in the domain, and marital consensus.

We also learned several lessons for a future WTP study. The most important is that the majority of couples appear to consider cost in their major decisions about children's health and, specifically, in response to our hypothetical question about decisions in response to a finding of high lead levels in the home. They were willing to make tradeoffs with effectiveness and cost. Another lesson is that there are gender differences in risk attitudes and risk perceptions, e.g., wives think lead exposure is more widespread than husbands and

are more worried about its impact on children. However, they both view the lead paint problem as equally controllable. A further lesson is that it is a viable strategy to administer a survey to couples and spouses separately.

There are many caveats to these conclusions. The sample is too small to do a more thorough test of decisionmaking styles using MNL techniques to capture the three styles: husband decides, wife decides, joint decision. With a larger sample we could remove the ambiguity of the "other" answer. Future work will involve almost ten times this sample size. In addition, we will need to address endogeneity issues associated with some of our explanatory variables. Further work is also needed on understanding differences between spouses responses – both on opinion and facts – across couples. For instance, spouses seem to have significant differences in perceptions about what household income is and different perceptions about their own and their spouses' income share.

Table 1. Descriptive Statistics of Selected Variables.

	Husbands				Wives				Total			
	Mean	Std Dev	Min	Max	Mean	Std Dev	Min	Max	Mean	Std Dev	Min	Max
Age	36.21	6.08	30.50	50.50	34.56	6.38	21.50	50.50	35.39	6.25	21.50	50.50
Race (Non-White=1)	0.29	0.46	0.00	1.00	0.31	0.47	0.00	1.00	0.30	0.46	0.00	1.00
Previous Marriage (%)	0.14	0.36	0.00	1.00	0.06	0.24	0.00	1.00	0.10	0.30	0.00	1.00
Number of Children	2.26	1.04	0.00	1.00	2.26	1.04	1.00	6.00	2.26	1.03	1.00	6.00
Age of Children	5.79	4.84	0.25	28.00	5.58	4.89	0.00	28.00	5.69	4.85	0.00	28.00
Age of Oldest Child	7.68	5.85	0.58	28.00	7.52	5.90	0.58	28.00	7.60	5.84	0.58	28.00
Education (Years)	16.06	2.45	12.00	19.00	16.09	2.23	12.00	19.00	16.07	2.32	12.00	19.00
Education Difference (Husband - Wife)	-	-	-	-	-	-	-	-	0.00	1.97	-4.00	5.00
Absolute Value	-	-	-	-	-	-	-	-	1.23	1.53	0.00	5.00
House Built (Year)	1954	22.24	1903	1979	1954	21.32	1907	1979	1954	21.61	1903	1979
Years in House	5.50	3.15	0.08	11.50	6.40	5.26	0.08	30.00	5.96	4.34	0.08	30.00
Employed (%)	0.91	0.29	0.00	1.00	0.63	0.49	0.00	1.00	0.77	0.43	0.00	1.00
Full-time (%)	0.74	0.44	0.00	1.00	0.17	0.38	0.00	1.00	0.46	0.50	0.00	1.00
Part-time (%)	0.14	0.36	0.00	1.00	0.46	0.51	0.00	1.00	0.30	0.46	0.00	1.00
Total Household Income ($000)*	83.29	35.15	22.00	142.00	79.86	35.49	22.00	142.00	81.57	35.11	22.00	142.00
Contribution to Income (%)*	72.61	27.52	0.00	95.50	26.51	32.56	0.00	95.50	49.90	37.85	0.00	95.50
Dyadic Scale of Marital Adjustment (32-Item)	113.06	15.95	61.00	138.00	112.94	15.25	59.00	143.00	113.00	15.49	59.00	143.00
13-Item DAS Scale	48.78	8.25	25.00	63.00	49.73	4.36	37.00	61.00	49.26	6.54	25.00	63.00
Index of Work Commitment	2.32	1.36	0.00	4.00	0.32	0.53	0.00	2.00	1.32	1.44	0.00	4.00

*Total Household Income and Contribution to Household Income values are created by taking the midpoint of respondent-selected income intervals.

Table 2. Income Related Disagreements.

	Husbands View		Wives View		Difference Within Each Couple (Husband-Wife)			
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Min	Max
Husband's Contribution to Income (%)	72.61	27.52	74.00	28.46	-0.44	18.52	-55.00	75.00
Wife's Contribution to Income (%)	28.68	31.61	26.51	32.56	0.77	10.87	-40.00	20.00
Total Household Income ($000)	83.29	35.15	79.86	35.49	3.43	14.59	-45.00	30.00

Table 3. Decisionmaking and Income Pooling Statistics.

	Husbands View (%)			Wives View (%)			Couple-Level Agreements (%)			
Decisionmaking	Husband DM	Wife DM	Joint DM	Husband DM	Wife DM	Joint DM	Agree Joint	Agree Non-Joint	Disagree	Disjoint
DM: General	9	3	88	3	9	88	76	0	0	24
DM: Home Repairs	50	3	47	47	3	50	38	41	0	21
DM: Decoration	6	67	26	6	70	24	12	61	3	24
DM: Childcare	0	41	59	0	41	59	38	21	0	41
DM: Paying Bills	38	35	26	35	29	35	21	59	0	21
DM: Kids to Doctor	9	59	32	3	71	26	15	53	3	29
DM: Kids Clothing	3	74	24	0	67	33	18	61	0	21
DM: Car	18	0	82	18	6	76	68	6	3	24
DM: Major Purchases	9	9	82	12	9	79	68	6	0	26
DM: Finance	18	0	82	23	6	71	65	12	0	24

	Husbands View (%)			Wives View (%)			Couple-Level Agreements (%)			
Income Pooling	Pooled Income	Separate	Assigned Manage-ment	Pooled Income	Separate	Assigned Manage-ment	Agree Pooled	Agree Assign	Disagree Pool vs. Assign	Disagree Pool vs. Separate
Pooled Income?	73	0	27	77	6	17	61	9	27	3

Table 4. Time Allocation Statistics.

	Husbands View (%)			Wives View (%)			Couple-Level Agreements (%)			
Decisionmaking	Husband Spends More Time	Wife Spends More Time	Joint TA	Husband Spends More Time	Wife Spends More Time	Joint TA	Agree Joint	Agree Non-Joint	Disagree	Disjoint
TA: Infant Care	3	84	13	4	82	14	7	79	0	14
TA: Sick Kids	9	83	9	6	83	11	6	83	3	9
TA: Kids to Doctor	3	74	23	3	83	14	0	63	0	37
TA: After School Care	6	68	26	0	75	25	15	63	4	19
TA: Kids Homework	0	52	48	0	78	22	24	57	0	19
TA: Teacher Meetings	4	44	52	7	54	39	28	32	0	40
TA: Cleaning House	6	54	40	6	69	26	20	54	0	26
TA: Seasonal Cleaning	17	37	46	12	50	38	26	32	9	32
TA: Decoration	9	54	37	6	68	26	15	44	6	35
TA: Home Repairs	71	3	26	69	6	26	14	63	0	23
TA: Home Renovation	59	6	34	60	6	34	25	56	0	19
TA: Paying Bills	40	40	20	34	40	26	17	71	0	11
TA: Finance	37	29	34	37	31	31	17	48	3	31
TA: Major Purchases	23	17	60	21	29	50	38	23	6	32

Table 5. Descriptive Statistics for the Dyadic Adjustment Scale (32 Questions).

	Husbands				Wives				Difference Within Each Couple (Husband-Wife)			
	Mean	Std Dev	Min	Max	Mean	Std Dev	Min	Max	Mean	Std Dev	Min	Max
Handling Finances	3.56	0.82	1.00	5.00	3.62	0.65	2.00	5.00	-0.06	0.69	-2.00	1.00
Recreation	3.79	0.73	2.00	5.00	3.94	0.34	3.00	5.00	-0.15	0.70	-2.00	2.00
Religious Matters	3.82	0.80	2.00	5.00	3.85	0.86	1.00	5.00	-0.03	0.80	-1.00	2.00
Showing Affection	3.85	0.99	1.00	5.00	3.91	0.67	2.00	5.00	-0.06	0.89	-2.00	2.00
Friends	4.03	0.87	1.00	5.00	3.94	0.55	3.00	5.00	0.09	0.93	-2.00	2.00
Sex Relations	3.76	0.94	1.00	5.00	3.68	0.81	1.00	5.00	0.09	0.98	-2.00	3.00
Proper Behavior	3.76	0.96	2.00	5.00	3.68	0.64	2.00	5.00	0.09	0.93	-1.00	2.00
Philosophy of Life	3.81	0.78	2.00	5.00	4.00	0.65	2.00	5.00	-0.16	0.85	-2.00	1.00
Dealing With In-laws	3.68	0.84	1.00	5.00	3.85	0.66	3.00	5.00	-0.18	0.87	-2.00	1.00
Aims and Goals	4.12	0.77	2.00	5.00	4.03	0.68	2.00	5.00	0.09	0.77	-1.00	2.00
Time Spent Together	3.74	0.90	1.00	5.00	3.94	0.60	3.00	5.00	-0.21	0.91	-2.00	2.00
Major Decisions	4.03	0.80	2.00	5.00	3.79	0.64	2.00	5.00	0.24	1.02	-2.00	2.00
Household Tasks	3.21	0.98	0.00	5.00	3.38	0.65	2.00	5.00	-0.18	1.03	-3.00	2.00
Leisure Time	3.56	0.99	0.00	5.00	3.82	0.52	3.00	5.00	-0.26	1.16	-4.00	2.00
Career Decisions	3.94	1.01	1.00	5.00	3.88	0.54	3.00	5.00	0.06	1.10	-3.00	2.00
Discuss Divorce	4.44	0.70	2.00	5.00	4.36	0.99	0.00	5.00	0.09	1.07	-3.00	4.00
Leave After a Fight	4.56	0.66	3.00	5.00	4.61	0.56	3.00	5.00	-0.06	0.70	-1.00	1.00
Things Going Well	3.32	1.32	0.00	5.00	3.45	1.25	0.00	5.00	-0.12	0.65	-1.00	1.00
Confide in Mate	3.65	1.57	0.00	5.00	3.61	1.69	0.00	5.00	0.00	1.25	-2.00	5.00
Regret Marriage	4.62	0.70	2.00	5.00	4.58	0.71	2.00	5.00	0.03	0.73	-2.00	1.00
Quarrel	3.45	0.75	1.00	5.00	3.39	0.75	1.00	5.00	0.09	0.59	-1.00	2.00
Get on Nerves	3.32	0.84	1.00	5.00	3.12	0.78	0.00	4.00	0.21	0.70	-2.00	1.00
Kiss your mate	3.56	0.86	1.00	4.00	3.61	0.75	1.00	4.00	-0.03	0.77	-3.00	2.00
Engage in Interests	2.56	0.75	1.00	4.00	2.58	0.75	0.00	4.00	0.00	0.87	-2.00	3.00
Exchange of Ideas	3.24	1.23	0.00	5.00	3.55	1.03	1.00	5.00	-0.31	1.53	-5.00	3.00
Laugh Together	4.27	0.72	3.00	5.00	4.21	0.82	2.00	5.00	0.09	0.89	-1.00	2.00
Calmly Discuss	3.97	0.98	2.00	5.00	4.09	0.80	2.00	5.00	-0.06	1.11	-2.00	2.00
Project Together	2.70	1.33	0.00	5.00	2.73	1.40	1.00	5.00	-0.06	1.66	-3.00	4.00
Too Tired For Sex	0.61	0.50	0.00	1.00	0.61	0.50	0.00	1.00	0.00	0.62	-1.00	1.00
Not Showing Love	0.76	0.44	0.00	1.00	0.79	0.42	0.00	1.00	0.00	0.44	-1.00	1.00
Overall Happiness	3.97	1.45	0.00	6.00	4.25	1.22	2.00	6.00	-0.23	1.48	-5.00	3.00
Future of Relationship	4.55	0.51	4.00	5.00	4.16	0.68	2.00	5.00	0.35	0.95	-1.00	3.00
Sum of Responses (32-Item Scale)	113.06	15.95	61.00	138.00	112.94	15.25	59.00	143.00	0.12	15.14	-23.00	47.00
Sum of Responses (13-Item Scale)	48.78	8.25	25.00	63.00	49.73	4.36	37.00	61.00	-0.97	7.49	-14.00	18.00

Table 6. Ranking of Hazard Controllability by Spouse - % of Respondents for each Possible Ranking (1=Most Controllable).

	Air Pollution		Climate Change		Radon		Lead Paint		Smallpox		Smallpox Vaccine		Anthrax		Influenza	
Rank	H	W	H	W	H	W	H	W	H	W	H	W	H	W	H	W
1	9	3	3	3	35	20	69	54	12	17	54	71	9	3	14	14
2	9	0	3	0	24	23	20	34	3	11	3	6	6	3	9	6
3	17	6	11	6	21	31	0	11	15	11	20	3	17	9	17	34
4	14	34	11	17	9	14	3	0	21	29	11	12	20	23	34	23
5	23	31	20	14	6	9	6	0	29	14	9	3	20	20	11	17
6	9	11	6	14	3	3	3	0	9	9	3	0	14	26	11	3
7	17	14	31	26	3	0	0	0	12	9	0	3	14	11	3	0
8	3	0	14	20	0	0	0	0	0	0	0	3	0	6	0	3

Table 7. Descriptive Statistics for Questions About Lead Knowledge.

	Husbands		Wives		Difference Within Each Couple (Husband-Wife)			
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Min	Max
Lead can be found throughout the environment. (TRUE)	0.86	0.36	0.97	0.17	-0.11	0.40	-1.00	1.00
Lead-based paint is rarely found in pre-1978 housing. (FALSE)	0.71	0.46	0.77	0.43	-0.06	0.59	-1.00	1.00
Young children are less vulnerable to lead poisoning. (FALSE)	0.80	0.41	0.91	0.28	-0.11	0.40	-1.00	1.00
Young children are more likely to come into contact with lead if it is in their environment. (TRUE)	0.69	0.47	0.83	0.38	-0.14	0.55	-1.00	1.00
Iron deficiency may increase vulnerability to lead poisoning. (TRUE)	0.29	0.46	0.54	0.51	-0.26	0.61	-1.00	1.00
Children absorb and retain relatively less lead than adults. (FALSE)	0.74	0.44	0.66	0.48	0.09	0.56	-1.00	1.00
Lead poisoning can decrease a person's IQ. (TRUE)	0.80	0.41	0.80	0.41	0.00	0.49	-1.00	1.00
Lead poisoning can cause respiratory problems. (TRUE)	0.71	0.46	0.74	0.44	-0.03	0.62	-1.00	1.00
Lead can be found in the blood, brain, and bones. (TRUE)	0.89	0.32	0.91	0.28	-0.03	0.38	-1.00	1.00
Lead poisoning can lead to lower school performance. (TRUE)	0.94	0.24	0.91	0.28	0.03	0.38	-1.00	1.00
Lead does not contribute to hyperactivity in children. (FALSE)	0.23	0.43	0.40	0.50	-0.17	0.57	-1.00	1.00
Lead dust can be found in windowsills in houses that are contaminated. (TRUE)	0.91	0.28	0.91	0.28	0.00	0.24	-1.00	1.00
Cleaning can help minimize lead dust. (TRUE)	0.80	0.41	0.60	0.50	0.20	0.63	-1.00	1.00
Number of Correct Answers	9.37	2.38	9.97	2.15	-0.60	2.75	-7.00	4.00

*1=Question answered correctly; 0=Question answered incorrectly.

Table 8. Variables Used in Regression Analyses.

Explanatory Variables	Definition
Non-White	0 = White; 1 = Non-White.
Male	0 = Female; 1 = Male.
Age	Age in Years (spouse variable as well).
Educational Difference	Husband's Age minus Wife's Age (with Absolute Value option).
Education	Years of Education (spouse variable as well).
Total Household Income	Income in ($000). Constructed from the midpoint of respondent-selected income intervals. Also used a dummy variable (1= Above median income)
Share of Household Income	Respondent's share (%) of household income. Constructed from the midpoint of respondent-selected income intervals.
Children	Number of Children.
Wife Employed	0 = Wife not employed; 1 = Wife Employed in household of the respondent.
Wife Fulltime Job	0 = Wife does not work fulltime; 1 = Wife works fulltime in household of the respondent
Interactions with Income & Income Shares	Wife Employed and Wife Fulltime Job Dummy interacted with Total Household Income and Wife's Share of Income.
Index of Work Commitment	Index of commitment to the work force generated from responses regarding desires for staying at home, careers, and who should be the breadwinner. Values range from 0 to 4 with higher values implying greater desired commitment to work.
Previous Marriage	0 = No previous marriage; 1 = Married Previously.
Age of Oldest Child	In years
Dyadic Scale of Marital Adjustment	Total Score on the 32-item "Dyadic Adjustment Scale" to assess marital adjustment. Possible range from 0 to 151, with higher numbers indicating better adjustment
DAS Subscale of Marital Consensus	13-item subscale of above
Performance on Lead Knowledge Questions	Number of correct answers about lead-health knowledge (true-false).
Index for Beliefs About Children's Risk Levels for Hazards	Index identifying beliefs about current risk levels over 8 hazards. Values range from 0 to 8 with higher values implying higher perceived risk.
Index for Beliefs on Controllability of Children's Risks	Index identifying beliefs about controllability of children's risks from 8 hazards. Higher values imply greater controllability.
Time Allocation Series	Variables for each time allocation domain (see Table 4). 0 = Joint Time Allocation; 1 = Non-Joint Time Allocation (i.e., either wife spends more time or husband spends more time in this domain).

Dependent Variables	Definition
DM: Kids to Doctor	Who decides when the children go to the doctor? 0 = Joint; 1 = Wife.
DM: Childcare	Who makes childcare decisions? 0 = Joint; 1 = Wife.
DM: Paying Bills	Who makes decisions about paying bills? 0 = Joint; 1 = Other.
DM: Financial	Who makes major financial decisions? 0 = Joint; 1 = Husband.
Income Pooling	Arrangement for managing household income. 0 = Pooled; 1 = Not Pooled.
DM: Total	Used in domain regressions (Table 10). All 10 DM domain variables are pooled so that there are 10 observations per respondent. 0 = joint; 1 = Other

Table 9. Selected Logit Regression Results for Different Specifications of the Decision Variables at the Individual Level (N=70 or less).

Independent Variables	Dependent Variables				
	DM: Doctor	DM: Childcare	DM: Paying Bills	DM: Finance	Income Pooling
	Wife vs. Joint	Wife vs. Joint	Other vs. Joint	Husband vs. Joint	Other vs. Pooled
Non-White		Joint*			
Educational Difference (Abs Value)	Joint	Joint*			
Years of Education			Other§ (dummy)	Husband (dummy)	Pooled*
Total Household Income	Wife§ (Dummy for above median income)	Wife*	Joint	Husband§	
Number of Children	Joint**		Joint§		
Wife's Share of Household Income		Wife**		Joint§ (dummy)	
Wife Employed		Joint**			Other§
(Wife Employed) x (Wife Inc Share)					Pooled
Index of Work Commitment	Joint§				Other
Dyadic Scale of Marital Adjustment (13 Questions)	Joint*				
# of Lead Questions Answered Correctly	Joint§		Joint**		
Index for beliefs about children's risk levels for hazards			NA	NA	NA
Index for Beliefs on Controllability of Children's Risks		Wife	Other* (dummy)		
Time Allocation for Related Tasks	Wife§ (TA: Doctor)		Other** (TA: Bills)	Husband** (TA: Purchases)	Other* (TA: Finance)

Note: To interpret this table, of a sample of couples where decision about childcare are made either jointly or predominately by the wife, non-white respondents are more likely to report that such decisions are made jointly.
Note: § indicates significance at the 10% level; * indicates significance at the 5% level; ** indicates significance at the 1% level.

Table 10. Logit Regressions on Pooled Decision-Making Domains (Joint = 0; Other = 1)

Variable	*Specification A*		*Specification B*		*Specification C*		*Specification D*	
	Coefficient	*Std error*	*Coefficient*	*Std error*	*Coefficient*	*Std error*	*Coefficient*	*Std error*
Intercept	0.36	0.25	-0.34	0.26	-3.18**	0.99	0.36	1.49
Domain Variables:								
DM: General	-1.66**	0.45	-1.66**	0.45	-1.93**	0.48	-1.91**	0.50
DM: Home Repairs	0.42	0.35	0.42	0.35	0.40	0.36	0.30	0.39
DM: Decoration	1.44**	0.37	1.44**	0.37	1.51**	0.39	1.54**	0.42
DM: Paying Bills	1.16**	0.36	1.16**	0.36	1.22**	0.38	1.15**	0.40
DM: Kids to Doctor	1.23**	0.36	1.23**	0.36	1.30**	0.38	1.41**	0.41
DM: Kids Clothing	1.28**	0.37	1.28**	0.37	1.36**	0.39	1.30**	0.41
DM: Car	-0.99*	0.39	-0.99*	0.39	-1.08**	0.40	-1.23**	0.44
DM: Major Purchases	-1.09**	0.39	-1.08**	0.39	-1.18**	0.41	-1.12**	0.43
DM: Finance	-0.82*	0.38	-0.82*	0.38	-0.90*	0.39	-0.92*	0.42
Male			-0.04	0.17	-0.05	0.18		
Age					-0.20	0.02	-0.04*	0.02
Household Income ($000)					0.00	0.00		
HH income dummy ($\geq$ 80,000)							0.54*	0.23
Education (years)					0.21**	0.05	0.19**	0.05
Number of Children					-0.13	0.13	-0.27*	0.14
Wife Employed							-0.52§	0.28
Wife Employed x wife income							-0.00	0.00
Dyadic scale of marital consensus (13 Questions)							-0.04*	0.02
-2*Log Likelihood	-768		-768		-709		-634	
Number of Observations	678		678		668		608	

Note: § indicates significance at the 10% level; * indicates significance at the 5% level; ** indicates significance at the 1% level.

Figure 1.

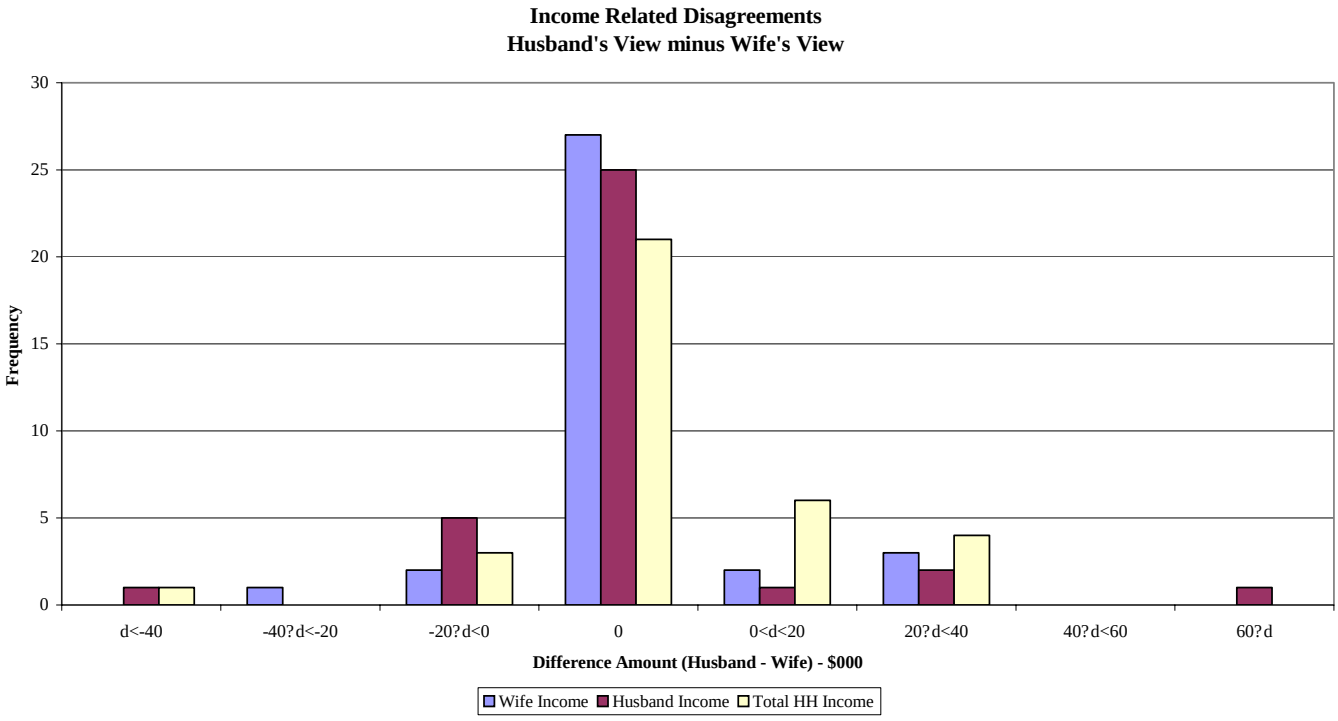

Income Related Disagreements
Husband's View minus Wife's View
Frequency
30
25
20
15
10
5
0
d<-40
-40?d<-20
-20?d<0
0
0<d<20
20?d<40
40?d<60
60?d
Difference Amount (Husband - Wife) - $000
Wife Income
Husband Income
Total HH Income

Figure 2.

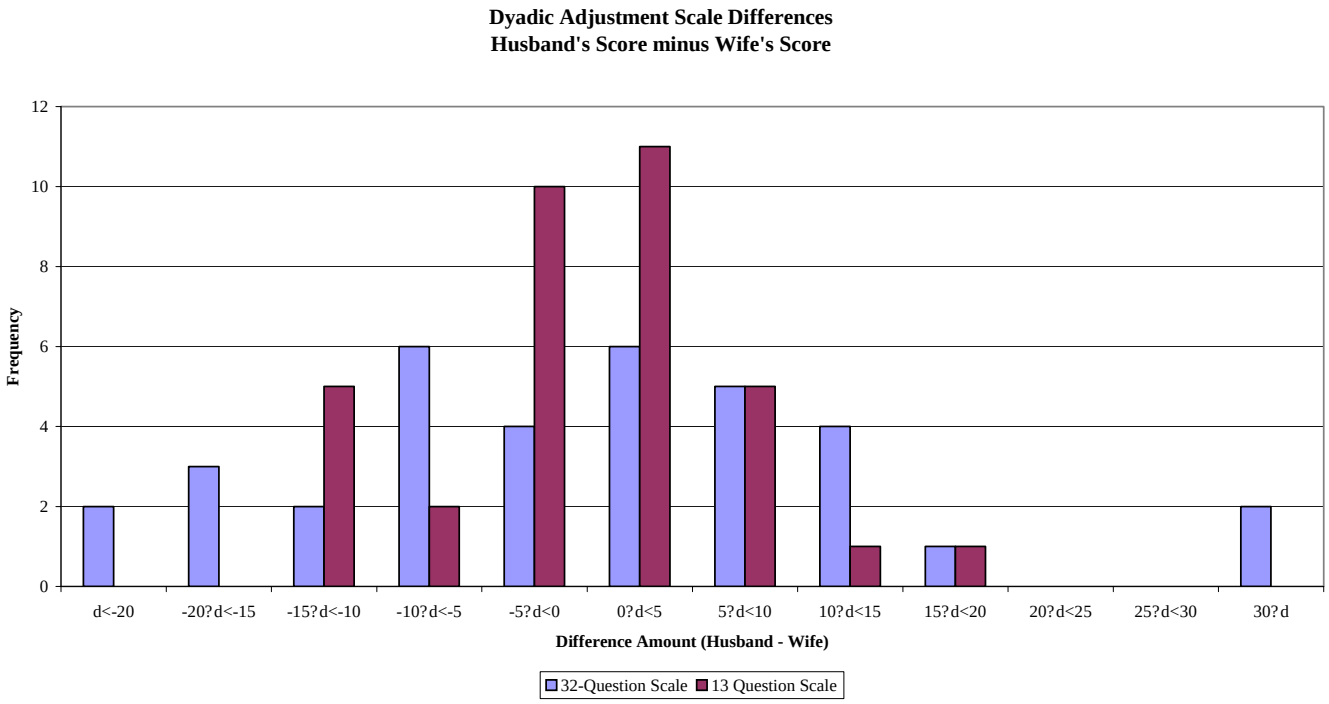
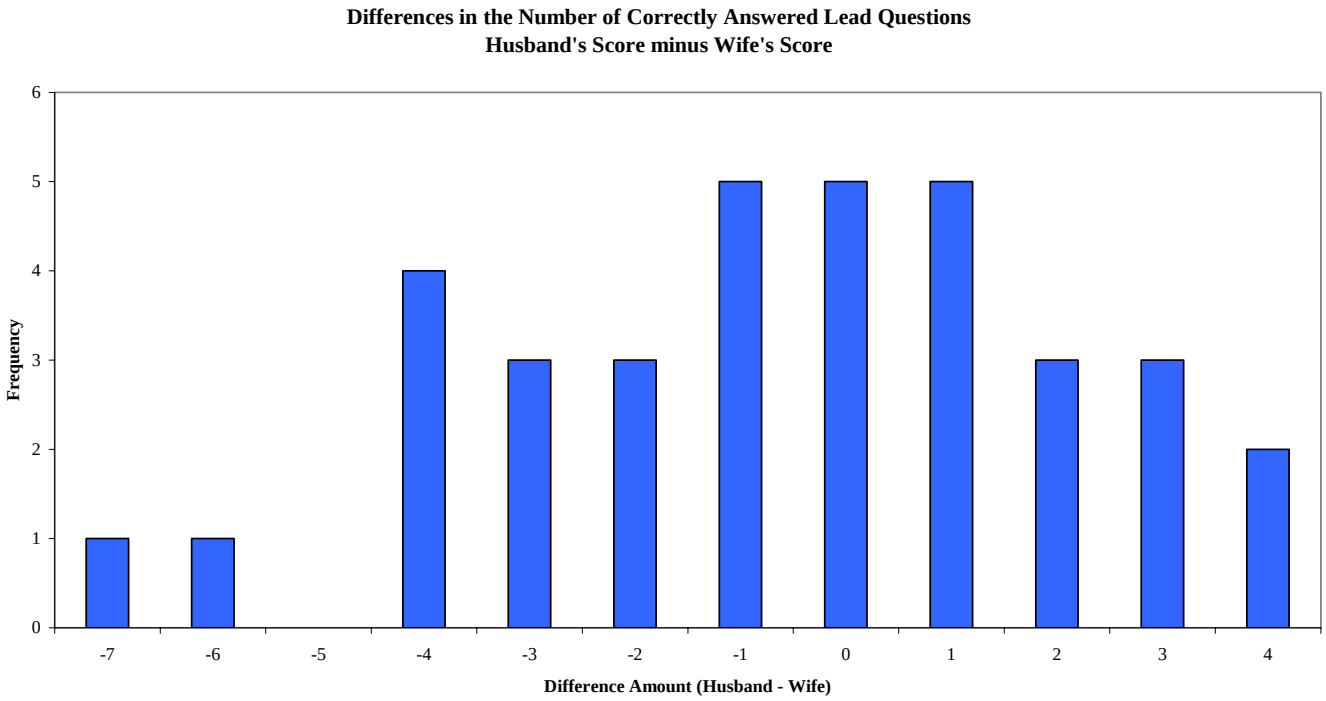

Dyadic Adjustment Scale Differences
Husband's Score minus Wife's Score
Frequency
12
10
8
6
4
2
0
d<-20
-20?d<-15
-15?d<-10
-10?d<-5
-5?d<0
0?d<5
5?d<10
10?d<15
15?d<20
20?d<25
25?d<30
30?d
Difference Amount (Husband - Wife)
32-Question Scale
13 Question Scale

Figure 3.


Differences in the Number of Correctly Answered Lead Questions
Husband's Score minus Wife's Score
Frequency
0
1
2
3
4
5
6
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
Difference Amount (Husband - Wife)

REFERENCES

Adelman, L., D.A. Zirk, P.E. Lehner, R.I. Moffett, and R. Hall. 1986. "Distributed Tactical Decisionmaking: Conceptual Framework and Empirical Results." *IEEE Transactions on Systems: Man and Cybernetics* 16: 794-805.

Apps, P. and R. Rees. 1996. "Labour Supply, Household Production and Intra-family Welfare Distribution." *Journal of Public Economics* 60: 199-219.

Apps, P. and R. Rees. 1997. "Collective Labor Supply and Household Production." *Journal of Political Economy* 105: 178-190.

Ashworth, J. and Ulph, D. 1981. "Household Models. In C. Bown (ed.), *Taxation and Labour Supply*. London: George Allen and Unwin, 117-133.

Bainbridge, L. 1991. "Mental Models in Cognitive Skill," in A. Rutherford and Y. Rogers (eds.), *Models in the Mind*. Academic, New York.

Bateman, I. and A. Munro. 2005. "An Experiment on Risky Choice amongst Households." *Economic Journal* 115(502): C176-89.

Becker, G. 1974. "A Theory of Social Interactions." *Journal of Political Economy* 82: 1063-93.

Bostrom, A., B. Fischhoff, and M. Granger Morgan. 1992. "Characterizing Mental Models of Hazardous Processes: A Methodology and an Application to Radon." *Journal of Social Issues* 48(4): 85-100.

Bourguignon, F. and P. A. Chiappori. 1992. Collective Models of Household Behavior: An Introduction, *European Economic Review*, 36, 1-10.

Bourguignon, F. 1999. "The Cost of Children: May the Collective Approach to Household Behavior Help?" *Journal of Population Economics* 12: 503-521.

Browning, M., F. Bourguignon, P.-A. Chiappori, and V. Lechene. 1994. "Income and Outcomes: A Structural Model of Intrahousehold Allocation," *Journal of Political Economy*. 102: 1067-96.

Browning, M. and P.-A. Chiappori. 1998. " Efficient Intra-Household Allocations: a General Characterization and Empirical Tests", *Econometrica*. 1241-78.

Browning, M. 2000. "The Saving Behavior of a Two-person Household," *Scandinavian Journal of Economics*, 102: 235-251.

Busby, D.M., C. Christensen, D.R. Crane, and J.H. Larson. 1995. "A Revision of the Dyadic Adjustment Scale For Use With Distressed and Nondistressed Couples:

Construct Hierarchy and Multidimensional Scales." *Journal of Marital and Family Therapy* 21(3): 289-308.

Casas, J. M., and S. Ortiz. 1985. "Exploring the Applicability of the Dyadic Adjustment Scale for Assessing Level of Marital Adjustment with Mexican-Americans." *Journal of Marriage and the Family* 47: 1023-1027.

Chiappori, P.-A. 1997. "Introducing Household Production in Collective Models of Labor Supply." *Journal of Political Economy* 105: 191-209.

Craik, K. 1943. *The Nature of Explanation*. Cambridge University Press: Cambridge, UK.

Crane, D.R., S.M. Allgood, J.H. Larson, and W. Griffin. 1990. "Assessing Marital Quality with Distressed and Nondistressed Couples: A Comparison and Equivalency Table for Three Frequently Used Measures." *Journal of Marriage and the Family* 52: 87-93.

Daniels, K., L. de Chernatony, and G. Johnson. 1995. "Validating a Method for Mapping Managers' Mental Models of Competitive Industry Structures." *Human Relations* 48: 975-991.

Dosman, D. and W. Adamowicz. Forthcoming. "Combining Stated and Revealed Preference Data to Construct an Empirical Examination of Intrahousehold Bargaining." *Review of Economics of the Household*.

Doss, C. 1996. "Testing among Models of Intrahousehold Resource Allocation," *World Development.* 24, 1597-1609.

Eddy J.M., R.E. Heyman, and R.L. Weiss. 1991. "An Empirical Evaluation of the Dyadic Adjustment Scale: Exploring the Differences Between Marital 'Satisfaction' and 'Adjustment'." *Behavioral Assessment* 13: 199–220.

Gentner, D. and A. Stevens (eds.). 1983. *Mental Models*. Lawrence Erlbaum Associates: Hillsdale, New Jersey.

Hoddinott, J. and L. Haddad. 1995. "Does Female Income Share Influence Household Expenditure Patterns?," *Oxford Bulletin of Economics and Statistics*, 57(1): 77-96.

Jacobson, N.S. 1984. "A Component Analysis of Behavioral Marital Therapy: The Relative Effectiveness of Behavior Exchange and Communication/problem-solving Training." *Journal of Consulting and Clinical Psychology* 52: 295–305.

Langan-Fox, J. and S. Code. 2000. "Team Mental Models: Techniques, Methods and Approaches." *Human Factors* 42(2): 242-271.

Lesvesque, L.L., J.M. Wilson, and D.R. Wholey. 2001. "Cognitive Divergence and Shared Mental Models in Software Development Project Teams. *Journal of Organizational Behavior* 22: 135-144.

Leuthold, J. 1968. "An Empirical Study of Formula Income Transfers and the Work Decision of the Poor," *Journal of Human Resources*, 3: 312-323.

Lundberg, S.J. and R.A. Pollak. 1993. "Separate Spheres Bargaining and the Marriage Market," *Journal of Political Economy*. 101(6): 988-1010.

Lundberg, S.J., R.A. Pollak, and T. Wales. 1997. "Do Husbands and Wives Pool Resources?: Evidence from the UK Child Benefit," *Journal of Human Resources.* 32(3): 463-480.

Lynch, C. 1995. "The Use of Repertory Grid Technique to Obtain the Views of Residents in a Continuing Care Setting." *International Journal of Therapeutic and Supportive Organizations* 18: 111-122.

Manser, M. and M. Brown. 1980. "Marriage and Household Decision-Making: A Bargaining Approach," *International Economic Review* 21(1): 31-44.

McElroy, M. 1990. "The Empirical Content of Nash-Bargained Household Behavior," *Journal of Human Resources.* 25(4): 559-83.

McElroy, M. and M.J. Horney. 1981. "Nash-bargained Household Decisions: Toward a Generalization of the Theory of Demand," *International Economic Review* 22(2): 333-349.

Mohammed, S. and B.C. Dumville. 2001. "Team Mental Models in a Team Knowledge Framework: Expanding Theory and Measurement across Disciplinary Boundaries," *Journal of Organizational Behavior* 22: 89-106.

Morgan, M.G., B. Fischhoff, A. Bostrom, and C.J. Atman. 2001. Risk Communication: A Mental Models Approach. Cambridge University Press: Cambridge, UK.

Orasanu, J. and E. Salas. 1993. "Team Decisionmaking in Complex Environments," in G.A. Klein, J. Orasanu, R. Calderwood, and C Zsambok (eds.), *Decisionmaking in Action: Mental Models and Methods*. Ablex: Norwood, New Jersey.

OECD (Organisation for Economic Co-operation and Development). 2006. *Economic Valuation of Environmental Health Risks to Children*. OECD Publishing: Paris, France.

Phipps, S. and P. Burton. 1998. "What's Mine is Yours? The influence of Male and Female Incomes on Patterns of Household Expenditure," *Economica* 65: 599-613.

Rouse, W.B., J.A. Cannon-Bowers, and E. Salas. 1992. "The Role of Mental Models in Team Performance in Complex Systems." *IEEE Transactions on Systems, Man, and Cybernetics* 22: 1296-1308.

Samuelson, P. 1956. "Social Indifference Curves." *Quarterly Journal of Economics* 70: 1-22.

Shek, D.T.L. 1994. "Psychometric Properties of the Chinese Version of the Dyadic Adjustment Scale." *Psychologia* 37: 7-17.

Sorkin, R.D., C.J. Hays, and R. West. 2001. "Signal Detection Analysis of Group Decisionmaking." *Psychological Review* 108: 183–203.

Spanier, G.B. 1976. "Measuring Dyadic Adjustment: New Scales for Assessing the Quality of Marriage and Similar Dyads." *Journal of Marriage and the Family* 38(1): 15-28.

Spanier, G.B., and L. Thompson. 1982. "A Confirmatory Analysis of the Dyadic Adjustment Scale." *Journal of Marriage and the Family* 44(3): 731-738.

Strauss, J., Mwabu, G. and Beegle, K. 2000. "Intrahousehold allocations: A Review of Theories and Empirical Evidence," *Journal of African Economies*, 9, AERC Supplement 1: 83-143.

U.S. EPA. 2000. National Center for Environmental Research Proceedings: EPA/NIEHS/CDC Centers for Children's Environmental Health and Disease Prevention Research Progress Review Workshop. November 5-7: University of California, Berkeley.

U.S. EPA. 2003. National Center for Environmental Economics National Workshop: Valuing Environmental Health Risk Reductions to Children. October 20-21: Washington, DC.

Vermeulen, F. 2004. "Collective Household Models: Principles and Main Results." *Journal of Economic Surveys* 16(4): 533-564.

Value of Reducing Children's Mortality Risk: Effects of Latency and Disease Type

James K. Hammitt
Center for Risk Analysis, Harvard University
IDEI and LERNA-INRA, Université de Toulouse
jkh@harvard.edu

Kevin Haninger
Center for Risk Analysis, Harvard University
haninger@fas.harvard.edu

March 2006

Abstract

Despite research showing children may be differentially susceptible to various environmental health hazards, and that risks to children may be of greater social concern than risks to adults, there have been relatively few studies that estimate the economic value of reducing risk to children's health. We propose to design and conduct a contingent valuation (CV) survey to estimate household willingness to pay (WTP) to reduce mortality risk from pesticides in food, and to compare WTP to reduce risks to children and risks to adults. We will examine how WTP depends on latency (the length of the period between exposure and development of symptoms), noting that childhood exposure may lead to childhood or adult disease and fatality, depending on latency. We will also evaluate how WTP depends on disease type, comparing terminal cancer and non-cancer illnesses that present similar symptoms and prognosis.

We will elicit values for risk reductions that vary across the following characteristics: whether the pesticide exposure is to a child or to an adult, whether the disease is latent or acute, whether the disease is cancer or not cancer. We will vary the level of detail provided about the disease to determine whether differences in WTP to reduce risks of cancer and non-cancer disease reflect differences in information. We will also vary the magnitude of risk reduction, and use sensitivity of WTP as a diagnostic criterion for validity of the results. Survey respondents will include both parents and non-parents to allow comparison with prior studies of the value of reducing risks to adults, and we will measure a variety of demographic variables that may influence WTP. By comparing estimated WTP between and within respondents, it will be possible to estimate the relative value of reducing health risks to children versus adults. The survey will be administered over the World Wide Web, which will facilitate the presentation of visual aids to assist in communicating the magnitude of risks to survey respondents.

This project is anticipated to provide estimates of the value of reducing food-borne pesticide risk to children versus adults, as well as analysis of how age, latency, and disease type influence the valuation. Policymakers can use such estimates to evaluate the benefits of programs aimed at reducing risks to children.

Keywords: Willingness to pay, health risk, stated-preference, children, cost-benefit analysis

1. Introduction

Despite evidence that children and adults differ significantly in their exposure and vulnerability to toxic substances, and observations that individuals may systematically place a different value on child health than they do on adult health (US EPA, 2001), most of the existing valuation estimates pertain to risks to adults. Moreover, most previous studies have focused on risks of traumatic fatality, such as workplace or transportation accidents, which differ qualitatively from the risks of cancer and other disease that are more often associated with environmental contaminants (Savage, 1993; Revesz, 1999; Sunstein, 1997).

This study is intended to complement previous studies by estimating household willingness to pay (WTP) to reduce environmental health risks to children, and by examining how the value of reducing risks to children compares with the value of reducing similar risks to adults. In addition, the study will investigate the effects of two risk characteristics that are particularly important in valuing environmental health risks to children: latency (the period between exposure to an environmental contaminant and development of adverse health effects) and disease type (cancer versus other degenerative, fatal diseases).

Many environmental risks are characterized by a latency period between exposure to the environmental contaminant and adverse health effects. The duration of the latency period can determine whether childhood exposure to a contaminant manifests as disease or death of the child, or of the adult. In contrast, adult exposure necessarily manifests as disease or death of the adult. We will investigate the effects on WTP of latency and of whether the exposure and/or disease manifestation occur to children or adults.

We propose to use contingent valuation (CV) to estimate the effects of age, latency, and disease type on WTP to reduce mortality risk. In particular, we will elicit parents' WTP to reduce fatal risks to their children associated with exposure to pesticides in food, and we will compare these values with parents' and other adults' WTP to reduce similar risks to themselves. In both cases, the risks presented will vary in latency, whether they cause cancer or another disease, and other attributes.

In the following section, we describe the theoretical and empirical background for the study. In Section 3, we describe the survey instrument and sample. In Section 4, we report the results of regression models relating WTP to the severity and duration of illness, reduction in its probability, other risk attributes, and to demographic and preference characteristics of the respondents.

2. Background

In this section, we describe the theoretical and empirical background for this study. First, we briefly review the literature on the value of reducing health risks to children. Second, we describe the reasons for selecting health risks of pesticide residues on food as the hazard whose reduction we will value. Third, we describe the economic theory and prior empirical results concerning the effects of latency and disease type on risk to adults and the implications for children's risk. Fourth, we justify the use of household WTP for valuing children's health.

2.1. Prior Work on Valuing Children's Health

There are two strands of prior work that relate to valuing children's health: estimates of altruistic WTP to protect another individual's health, and estimates of household spending on children's health (Becker, 1981; Johansson, 1994). Viscusi et al. (1987) used CV to estimate WTP to prevent the risk of injury associated with household pesticides. They found that WTP to reduce risks to one's children exceeds WTP to reduce risks to oneself, but could not distinguish between the effects of parental altruism and injury severity. Viscusi et al. (1988) examined household WTP to reduce risks of injury associated with household insecticides, for injuries to adults and children within and outside the household. They found that household values for a statistical case of child inhalation poisoning were about 75 percent larger than for a statistical case of adult skin poisoning. Unfortunately, this research does not allow estimation of the relative value of adult and childhood risks of the same injury.

In the same study, Viscusi et al. (1988) elicited WTP to reduce these risks to people in other households, both in the same state (North Carolina) and in the United States as a whole. Viscusi et al. found that altruistic WTP to reduce risks to other households was substantial and was greater for reducing risks to children than for reducing risks to adults. In particular, the probability of contributing to a program to reduce risks in the state was 79 percent for a program that reduced risks to children, and 57 percent for a program that reduced risks to adults. Average contributions to each program, accounting for the probability of contributing, were $11.53 for reducing risks to children and $8.75 for reducing risks to adults.

Agee and Crocker (1996) estimated parental WTP to reduce the risk of neurological impairments from childhood exposure to lead using a revealed-preference approach based on the parents' decision to obtain chelation therapy for their child. They did not examine WTP to reduce risks of neurotoxicity to adults, which are much smaller than the risk to children.

A more recent study by Liu et al. (2000) used CV to estimate mothers' WTP to protect themselves and their children from suffering a cold. WTP was positively associated with the severity of symptoms and the duration of illness. In addition, mothers' WTP to protect their child from a cold was nearly twice as large as their private WTP to protect themselves from a cold of equivalent severity and duration, an indication that mothers value their children's health more than their own.

2.2. Pesticide Risks

We propose to study WTP to reduce health risks from residual pesticides on food for a variety of reasons. First, pesticide contamination of food is a topic of major public concern. Opinion polls show that pesticides consistently rank as one of the greatest concerns about food safety in the US (Buzby et al., 1995; Bruhn et al., 1992; Ott et al., 1991). In part as a result of this concern, the market for "organic" or foods grown without use of synthetic pesticides has grown to approximately 2% of the US food market (US Department of Agriculture, 1997).

Second, to compare WTP to reduce risks to children and adults, we require a hazard that allows us to distinguish actions that reduce risks to different members of the household. Exposures to many environmental health risks are similar to all household members (e.g., air, drinking and bathing water). Even though some household members are more highly exposed to certain environmental media (e.g., children may be more exposed to dust and soil than adults), it is

difficult to construct plausible scenarios for a CV study that reduce risks to children, or to adults, but not to both. In this respect, foodborne risks are attractive because it is often the case that children and adults in a household will consume different foods (at least in part), and so it is plausible to imagine reducing pesticide concentrations on a food that only the children eat, or a food that only the adults eat.

2.3. Theoretical Background

The economic approach to valuing mortality risk was developed by Schelling (1968) in an article suggestively entitled "The Life You Save May Be Your Own." Several years earlier, Drèze (1962) proposed a similar approach in a French operations research journal, but his work has received little attention among English-speaking economists. Schelling observed that for environmental regulations and other life-saving programs, one cannot know whose life will be "saved." The question is not how to value prevention of a specific death, but how to value small changes in mortality risk across a population.

The value per statistical life (VSL) is defined as an individual's marginal rate of substitution between mortality risk and wealth. VSL is not a universal constant but varies by individual and circumstance. The standard economic model of preferences for wealth and mortality risk (Jones-Lee, 1974; Weinstein et al., 1980; Drèze, 1962) assumes that an individual's welfare can be represented as:

$$EU(p,w) = (1-p)u_a(w) + pu_d(w) \tag{1}$$

where p is the individual's chance of dying during the current period and $u_a(w)$ and $u_d(w)$ represent his utility as a function of wealth conditional on surviving and not surviving the period, respectively. The function $u_a(w)$ incorporates the individual's preferences for bequests and can incorporate any financial consequences of dying (such as medical bills or life-insurance benefits). In this one-period model, wealth and income are treated as equivalent, but the difference between them can be important in multiple-period models.

The individual's VSL is derived by differentiating Equation (1) holding expected utility constant to obtain

$$VSL = \frac{dw}{dp} = \frac{u_a(w) - u_d(w)}{(1-p)u_a'(w) + pu_d'(w)} = \frac{\Delta u(w)}{Eu'(w)} \tag{2}$$

where prime indicates first derivative.

The numerator in Equation (2) is the difference in utility between surviving and dying in the current period. The denominator is the expected marginal utility of wealth, i.e., the utility associated with additional wealth conditional on surviving and dying, weighted by the probabilities of these events. Assuming that life is preferred to death and that greater wealth is preferred to less, both numerator and denominator are positive and so VSL is positive. If the marginal utility of wealth is non-negative, and greater in the event of survival than death (i.e., $u_a'(w) > u_d'(w) \geq 0$), then VSL increases in mortality risk p. Weak risk aversion with respect to wealth, conditional on survival and on death (i.e., $u_a''(w) \leq 0, u_d''(w) \leq 0$), is a sufficient condition for VSL to increase with wealth.

In the following subsections, we describe what the theory tells us about how VSL depends on age, latency, and disease type.

Effect of Age on VSL. Theoretical and empirical studies of VSL have generally focused on own WTP for own risk, treating the individual as the economic agent. Some of these studies have evaluated the effect of age, but only within adults, and have not considered the valuation of risks to children. Nevertheless, it may be informative to consider extrapolating results from young adulthood to childhood.

Theoretical models (e.g., Shepard and Zeckhauser, 1984; Rosen, 1988; Ng, 1992) represent the individual's lifetime utility as the expected present value of his utility in each time period. Utility within a period depends on consumption, which is limited by current income, savings and inheritance, and ability to borrow against future earnings. The individual seeks to maximize lifetime utility by allocating his wealth to consumption, savings, and reductions in current-period mortality risk.

Two factors influence the life-cycle pattern of VSL. First, the number of life years at risk declines as one ages, so the benefit of a unit decrease in current-period mortality risk declines. Second, the opportunity cost of spending on risk reduction also declines with age as savings accumulate and the investment horizon approaches. The net effect may cause VSL to fall or rise with age.

In models that assume an individual can borrow against future earnings, VSL declines monotonically with age. For example, Shepard and Zeckhauser (1984) calculate that VSL for a typical American worker falls by a factor of three from age 25 to age 75. If individuals can save but not borrow, VSL rises in early years as the individual's savings (and earnings) increase before it ultimately declines. In this case, Shepard and Zeckhauser find that VSL peaks near age 40 and is less than half as large at ages 20 and 65.

Ng (1992) argues that the rate at which individuals discount their future utility is likely to be smaller than the rate of return to financial assets, whereas Shepard and Zeckhauser (1984) assume these rates are the same. If the utility-discount rate is less than the rate of return, individuals should save more when they are young and consume more when old. Under these conditions, VSL may not peak until age 60 or so (Ng, 1992). Even if individuals discount future utility at the rate of return, if they are prudent (Kimball, 1990), younger people might be anticipated to save more, and spend less on reducing mortality risk, because of the greater range of future financial contingencies they face.

Although many CV studies include age as one of several covariates in a regression model explaining WTP for risk reduction, these studies have not typically focused on estimating the effect of age on VSL. The results of these studies are somewhat contradictory, with several finding VSL increases with age (Gerking et al., 1988; Johannesson et al., 1997; Lee et al., 1997) and others finding VSL decreases with age (Buzby et al., 1995; Hammitt and Graham, 1999). Jones-Lee et al. (1985) included both linear and quadratic age terms in their regression models and concluded that VSL peaks at about the mean age in their sample (which is not reported).

Several studies have attempted to empirically estimate the effect of age on the benefits of public life-saving programs, by asking respondents to choose between hypothetical lifesaving programs that protect people of different ages at different dates. These results do not necessarily reflect individual WTP to reduce different risks to oneself, since it is implausible to assume that survey

respondents compare programs solely in terms of their own private benefits. Cropper et al. (1994) asked survey respondents about programs to save people of different ages. Their results suggest that respondents most prefer to protect people in young middle age. Lives of 30 year olds were valued about 11 times more highly than lives of 60 year olds. For comparison, lives of 20 and 40 year olds are valued as equal to about 8 and 7 60 year olds, respectively. Risks to children were not evaluated explicitly, but extrapolating the relations found for other ages suggests that risks to children would be valued as less than risks to young adults. Interestingly, these results were not sensitive to the age of the respondent.

Two recent empirical studies are specifically directed toward estimating the effect of age on VSL. Krupnick et al. (2002) conducted a CV study of WTP for a hypothetical intervention that would reduce the respondent's risk of dying in the next 10 years by either 1 in 1,000 or 5 in 1,000. The sample was restricted to individuals aged 40 years and above. Krupnick et al. estimate that VSL is roughly constant for ages 40-69, and is about 30 percent smaller for individuals aged 70 and above. Smith et al. (2001) estimate compensating-wage differential estimates using data from the Health and Retirement Survey. Their estimates of VSL for individuals aged 51-65 are not sensitive to age and are comparable to standard estimates for younger populations.

Accounting for Latency. In Equation (2), VSL is defined in terms of wealth and mortality risk in a single period. Many environmental risks are characterized by a latency period between the time an individual is exposed to an agent and the time when he may die from its toxic effect. Since preventive measures must be undertaken before the exposure occurs, there is often a need to determine WTP now to reduce the risk of fatality in a future period.

Standard economic theory suggests that the appropriate procedure to account for latency is to value the risk change using the VSL representing the individual's value when the risk manifests, and to adjust for the time-value of money and the chance that the individual will die before then (Cropper and Sussman, 1990; Cropper and Portney, 1990). The adjustment is made by discounting the future value of the risk reduction back to the time when the expenditure must be incurred (at the individual's rate of interest). For example, assume that pollution-control equipment that could be installed today would reduce an individual's risk of dying from cancer by 1 chance in 100,000, that the cancer would prove fatal 20 years after exposure, that his VSL in 20 years will be $8 million, and that the individual can earn a 5 percent annual return on investments. In 20 years, he would be willing to pay $80 to reduce a contemporaneous fatality risk of 1 in 100,000. The amount he would be willing to pay now is the present value of $80, about $30 (= $80 x 1.05^{-20}). This amount should be multiplied by the probability that the individual will survive the intervening 20 years, since the cancer-risk reduction is of no benefit in the event that he dies of other causes before the environmental pollutant could have killed him. In many cases, this survival factor is much less important than the discount factor. For the average American, the probability of surviving 20 years is greater than 0.7 if the individual is younger than 55 (National Center for Health Statistics, 1998).

The effect of calendar time on VSL has received relatively little attention in the literature, except to observe that if economic welfare grows over time, VSL would be expected to increase. The United States Environmental Protection Agency (EPA) has sometimes accounted for the anticipated growth of income and VSL in regulatory impact assessments, especially when benefits extend across generations. For example, in evaluating the effects of restrictions on use of

CFCs to protect stratospheric ozone, EPA assumed that VSL would grow at annual rates of 0.85-3.4 percent (U.S. EPA, 1987).

The rate at which VSL increases with income growth (the income elasticity[1]) is not well estimated. The primary source of VSL estimates, compensating-wage-differential studies, usually do not provide information about the income elasticity, because the wage rate is the dependent variable and so income cannot be used as an explanatory variable.

The income elasticity can be estimated by meta-analysis of compensating-wage-differential studies where the study populations differ in income, risk, and other factors, but these studies lack power. Liu et al. (1997) estimated the relationship between VSL, income, and workplace-fatality risk for a sample of 17 compensating-wage-differential studies in the US and other industrialized countries. Their point estimate for the income elasticity is 0.54, with a standard error of 0.85. Mrozek and Taylor (2002) expanded on this approach by including multiple VSL estimates from each of 33 wage studies and controlling for the average wage, risk, and other factors. They report four specifications yielding estimated elasticities of VSL with respect to the wage rate between 0.36 and 0.49 with standard errors of 0.20 and above.

CV studies elicit WTP directly and can be used to estimate the income elasticity of VSL. Typical estimates range from 0.2 to 0.5. For example, Jones-Lee et al. (1985) estimated values of 0.25 to 0.44, Mitchell and Carson (1986) estimated 0.35, and Corso et al. (2001) estimated 0.41.

Subramanian and Cropper (2000) asked respondents to choose between different public programs to reduce health risks, and then asked how much more effect (in terms of lives saved) the less preferred program would need to be to make the respondent indifferent between programs. In each case, the risks presented the same health endpoint but differed in delay until benefits would be achieved, voluntariness, controllability, and other factors. Using a multivariate regression to control for the effects of various factors, Subramanian and Cropper (2000) found that people discounted for delay. They estimated a marginal rate of substitution of –0.15, which implies that a 1.5 percent increase in the number of lives saved would compensate for a 10 percent increase in delay.

Hammitt and Liu (2004) use CV to test for the effect of latency on WTP to reduce the risk of a fatal disease from environmental pollution in Taiwan. The authors find that respondents discount for the latency period between exposure to environmental contaminants and development of any resulting disease at a rate of 1.5 percent per year, and that WTP depends on the payment mechanism, affected organ, and environmental pathway.

WTP to reduce exposure to environmental pollution was not sensitive to the latency period between exposure and manifestation of disease. The insensitivity of WTP to latency suggests that respondents anticipate that their VSL will grow over time at a rate about equal to their discount rate.

In summary, the effects of latency on WTP to reduce own mortality risk are unknown. In theory, latency increases WTP if individual VSL increases faster than the interest rate, and decreases WTP otherwise. Empirical studies have not resolved this ambiguity.

[1] Carson et al. (2001) note that the income elasticity of demand and income elasticity of WTP are fundamentally different. The former describes how the quantity demanded increases with income while the latter describes how WTP for a fixed quantity of a good changes as income increases.

Magnitude of Cancer Premium. The value of preventing a fatal cancer is often considered to be greater than the value of preventing a fatal trauma in a workplace or transportation accident. Cancer is also frequently viewed as more threatening than other degenerative conditions, such as heart disease. A striking example is provided by the controversy over whether to encourage hormone replacement therapy for postmenopausal women. Therapy reduces risk of heart disease and hip fracture but increases the risk of breast and endometrial cancers. Because heart disease is five times more likely to kill a woman than is breast cancer, the net effects of treatment are substantial with gains in life expectancy as large as three years (Col et al., 1997).

There are a number of differences between cancer and accidental fatalities that might affect relative WTP to reduce each risk, including the often protracted suffering from cancer before death and the knowledge with cancer that one's condition will deteriorate and lead to death. Despite the plausibility that there may be a "cancer premium," the empirical literature supporting this supposition is limited. There are a few studies that provide information about the relative value of reducing risks of cancer and of acute trauma (e.g., motor vehicle fatality) but no studies of which we are aware have compared the value of reducing risks of cancer and of other fatal disease.

Jones-Lee et al. (1985) asked respondents to choose between public programs that would reduce the number of people dying in the next year by 100 from one of three causes (motor-vehicle accidents, heart disease, and cancer), and to indicate how much they would voluntarily contribute to reducing the number of deaths from the cause they selected. A large majority of respondents (76 percent) chose to reduce cancer deaths, and the mean voluntary contribution was larger for cancer than for the other causes. Interpreting the mean contributions as estimates of WTP yields a VSL of £23 million for cancer, £13 million for heart disease, and £7 million for motor vehicle accidents.

Savage (1993) asked survey respondents to allocate a hypothetical $100 contribution to research intended to reduce risks of stomach cancer, household fires, commercial-airplane accidents, and automobile accidents. He found that respondents would allocate the largest amount to stomach cancer ($47) with much smaller amounts ($15-$21) to the other risks. Although this study suggests greater WTP to reduce cancer risks, it does not measure individual WTP to reduce own risk. The value of research on methods to reduce risk of cancer (or the other fatality risks) depends on the probability that the research will identify interventions to reduce the risk, the magnitude of the risk reduction produced by the interventions, and the cost of implementing them. None of these parameters were specified, and so we cannot know what assumptions respondents made about them. In addition, the pattern of responses seems inconsistent with a measurement of WTP. The optimal response is to allocate all $100 to whichever risk the respondent believes will benefit most, since significant diminishing marginal efficacy of spending is implausible for contributions of $100.

McDaniels et al. (1992) conducted a CV study with only 55 respondents to estimate WTP for programs to reduce a wide range of health risks. The programs were described as public goods that would reduce risks to the relevant populations, not only to the respondent. The authors also elicited risk-perception variables, such as dread. They found that WTP to reduce risk was positively associated with dread.

Magat et al. (1996) used a risk-risk survey to elicit preferences for reductions in the risk of fatal automobile accidents and three chronic diseases: terminal lymph cancer, curable lymph cancer,

and non-fatal nerve disease. The latency periods for the diseases were not specified in the survey instrument. The median respondent was indifferent between equal reductions in the probability of terminal lymph cancer and of fatal automobile accident, suggesting that there is no cancer premium or that any cancer premium is offset by an assumed difference in latency. The loss in utility due to curable lymph cancer and non-fatal nerve disease were estimated as 58 percent and 40 percent as great as the loss from a fatal automobile accident, respectively, which suggests that the utility loss from lymph cancer morbidity is 45 percent larger than the loss from nerve disease.

Hammitt and Liu (2004) also examined whether respondents were willing to pay more to reduce liver cancer versus liver disease associated with contaminated drinking water, as well as lung cancer versus lung disease associated with industrial air pollution. The authors estimate that WTP to reduce the risk of cancer is about one-third larger than WTP to reduce risk of a similar chronic, degenerative disease.

2.3. Household WTP as a Measure of the Value of Children's Health

There are a variety of reasons why children's own WTP for health and safety initiatives are not appropriate measures of the value of these goods to children. One obvious issue is that society does not generally view children as autonomous economic agents. Most children do not earn income or make economic choices regarding their health and well-being. Children also differ from adults in their view of death, and may exhibit higher degrees of risk-taking behavior, perhaps because of their undeveloped cognitive abilities and limited practical experience (Harbaugh, 1999). Young children often have difficulty imagining and understanding death in the same way that adults do. They may instead view death as a type of sleep or as an event that happens only to bad people (Carey, 1985). Another difference from adults is that both children and adolescents have shorter time horizons, discount the future at higher rates, and often underestimate the value of future consumption (Krause and Harbaugh, 1998; Harbaugh, 1999). In short, all of these observed differences present problems for the standard economic assumptions of informed and rational behavior.

While children's own WTP may be an inappropriate measure of value, household WTP is an appropriate starting point. Understandably, parents know and care about their children's health, and they are accustomed to making economic decisions that will affect their children. To some extent, economists may view parental choices as altruistic behavior, but they may also regard households as unitary economic agents, with preferences and behaviors that are the result of some intra-household decision-making process.

Indeed, although most of the literature on the value of statistical life treats the concept as measuring an individual's rate of substitution between income and mortality risk, in both theory and practice it seems equally tenable to interpret this literature as measuring household WTP for changes in mortality risk. In some cases, the change in mortality risk is to a defined individual (e.g., the worker in studies of compensating wage differentials). In other cases, the risk change may benefit the entire household (e.g., studies valuing the risk of residential proximity to hazardous-waste sites, Smith and Desvousges, 1987). In all cases, the opportunity cost of a mortality risk reduction is smaller household income. Depending on how households allocate consumption among their members, some or all of them may have lower consumption as a result.

3. Survey

We will design and conduct a stated-preference survey to elicit values for reductions in mortality risks that vary in the baseline probability of illness, reduction in probability, latency of symptoms, disease type, symptom detail, and whether the exposure occurs to a child or to an adult. This section describes the survey instrument and sample.

3.1. Survey Instrument

The survey includes a dichotomous-choice experiment in which respondents decide whether to purchase a safer but more expensive food. The survey instrument is organized as follows. First, respondents are asked about their knowledge of foodborne pesticide risk and their perception of how common it is compared with other health and safety risks. Second, respondents assess their current health using a visual analogue scale (VAS) and the Health Utilities Index Mark 3 (HUI). The VAS is a numbered line with endpoints of 0 and 100 labeled "equivalent to dead" and "perfect health," respectively. The HUI is a generic, preference-based, multiattribute health-status classification system and index that is widely used as a measure of HRQL in clinical studies, population health surveys, and economic evaluation (Feeny et al., 2002). The HUI classifies health according to the degree of function on eight dimensions: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain. For each dimension, there are five or six levels of functional impairment that range from complete function to severe impairment.

Third, respondents complete a tutorial designed to help them practice making tradeoffs between the price and safety of food. The tutorial also familiarizes respondents with a visual aid that communicates the probability of risks (Corso et al., 2001). The visual aid contains red and white areas that represent 10,000 apples, where the fraction of the area that is colored red equals the probability that an apple contains unsafe levels of pesticide.

Fourth, respondents are asked to consider buying food for a meal that only they will eat. Respondents are asked whether they eat a type of food randomly selected from the set {apples, grapes, lettuce}. If they do not eat the selected food, respondents are asked about another randomly-selected food. After answering questions about how often they eat the food and how much they typically eat, respondents are presented with a description of the symptoms of a fatal disease caused by consuming pesticide in the food. Respondents are then told their baseline probability of illness (either 2 in 100,000 or 4 in 100,000 per year) and informed that they could reduce their risk to 1 in 100,000 per meal by purchasing a safer but more expensive brand of food. The baseline probability of illness and reduction in probability are communicated using the visual aid described above. The risk reduction is described as produced by a stringent pesticide safety program established and monitored by the United States Government. Respondents are told that while the food produced by the pesticide safety program is safer to humans than conventional food, the program is not an organic farming practice, nor does it affect other animals or the environment any differently than conventional farming. WTP to reduce the probability of illness is elicited using double-bounded, dichotomous-choice questions. Each respondent is asked if he would purchase the safer food if the extra cost per year were a randomly selected amount from the set {$10, $20, $50, $80, and $100}. There is one follow-up question, in which the bid is equal to twice the initial bid if the respondent is willing to pay the initial amount, and equal to half the initial bid otherwise. Finally, respondents answer follow-up

questions about their food-handling practices, acceptance of the hypothetical scenario, and relevant personal characteristics.

Each respondent is asked to value three health-risk reductions that vary in baseline probability of illness, reduction in probability, severity and duration of symptoms, conditional probability of mortality, and type of food affected. Using a full factorial design, the risk attributes are randomly assigned so that each of the possible combinations is asked of some respondents. Table 1 shows the risk attributes, which we describe in more detail below.

Table 1. Risk Attributes (Full-Factorial Design)

Individual Exposed	Annual Risk Reduction	Latency	Disease Type	Symptom Detail	Type of Food
Self	1 in 100,000	1 year	Cancer	Brief	Apples
Child	3 in 100,000	10 years	Non-Cancer	Detailed	Grapes
Other Adult		20 years			Lettuce

Person Exposed. Depending on their household composition, respondents are asked about reducing risks to their own health, the health of a child, or the health of another adult. Respondents who live in a household with at least one child under the age of 18 and at least one other adult are asked about reducing one risk to their own health, one risk to the health of a randomly-selected child from their household, and one risk to the health of a randomly-selected adult from their household (in random order). Respondents who live in a household with at least one child under the age of 18 and no other adults are asked about reducing one risk to their own health and two risks to the health of a randomly-selected child from their household (in random order). Respondents who live in a household with at least one other adult and no children under the age of 18 are asked about reducing two risks to their own health and one risk to the health of a randomly-selected adult from their household (in random order). Respondents who live alone are asked about reducing three risks to their own health, but are not presented with the same food twice.

Latency. The risks presented will differ in latency, defined as the period between the time when an individual is exposed to an environmental contaminant and the time when he or she develops symptoms of disease or is diagnosed. Three latency periods (1 year, 10 years, and 20 years) will be considered. In the short latency case, respondents will be told that, if they develop the stated disease, symptoms will begin within a year and they will live only about two years longer. In the long latency cases, respondents will be told they will not know if they were sufficiently exposed to develop the disease until they experience symptoms about 10 years (or 20 years) in the future. After developing symptoms, the prognosis is identical to the short latency case.

Disease Type. WTP will be elicited for one or more disease pairs that consists of a specific form of cancer and a non-cancer disease that affects the same organ and has similar symptoms and prognosis. All diseases will be terminal. The symptom descriptions presented to respondents will be identical except for the name of the disease.

Symptom Detail. The symptom descriptions will be varied to provide different levels of detail. Our hypothesis is that the cancer premium may be sensitive to the comprehensiveness of the

symptom description. When respondents are given little or no information about the symptoms and prognosis of a disease other than its name, they may have a higher WTP to reduce the risk of "cancer," if cancer is generally perceived to lead to more severe morbidity than other fatal diseases. In this case, we might observe a substantial cancer premium. Alternatively, when the respondent is given extensive information about the symptoms associated with a disease, the additional information associated with knowing that the disease is a form of cancer rather than another fatal disease may have less impact, and so the magnitude of the cancer premium may be much smaller or non-existent.

Magnitude of Risk Reduction. The magnitude of the risk reduction will be varied across valuation tasks to provide information about whether the CV instrument produces WTP estimates that are sensitive to scope. Under conventional economic theory, WTP for a small reduction in mortality risk is nearly linear in the magnitude of the risk reduction. The sensitivity of estimated WTP to magnitude of risk reduction can be used as a diagnostic test of the performance of the survey instrument (Hammitt and Graham, 1999; Hammitt, 2000; Corso et al., 2001). If WTP is not proportional to the magnitude of risk reduction, then estimated VSL is sensitive to the arbitrary magnitude of the risk reduction offered.

Inadequate sensitivity of estimated WTP to magnitude of risk reduction has been a substantial problem in almost all CV studies of health risks. Hammitt and Graham (1999) identified 14 CV studies published from 1980 through 1998 that either reported a test of sensitivity to magnitude or provided enough information to enable them to conduct such a test. They found that although estimated WTP was sensitive to the magnitude of risk reduction (i.e., the estimated value of a larger reduction exceeded the estimated value of a smaller reduction) in 11 cases, WTP was inadequately sensitive (i.e., less than proportionate to magnitude of risk reduction) in all cases.

To test whether inadequate sensitivity to magnitude is a result of difficulties in communicating small risk changes to survey respondents, Corso et al. (2001) asked respondents to value reductions in automobile fatality risk. Corso et al. presented respondents with one of three visual aids (a field of 25,000 dots, a logarithmic risk ladder, or a hierarchical linear risk ladder) or no visual aid, and then elicited values for reducing annual risk by 5 or 10 in 100,000 from separate sub-samples. Corso et al. found that estimated WTP was sensitive to risk reduction for respondents presented with any of the visual aids, but not for the control group. Moreover, the hypothesis that estimated WTP was proportionate to the risk reduction could not be rejected for the groups of respondents presented with either the dots or the logarithmic risk ladder. The study by Corso et al. suggests that CV can be used to estimate WTP for small risk reductions that are consistent with economic theory, and hence that near-proportionality of estimated WTP to risk reduction may be used as a test for the validity of CV estimates (Hammitt, 2000).

For the valuation tasks, we anticipate using two magnitudes of risk reduction: 1 in 100,000 per year and 3 in 100,000 per year. These risk reductions are small enough to be relevant to the pesticide risks of concern, yet are sufficiently far apart that WTP should differ substantially (by a factor of three). The risk reductions will be accompanied by visual aids that were found to work well by Corso et al. (2001). In addition, describing risks using a common denominator is anticipated to assist respondents in recognizing differences between the two risk magnitudes.

WTP will be elicited using double-bounded discrete-choice questions (Hanemann et al., 1991). Each respondent will be randomly assigned to one of five initial bid values ($10, $20, $50, $80, and $100) that represent the additional cost of meals made with food containing reduced

pesticide levels. There will be one follow-up question, where the bid is equal to twice the initial bid if the respondent indicates he would be willing to pay the initial amount, or equal to half the initial bid otherwise. Respondents will receive different initial bids for the first and second valuation questions to minimize follow-up effects (e.g., giving the same "yes" or "no" response to the second valuation question as given to the first).

Discrete-choice questions are often preferred to open-ended questions because they appear to be easier for respondents to answer. The referendum format is incentive-compatible and was recommended by the NOAA panel (Arrow et al., 1993). In addition, dichotomous-choice questions are often considered superior to open-ended, bidding-game, and payment-card formats, because they do not create anchoring effects. The double-bounded format provides substantially greater information per respondent than a single-bounded format. The corresponding double-bounded or interval-data models of WTP have been shown to produce more efficient estimates than those obtained using only the single-bounded payment format (Hanemann et al., 1991; Alberini, 1995). Although the initial bid may influence responses to the follow-up question (Alberini et al., 1997), we will calculate single-bounded estimates using only the response to the first valuation question to investigate the magnitude of any follow-up effect.

3.2. Sample

The survey will be fielded to members of a demographically representative panel maintained by Knowledge Networks. Households are recruited to the panel using random digital dialing and provided free Internet access and hardware, such as MSN® TV, as a participation incentive. In total, 2,000 interviews will be completed. We plan to over-sample households with children so that we have sufficient responses about reducing risks to children's health.

4. Analysis

Using theory to inform model specification, we will develop an empirically estimable model relating WTP to health risk attributes, the respondents' socioeconomic characteristics, and variables characterizing risk attitudes. For the purposes of illustration, consider the following model:

$$\log(WTP) = \alpha + \beta_i X_i + \gamma_i R_i + \varepsilon \tag{3}$$

where X_i is a vector of covariates describing the respondent (e.g., age, sex, health, education, marital status, household income) and the person at risk (e.g., age, sex, health), R_i is a vector of risk characteristics (e.g., latency, disease type, magnitude of risk reduction), and ε is an error term.

Because WTP is elicited using double-bounded binary choice questions, individual WTP is interval censored. We observe only the upper and lower bounds on an individual's WTP (which may be infinite and zero, respectively). Equation (3) will be estimated using maximum likelihood methods (Alberini, 1995) implemented in standard statistical software (e.g., SAS). Estimates will be obtained using alternative parametric assumptions regarding the distribution of the error term, including a "mixed model" which allows for the possibility that a finite fraction of respondents have WTP equal to zero (Werner, 1999).

In order to test for differences in WTP to reduce risks to children versus adults, we will include a dummy variable indicating whether the individual who benefits from the reduced pesticides is a child or adult. The effects of age will also be evaluated using dummy variables; e.g., the "child" dummy may be replaced by a series of dummy variable for age category (e.g., 0-5, 6-12, and 13-18 years). Adult age will be represented using dummy variables for age categories and, alternatively, using simple polynomial functions (e.g., age, age^2).

To determine how WTP depends on other characteristics of the health risks, we will estimate a regression that includes dummy variables for various risk characteristics, such as the degree of latency, whether the risk causes cancer or not, and the level of detail of the symptom description provided. We will also interact the dummy variable for long-latency with the child age dummy variables, to determine whether the valuation of latent risks (where exposure occurs to a child but the risk only manifests to the adult), is sensitive to the age of the child at time of exposure.

We will incorporate several methods to test the validity of estimated WTP. First, we will estimate the coefficient on risk magnitude to determine how WTP depends on the magnitude of risk reduction. Under standard economic theory, WTP should be almost exactly proportional to the magnitude of risk reduction for small risk reductions (where income effects are negligible) (Corso et al., 2001; Hammitt, 2000). Hence, if we can reject the hypothesis that WTP for the 3 in 100,000 risk reduction is not three-times WTP for the 1 in 100,000 risk reduction, this will provide evidence suggesting that respondents did not accurately report their WTP for risk reduction. Given the difficulties in communicating and comprehending small risk changes, this proportionality test is quite demanding and has only once been satisfied, to our knowledge (Corso et al., 2001). A weaker test is to require that estimated WTP be statistically significantly larger for the larger risk reduction. Even this test is frequently not satisfied by prior studies, perhaps because of inadequate attention to communicating the magnitude of risk changes (Hammitt and Graham, 1999).

Additional evidence regarding the validity of estimated WTP will come from use of follow-up questions and examining the relationship between individual WTP and covariates that are anticipated to be associated with it. Follow-up questions will include some addressed to accuracy of risk perception (e.g., asking respondents if they believe they are more likely to get sick or injured, or to die, from, e.g., pesticides on food, microbial contaminants on food, heart disease, or other causes). Previous studies have found some ability to accurately rank these risks (Williams and Hammitt, 2001), and we anticipate that respondents with a better sense of the relative probabilities of these events would give more valid answers about WTP. Other questions will address respondents' health habits both for themselves (e.g., dietary choices, smoking, drinking, exercise, preventive care, seatbelt use) and for their children (e.g., dietary choices, preventive care, seatbelt and child seat use, bicycle helmet use, childproofing home by storing hazardous materials carefully and covering electrical sockets). We anticipate that people who adopt healthier habits may also have greater WTP for reductions in pesticide-related risk. There is some collaborating evidence that those with poorer health habits (smokers and those who do not use automobile seatbelts) have smaller WTP to reduce risk of workplace injury (Hersch and Viscusi, 1990; Hersch and Pickton, 1995; Viscusi and Hersch, 2001).

References

Agee, M.D., and T.D. Crocker, "Parental Altruism and Child Lead Exposure: Inferences from the Demand for Chelation Therapy," *Journal of Human Resources* 31: 677-91, 1996.

Alberini, A., "Efficiency vs Bias of Willingness-to-Pay Estimates: Bivariate and Interval-Data Models," *Journal of Environmental Economics and Management* 29: 169-180, 1995.

Alberini, A., B. Kanninen, and R.T. Carson, "Modeling Response Incentive Effects in Dichotomous Choice Contingent Valuation Data," *Land Economics* 73: 309-324, 1997.

Arrow, K., R. Solow, P. Portney, E. Leamer, R. Radner, and H. Schuman, "Report of NOAA Panel on Contingent Valuation," *Federal Register* 58(10): 4601-4614, 1993.

Becker, G., "Altruism in the Family and Selfishness in the Market," *Economica* 48: 1-15, 1981.

Bruhn, C.M., K. Diaz-Knauf, N. Feldman, J. Harwood, G. Ho, E. Ivans, L. Kubin, C. Lamp, M. Marshall, S. Osaki, G. Stanford, Y. Steinbring, I. Valdez, E. Williams, and E. Wunderlich, "Consumer Food Safety Concerns and Interest in Pesticide-Related Information," *Journal of Food Safety* 12: 253-262, 1992.

Buzby, J., R. Ready, and J. Skees, "Contingent Valuation in Food Policy Analysis: A Case Study of a Pesticide-Residue Risk Reduction," *Journal of Agricultural and Applied Economics* 27: 613-625, 1995.

Carey, S., *Conceptual Change in Childhoo*d, MIT Press, Cambridge MA, 1985.

Carson, R.T., N.E. Flores and N.F. Meade, "Contingent Valuation: Controversies and Evidence," *Environmental and Resource Economics* 19: 173-210, 2001.

Col, N.F., M.E. Eckman, R.H. Karas, S.G. Pauker, R.J. Goldberg, E.M. Ross, R.K. Orr, and J.B. Wong, "Patient-Specific Decisions About Hormone Replacement Therapy in Postmenopausal Women," *Journal of the American Medical Association* 277: 1140-1147, 1997.

Corso, P.S., J.K. Hammitt, and J.D. Graham, "Valuing Mortality-Risk Reduction: Using Visual Aids to Improve the Validity of Contingent Valuation," *Journal of Risk and Uncertainty* 23: 165-184, 2001.

Cropper, M.L, and P.R. Portney, "Discounting and the Evaluation of Lifesaving Programs," *Journal of Risk and Uncertainty* 3: 369-379, 1990.

Cropper, M.L., and F.G. Sussman, "Valuing Future Risks to Life," *Journal of Environmental Economics and Management* 19: 160-174, 1990.

Cropper, M.L., S.K. Ayded, and P.R. Portney, "Preferences for Life Saving Programs: How the Public Discounts Time and Age," *Journal of Risk and Uncertainty* 8: 243-265, 1994.

Drèze, J., "L'Utilitè Sociale d'une Vie Humaine," *Revue Française de Recherche Opèrationelle* 6: 93-118, 1962.

Fu, T.-T., J.-T. Liu and J.K. Hammitt, "Consumer Willingness to Pay for Low-Pesticide Fresh Produce in Taiwan," *Journal of Agricultural Economics* 50: 220-233, 1999.

Gerking, S., M. De Haan, and W. Schulze, "The Marginal Value of Job Safety: A Contingent Valuation Study," *Journal of Risk and Uncertainty* 1: 185-199, 1988.

Gray, G.M., and J.K. Hammitt, "Risk/Risk Tradeoffs in Pesticide Regulation: An Analysis of the Public Health Effects of a Ban on Organophosphate and Carbamate Pesticides," *Risk Analysis* 20: 665-680, 2000.

Hammitt, J.K. and J.-T. Liu, "Effects of Disease Type and Latency on the Value of Mortality Risk," Harvard School of Public Health, 2002.

Hammitt, J.K., "Consumer Willingness to Pay to Avoid Pesticide Residues," *Statistica Sinica* 3: 351-366, 1993.

Hammitt, J.K., "Risk Perceptions and Food Choice: An Exploratory Analysis of Organic- Versus Conventional-Produce Buyers," *Risk Analysis* 10: 367-374, 1990.

Hammitt, J.K., "Using Willingness to Pay to Value Food Safety: Theory and Empirical Methods," in *Setting Food Safety Priorities: Toward a Risk-Based System*, Resources for the Future, Washington, DC, in press, 2002.

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

Hammitt, J.K.,"Evaluating Contingent Valuation of Environmental Health Risks: The Proportionality Test," *Association of Environmental and Resource Economists Newsletter* 20(1): 14-19, May 2000. Reprinted in *Stated Preference: What Do We Know? Where Do We Go? (Proceedings)*. Report number EE-0436, U.S. Environmental Protection Agency, October 2000.

Hanemann, W.M., J. Loomis, and B. Kanninen, "Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation," *American Journal of Agricultural Economics* 73: 1255-1261, 1991.

Harbaugh, W.T, "Valuing Children's Health and Life: What Does Economic Theory Say about Including Parental and Societal Willingness to Pay?" EPA workshop on Valuing Health for Environmental Policy with Special Emphasis on Children's Health Protection, March 1999.

Hersch, J., and T.S. Pickton, "Risk-taking Activities and Heterogeneity of Job-Risk Tradeoffs," *Journal of Risk and Uncertainty* 11: 205-217, 1995.

Hersch, J., and W.K. Viscusi, "Cigarette Smokers, Seatbelt Use, and Differences in Wage-Risk Tradeoffs," *Journal of Human Resources* 25: 202-227, 1990.

Johannesson, M., P-O Johansson, and K-G Lofgren., "On the Value of Changes in Life Expectancy: Blips Versus Parametric Changes," *Journal of Risk and Uncertainty* 15: 221-239, 1997.

Johansson, P-O., "Altruism and the Value of Statistical Life: Empirical Implications," *Journal of Health Economics* 13: 111-118, 1994.

Jones-Lee, M., "The Value of Changes in the Probability of Death or Injury," *Journal of Political Economy* 82: 835-849, 1974.

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

Comments on "Use of Contingent Valuation to Elicit Willingness-to-Pay for Benefits of Developmental Health Risk Reductions"
by Katherine von Stackelberg and James K. Hammitt

- Why not ask questions about the household, or if the respondent is a parent? This would impact how well the respondent could identify with questions about a hypothetical child.

- It is an important result that respondents were willing to increase their bids from their initial ecological bid when asked for a total bid (ecological and health), but not when the health bid was asked for first (especially since 63-74% indicated that they could separate the two endpoints).

- The standard gamble and time tradeoff questions seem like they would be difficult for respondents to truly understand and answer. Could a parent of a real 10 year old child really answer a question that trades off a small probability of death (or weeks of longevity – this one might be easier) to a reduced cognitive deficit that is relatively mild?
  - Those types of questions may possibly be easier for a non-parent to answer, however a non-parent, or maybe even to some extent a parent of only a baby, may not fully understand the implications of the trade-off
  - Because the QALY questions turn out to be significant in most of the models, I think the responses could be viewed as representing respondents' perceptions about how a cognitive deficit would affect a child's quality of life.

- Overall, I found the paper interesting and could be a useful approach in getting values for mild developmental effects.

Comments on "Parental Decision-Making and Children's Health," by Ann Bostrom, Sandra Hoffmann, Alan Krupnick and Wictor Adamowicz with Robin Goldman and Michael McWilliams

- Well-written and very fun to read – certainly made me reflect on decision making in my own household.

- Results highlight that there is a lot of disagreement in marriages/households about factual information as well as about how household decisions are made.
  - Even factual information provided by couples separately contained differences.
  - Couples not knowing exact percentages of contributions to household income, spouse's income, or total household income didn't surprise me.
  - It makes sense that if spouses specialize in decision domains such as paying bills or managing finances that the "specialist" would know more (e.g. I pay the bills in our house and my husband doesn't know exactly how much I make)
  - Spouses may have different concepts of income (e.g., I would answer an annual household income question assuming just my husband's salary however when he answers, he includes bonuses and extra fees). Respondents having jobs in sales

where a significant part of salary is based on commission, could introduce answers that vary between spouses.

- The survey collected a lot of information about decision-making behavior with a section specifically geared towards marital adjustment, but I'm wondering if you could ask questions to reveal the personality traits of each spouse as well. Personality could influence decision-making in household.

- The results imply that it does matter who in a household is interviewed for a survey. It would be nice to see more discussion on how the couples separately and together dealt with the hypothetical lead paint decision scenario and did it correspond to the results from the rest of the survey. Does a respondent consider other household members' preferences when answering individually?

- I'm excited to see results of the future WTP survey and answers to the questions: How does separate WTP for each spouse compare to each other and to a jointly arrived at WTP? What are some questions that could be asked of individuals to determine how representative of household preferences their own answers are?

Comments on "Value of Reducing Children's Mortality Risk: Effects of Latency and Disease Type," by James K. Hammitt and Kevin Haninger

Paper did not yet include results so I only have a few comments.

Nice survey of the literature on several different dimensions of WTP for mortality risk (exposure to child or adult; exposure to self or other household member; the fatality from disease is immediate or latent; the fatal disease is a cancer or non-cancer; the amount of information provided about the fatal disease).

How much information are respondents given about the pesticide safety program? Are they told specifics about how it works? For example, if they are told that there is a special wash applied to produce after it is harvested, there is clearly no ecological benefit. But if the program is less or different pesticide use, then respondents may still confer an ecological benefit to the program even if you state there isn't any.

Are respondents asked about organic food purchases?

I'm not sure how able respondents will be at comprehending a risk reduction to only one member of the household – most food brought into a household is consumed by everyone in the household (with some exceptions). Could you also ask a question about reducing the risk to the entire household?

Childlike Values:
Measurement Strategies for Children's Health Values

F. Reed Johnson
Senior Fellow
Research Triangle Institute

Discussant Remarks
U.S. EPA NCER/NCEE Workshop
April 2006

The goal of the STAR program is to support research that translates existing methods and findings into policy-relevant research and to fill in gaps in knowledge that limits our ability to assess the efficiency of environmental regulations. After three decades of environmental and health valuation research, we have acquired some respect for the difficulties inherent in nonmarket valuation. These difficulties are magnified when we attempt to estimate willingness to pay to reduce risks to health and safety.

From an individual's point of view, most environmental regulations reduce relatively small risk exposures by relatively small amounts. We thus encounter various impediments to obtaining valid and reliable values for such risk reductions, including among other challenges, respondent innumeracy, sensitivity to risk framing, sensitivity to features of the risk that are independent of probability or health endpoint, poor descriptive power of the standard expected utility model. As evidence accumulates regarding the differential sensitivity of children to environmental hazards, demand has increased for valid and reliable estimates of the value of reducing such risks. The papers presented in this workshop evaluate the extent to which people are willing to accept tradeoffs between money and children's health risks and what methods are likely to give us valid, policy-relevant estimates.

The three papers in this session offer different strategies for answering such questions. Von Stackelberg and Hammitt compare classic contingent valuation, standard gamble, and time tradeoff elicitation formats. They obtain estimates of $466 per IQ point for developmental impairment, or $109,000 per QALY. Hammitt and Haninger offer a research prospectus to evaluate risks from pesticide contamination of food using classic contingent valuation, visual analog scale, and health utilities obtained from the Health Utilities Index Mark 3 health-related quality of life instrument. They propose to evaluate the effect of outcome latency, disease type, and information treatment on values measured in each way. Finally, Bostrom, Hoffman, Krupnick, and Adamowicz offer some preliminary results from a survey of household decision patterns. They find that about 32% of surveyed couples' preferences were disjoint for major purposes and for financial decisions generally. They also find that most couples were willing to consider cost-efficacy tradeoffs for lead exposure.

Von Stackelberg and Hammitt, "Use of Contingent Valuation to Elicit Willingness to Pay for the Benefits of Developmental Health Risk Reductions"

The authors set out to determine whether WTP is proportional to risk reduction and to obtain WTP per QALY. The standard-gamble (SG) elicitation format is relatively unfamiliar to environmental economists. This method obtains a von Neumann-Morgenstern utility index scaled between death, assumed to have utility equal to zero, and perfect health, assumed to have utility equal to one. The utility index is the probability for a lottery between perfect health and instantaneous, painless death that makes respondents indifferent between the lottery and a sure outcome—in this case a specified developmental disability. The elicitation generally is assumed to be independent of the usual factors we generally use to condition utility such as income, demographic factors such as age and gender, duration of the certain condition, treatment options, and other context factors. Moreover this approach requires assuming preferences conform to the expected-utility model that generally performs poorly in describing actual behavior under risk. While SG is popular among (mostly non-economist) health researchers, it is hard to justify suspending so many considerations that guide preference research in virtually every area of applied economics other than health.

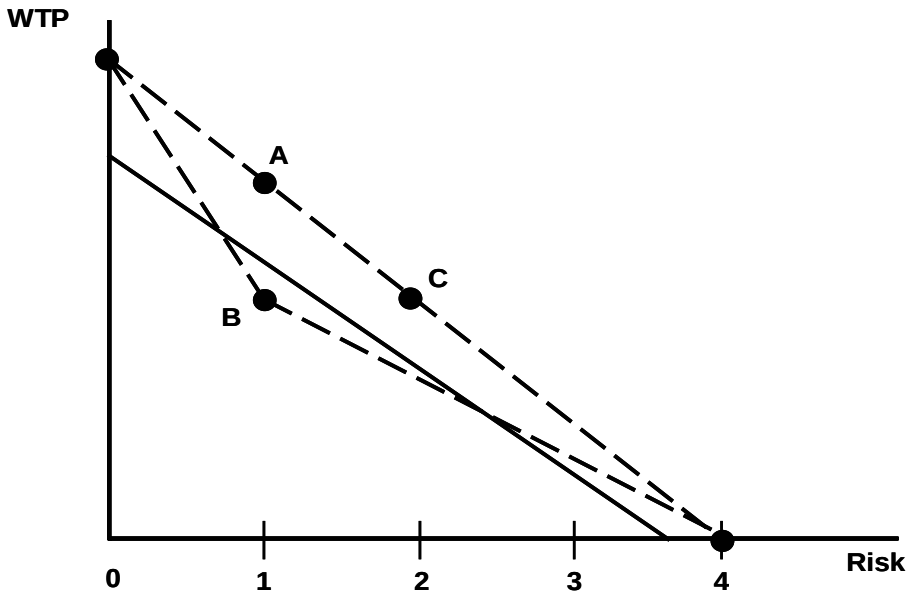
The authors follow the environmental economics convention of using a double-bounded format for both the standard gamble and CV questions. The convention in health economics is to use a bidding game for standard gamble elicitations. It is likely that the two methods would yield different utility weights. The authors acknowledge known problems with double-bounded CV formats. It isn't clear later whether they found no significant anchoring bias and used the double bounded estimator or appealed to Alberini's finding and pooled the first and second bids. The strategy for the second-bid starting point conditions on the first-bid starting point. While logical, it also imposes some degree of monotonicity and consistency in responses that might not have resulted from randomization.

Economical administration of stated-preference surveys conflicts with OMB requirements that for high response rates and validated claims of representativeness. OMB appears uncompromisingly opposed to using web panels to collect data in support of regulatory decisions. Nevertheless, the authors assert that the Knowledge Networks panel "is the

only available method for conducting internet-based survey research with a nationally representative probability sample." It is worth noting that (1) random-digit dialing no longer ensures reaching a representative sample; (2) there is selection bias in the sample that agrees to join the KN web panel once contacted; (3) there is selection bias in attrition from the panel. That doesn't mean we shouldn't use web panels, however. Both Knowledge Networks and other web panels use sophisticated weighting techniques to correct for possible selection bias. It is difficult to imagine any other alternative that is consistent with the actual resources available to conduct stated-preference studies.

The assertion that the estimated WTP is approximately proportional to risk reduction appears to rely on a weak test. In fact, there are competing hypotheses to support an expectation that WTP is nonlinear in probability. One possibility is that risk preferences follow rank-dependent utility axioms rather than expected-utility axioms. Rank-dependent utility overweights small probabilities and underweights large probabilities. Figure 1 indicates the possible effect of such weighting. Expected utility dictates that WTP at risk level 1 be at point A. However, if probabilities between 0 and level 1 are weighted more heavily than probabilities between levels 1 and 4, then WTP at risk level

Figure 1. Nonlinear Effect of Risk on WTP



1 will be at some point B. Alternatively, if the risk levels 1-4 are very small probabilities, respondents may find it difficult to discriminate between absolute differences. They may

simply recode 0 as “low”, 2 as “medium”, and 4 as “high” and set the utility differences to be equal. That would yield WTP at the medium level at point C. If plotted against nominal risk, C would look like B. If plotted against equally spaced categories, C would lie on a straight line. If preferences follow either B or C, estimating WTP as a linear function, as shown by the solid line, might not detect the kink and fail to reject a hypothesis of linearity. A better practice is to estimate the model using categorical risk levels and test whether utility differences are proportional to nominal risk values or not.

Cost per QALY is widely computed in health economics to evaluate the relative efficiency of alternative interventions. However, knowing that the cost per QALY for one policy is less than that for another policy does not provide any guidance about whether either policy is worth adopting. I am troubled by using WTP/QALY to solve the lack of a cost-effectiveness threshold. Lack of a threshold is the result of resistance to monetizing benefits to facilitate a real cost-benefit analysis in health economics, much as environmentalists have resisted monetizing environmental benefits for environmental policy analysis. Practitioners argue QALYs avoid all the equity baggage of WTP. If QALYs are all we need, why try to find a WTP value to do the analysis in QALY terms? Doing so combines incompatible conceptual models (Johnson, 2005).

The authors perpetuate a common confusion in comparing their WTP per QALY estimates with calculations reported in the literature based on the value of a statistical life (Hirth, 2003). Apart from the well-known problems in obtaining valid VSL estimates, it is inappropriate to divide VSL by life expectancy and interpret that as WTP per QALY. A statistical life year is not the same as a year of life, much less the same as a year of life in perfect health. That is exactly the misinterpretation that scandalizes non-economists when they hear us argue about the dollar value of a (statistical) year of life.

While the analysis in this paper is carefully done, there are several puzzling results that might warrant additional thought. For example, the significant negative sign on the reading-comprehension health endpoint is counter-intuitive and would benefit from some explanation. The statement that WTP was 33% lower for IQ compared to reading comprehension seems inconsistent with the wrong sign on reading comprehension. Furthermore, the significance of the IQ endpoint parameters is weaker than expected and values per unit IQ loss are an order of magnitude lower than the expected income

loss. The authors speculate that respondents are discounting the effect on expected lifetime income inappropriately, but there may be other explanations.

Hammitt and Haninger, "Value of Reducing Children's Mortality Risk: Effects of Latency and Disease Type"

Jim Hammitt conducted a well-conceived study for EPA in 1986 entitled "Organic Carrots: Consumer Willingness to Pay to Reduce Food-Borne Risks." I was interested in seeing how this plan to conduct a study on a similar topic reflected how much his and our understanding of risk-preference elicitation methods has evolved over the intervening 20 years. I think he would agree that we have not progressed as far as we would have liked.

The authors propose a repeated-CV design, along with visual analog scale and HUI-Mark 3 QALY weights to obtain QALY estimates. They propose to evaluate the insensitivity to latency noted in previous studies, although they appear to be unaware of the latency results reported in papers by Cameron and DeShazo. The proposed risk reduction from 2 or 4/100,000 to 1/100,000 may invite respondents to recode such small numbers into low, medium, and high categories. It might be prudent to include a scope test to see whether respondents are paying attention to absolute risk levels.

Asking only 3 repeated CV questions doesn't impose much of a cognitive burden on respondents. It is likely they could answer 10 or 12 questions, which would greatly increase the power of the sample. With careful attention to the experimental design, the data might provide enough information to estimate hierarchical Bayes individual-level estimates of WTP.

Bostrom, Hoffmann, Krupnick, and Adamowicz, "Parental Decision Making and Children's Health"

This study is an interesting first start at understanding how to interpret household preferences based on responses from one member of the household. This work is long overdue. The standard practice in stated-preference research is to administer the survey to one household member. The preference-elicitation question may or may not explicitly ask the respondent to indicate household preferences. In any case, in the absence of data or theory to help discriminate among household members, we simply

assume the observation represents an aggregation of household values. However, if spouses in the household specialize in decision making, then asking different spouses a WTP question is likely to lead to different answers.

In the next draft of the paper, it would help to be more explicit about what insights were obtained from the data about the basic research problem and how the results will be used to develop a better stated-preference instrument. For example, how might one adapt the standard time-to-think experiment? One possible explanation for differences between an immediate and a "considered" response is that the respondent takes the extra time to consult other decision makers in the household and construct a value that is a better aggregation of household preferences. Could the decision questions in this survey be adapted to measure what preference-aggregation process was used during the time to think?

There are several published studies on income-pooling experiments. (See, for example, Bateman and Munro, 2005.) Such experiments rely on actual decisions on lotteries with payoff rules designed to reveal how income is controlled within the household. It may be possible to extend these methods to explore how responsibility for expenditures in particular categories is allocated within a household.

The authors attribute the allocation of responsibilities on the basis of utility and bargaining power and thus the locus of decision making authority reveals the implicit weights attached to household members' utility functions. However, suppose spouses are highly altruistic and have good information about each other's preferences. Then allocation of decision making responsibilities might reflect comparative technical advantages—i.e. production-function factors—not welfare weights. A common example of the separation of preferences and allocation of responsibility is the "honey-do" list, suggesting that the wife's preferences dominate prioritizing household tasks, but the husband has responsibility for actually doing the tasks.

The introduction to this draft promises to employ a mental-models framework, but the focus is primarily on the cooperative household decision model. It is not clear to what extent these two frameworks are complements or substitutes. In any case, people may not be good at explaining decision processes after the fact. Well-known problems with

recall bias are likely to be even more serious in reconstructing subjective thought processes. Curiously, the authors average couple's answers in many cases, implying equal utility weights, which their conceptual framework suggests that is unlikely. The inverse correlation between income and joint decision making may simply indicate that joint decisions are time-intensive and the opportunity cost of time rises with income.

The evidence on disjoint reporting of supposedly factual data may be the most interesting feature of this study. I would have liked to see more effort to explain the direction and magnitude of disjoint responses. It is curious that 88% said decisions were made jointly, which isn't consistent with evidence on specific decisions. How do these results relate to the theoretical material? How might disjoint perceptions affect household decision making? It might be interesting to ask how responsibilities have changed over time. Suppose decision-making responsibility evolves over time as family circumstances change or couples gradually specialize. It is possible that disjoint responses are partly explained by husbands and wives averaging over different time periods. Perhaps the wives are recalling recent history and husbands are averaging over a longer period.

The main result from the quoted interview material seems to be that "a majority" were willing to consider tradeoffs. Of course, that result should be evident in a pretest of the instrument. Some of the quotes may reflect socially acceptable attitudes. We're actually less interested in their willingness to trade in the abstract than whether they are willing to accept tradeoffs in the specific context of a preference elicitation. I look forward to seeing how insights obtained from this study influence the design of a stated-preference survey to obtain true household values, including values for children's health.

References

Bateman I. and Munro A. 2005. An Experiment on Risky Choice Amongst Households, *The Economic Journal*. 115, C176-C89.

Hirth, R.A. et al. 2000. Willingness to pay for a quality-adjusted life year: In search of a standard. *Medical Decisionmaking* 20: 332–42.

Johnson, F.R. 1005. Einstein on Willingness to Pay per QALY: Is There a Better Way? Med Decis Making25: 607-608.

Summary of the Q&A Discussion Following Session IV

J.R. DeShazo, (UCLA)
NOTE: Dr. DeShazo's comments/questions were inaudible at first and are picked up here toward the end.
"I think, very importantly, people come to choices with subjective expectations that arise out of information they collected based on mental models they currently use. So, very often, subjective expectations about risk levels and risk reductions associated with different hazards and different programs are brought into the survey environment, and we have no idea really what's going on there.

Finally, in terms of the parent-child relationship, whether the parent is practicing altruistic paternalism or not is probably going to be a function of the age of the child. I can force my five-year-old to eat her vegetables, but I probably won't feel a responsibility to do that for my 25-year-old daughter. So, understanding the nature of the parental responsibility comes from understanding how they represent their role as a parent in their child's health."

Sandra Hoffmann, (Resources for the Future)
"One comment I'd like to make is on the relationship between the hazard and the health outcome: This is a classic way in which mental models are used. We didn't discuss this in our presentation today, but that's a major focus of the mental model study that we conducted. We structured what is called an "expert mental model" of the relationship between the environmental hazard and the risk that was peer-reviewed by a number of leading experts on children's lead hazards. That is being used as a basis to compare the parents' understanding of the relationship between lead exposure and health outcomes—and between mitigation and health outcomes. Our intention is to use that to help refine the way the risk is presented, and it's been used that way to improve risk communications in the past."

Alan Krupnick, (Resources for the Future)
Dr. Krupnick added, "Of course, we're planning on getting into the decision-making process mental model," and noted that they would be refining the work that was presented at the workshop. Addressing Dr. DeShazo's comments more directly, he stated, "I like the idea of asking perhaps some direct questions to try to get at their mental model for parental responsibility for the child. We thought we could get at that by just asking decision-making questions with respect to children's health and so on, but it's not enough. We can maybe get at it more directly."

Bryan Hubbell, (U.S. EPA)
Addressing his questions to Dr. Hammitt, he commented, "When we're dealing with the IQ evaluation, one of the things that struck me is when you asked the parents for their willingness to pay, and the reason it might be different than the cost of illness, is that you're essentially asking them to be able to project the relationship between IQ loss and future earnings. If they don't actually know that relationship, you're asking them to

somehow figure out what that six-point difference means. A question I have is: Could it instead be offered in showing them the information that's in the epidemiological literature relating the two?" He went on to phrase the question another way also: "If they're really not giving you their expectations of earnings loss, should this willingness to pay actually be *additive* to the cost of illness—so that there is some kind of estimate of utility loss beyond earnings?"

Still addressing Dr. Hammitt, Dr. Hubbell continued by saying he was also concerned about "the payment vehicle, in that you had it be a one-time payment in a particular year for what is essentially a lifetime impact." His question was: "If you would ask them instead what they would be willing to pay annually up through their child's eighteenth birthday in order to prevent this kind of exposure, would you be able to get a different value per IQ point? Again, this would reflect a lifetime impact rather than just a one-time payment, because you start getting into budget constraint issues and current trade-offs versus future earnings potential and future impacts."

Dr. Hubbell continued, "On your pesticide questionnaire one thing I'm really concerned about is the payment vehicle, again." He cited a study done by Kerry Smith and colleagues back in 1994 (he believes), in which they looked at the willingness to pay for avoiding risks from pesticides, focusing on grapefruit. Dr. Hubbell stated, "If you calculate a VSL based on their results, you get something like $80,000 or perhaps something even lower. Part of the reason for this is because it's tied to the specific product or to a particular sub-category of your budget. In those cases, in order to get a VSL that is more typical of what we get for environmental policy, you would have had to pay something like a hundred times the price of a grapefruit. Clearly, people are going to reject that. They're either going to hit the reservation price, or they're going to substitute, or something else." He closed by saying that his concern is that "you're going to run into the same problem here. While it still may be good to test the latency question, I wouldn't want to be able to use that VSL for anything—it's not really a VSL. The other related question is: While you say that you're not going to focus this on organics, people use organics as sort of a reference point. They know what organic foods cost and they've already made the decision one way or the other, so you can see that as a bounding on their willingness to pay extra for products. In fact, what they may do if you tell them a price that is higher than the organics is decide just to go to organics to get the health benefits plus the eco-benefits. Again, there's a bounding question there."

James Hammitt, (Harvard University)
Saying that those were "all good points," Dr. Hammitt first addressed the willingness to pay per IQ questions. He stated, "Clearly, I don't mean to suggest that EPA should use our value instead of the cost of illness. I think it's clear that people don't appreciate how much IQ apparently contributes to lifetime earnings. Whether some CV value should be added to the cost of illness value, I don't know—it might be that some part of the cost of illness is already in the CV. That's a good question."

Turning to the one-time payment issue, Dr. Hammitt clarified: "The willingness to pay amounts, the bids we offered people, were not extraordinarily high—they were a few hundred dollars. If you think of that as part of a tax payment, I don't think the income effect is going to be really important there. The one-time payment is consistent with the intervention as a one-time cleanup that will provide a long-stream term of benefits. So, asking about a one-time payment is not unreasonable on its face. These one-time cleanups of course could be financed by bonds, thereby spreading the cost to the taxpayers over many years, so one could do it many ways."

Reiterating that "everything matters," Dr. Hammitt continued, "I mentioned in the ERS study we asked about paying per meal or paying per month, where we had information on the frequency with which people consumed the various foods. So, we told them what the risk reduction would be on a per-month basis as well. I think our estimates of willingness to pay per meal are implausibly high—I think they're off by a couple of dollars per meal. That may be due to error in the sense that we tell them that the risk of getting sick from this one particular meal . . . –so there's a huge amount of salience there and maybe that's why they're paying a lot." He summarized that a $3 per meal increase over a month period really adds up to some money, but the gauge also involves "much bigger risks—these microbial illness risks are huge. So, as it turns out, our willingness to pay per unit of risk reduction is actually a little bit higher on the per-month basis than on the per-meal basis. But this is a general issue—how we allocate the timing of payments and what the benefits are, I think, is going to matter to our results."

———————

Susan Chilton, (University of Newcastle, United Kingdom)
Addressing her comment to Alan Krupnick and Sandra Hoffmann, Dr. Chilton said, "The issue about whether the mother's and the father's willingness to pay is the same—if it follows some empirical work that I've just completed—they won't be. In my study, they were asked separately and there were differences. Another interesting thing we found was that for an injury of low severity the mother's willingness to pay was higher than the father's in the same household. As the injury became more severe—this was in the context of child farm safety—the father's willingness to pay became higher than the mother's willingness to pay. It may be that the major decision maker in a household changes across the scope of an injury or illness, so that may be something to bear in mind."

———————

Mary Evans, (University of Tennessee)
Stating that she had "just a quick clarification question" for Drs. Krupnick and Hoffmann, Dr. Evans asked, "Can you talk a little bit about the level of information of respondents when they go into the initial interview? For example, are they aware of the fact that they will first be interviewed separately and then jointly—or are they expecting only to be interviewed by themselves?"

Alan Krupnick, (Resources for the Future)
Dr. Krupnick responded that the participants are aware of the format of the interview. He went on to clarify: "Actually, before the interview starts they are brought in together and asked to write down three recent decisions they've made regarding their children's health. The interviewers then get together and look at the responses and find one that's the same (or if not, they go back to the participants). Then, they use that common decision as the basis for the discussion of the decision making styles in the separate interviews. They know that they will then be coming back together to complete a second interview, so, yes, there is full information on that."

Dr. Krupnick added, "I'm not sure what your concern was—why don't you go a little further on that one?"

Evans
Dr. Evans clarified, "I guess I was just thinking about the broader implications of the question on who should we survey? Even in that context, if you find willingness to pay's to be equal, it still is not surprising that in a context where they're interviewed separately and those answers will never be rectified that we can see differences."

Lauraine Chestnut, (Stratus Consulting)
Addressing Dr. Hammitt, Ms. Chestnut said, "Maybe I need to see how you get from the question you asked about the IQ to the dollar per IQ to clarify this, but weren't you asking people about their willingness to pay for a cleanup program that's going to reduce risks to *somebody's* children but not necessarily their own? How many children were in the community? I guess I'm fuzzy about how we get from that to dollar per IQ—is that per one kid or per the community? Are we comparing apples and oranges?"

James Hammitt
Dr. Hammitt answered, "There is potentially a little ambiguity on that, but the idea is: What would you pay to reduce the risk that your child has this? So, it's one child—and then it's a reduction in the risk of suffering the six-point IQ deficit. So, it's willingness to pay divided by the change in probability divided by the six IQ points."

Chestnut
"So, it's: Suppose you had a child, and then . . ."

Hammitt
"Yes, right."

Sylvia Brandt, (University of Massachusetts)
Dr. Brandt asked this question of Drs. Krupnick and Hoffmann: "How are you going to connect your theoretical model to an empirical study? The reason I ask is because I have a concern. In building your theoretical model, you're working with a group of homogeneous, very traditional households. I understand why you wanted that group to be homogeneous. However, when I think about the population that we worry about when

we think about lead, I think about two things. One is housing structures of poor quality, typically in inner-city, lower-income neighborhoods. The second thing is poor nutrition, because the lower the iron level in your blood, the more likely it is that lead will bond to red blood cells. Both of these are more likely to occur in low-income, non-white populations. I know from personal experience in the Springfield, Massachusetts area, where we have a lead paint problem, eighty percent of our group were single-parent households. They were typically female, but they varied from being an aunt to a foster parent to a grandparent, so there was a lot of variation in the household structure. I wonder how you're going to make that leap from a model built on what I think of as a suburban setting to where the real problem is." Dr. Brandt went on with a second comment related to how participants were asked to rank health effects. She stated, "Again, building on my experience in Springfield and Oakland, when we ask households to rank health effects or health risks, they all might be ranked pretty low. For example, asthma morbidity, which in the suburb we may think is just outrageously out of control, may not be ranked as a high stress in inner-city households because they have competing stressors that are more basic than improved health—maybe it's making the rent payment or dealing with spousal abuse or kids' school issues, whatever. So, I would encourage you in asking about what are concerns to include, along with the health issues, also other things that may be important and that may completely dominate any health-related concerns in those settings where lead is a real problem."

Sandra Hoffmann
"In response to the first question, the focus of the study is really to try to get at the methodological question about whether we're taking the right approach in stated-preference surveys when we're trying to get at parental willingness to pay. The sample size that we can do, given the grant size, is fairly small, so it's always been conceived of as a pilot study that is focused on trying to examine this household modeling question. So, no, I don't think we're going to get really good measures of willingness to pay for reduction in neurotoxins that are representative of the entire population. That said, twenty –five percent of children in our country do live in homes that have lead paint as a potential hazard. I know in interviewing physicians in the Washington, DC area, they say that while one would expect that the risk is going to be highest in low-income households, they also see a lot of problems still in middle- and higher-income housing. So, what we're looking for are housing settings in which it *could* be a problem and family settings that raise a scenario in which we can test the alternative household hypotheses. Further work will have to be done to get more representativeness in income on neurotoxin hazards."

Alan Krupnick
Dr. Krupnick added, "Your second point is well taken, and we'll think about how to do that. On the first point I just wanted to add that we have no intent of generalizing these results beyond the group that we're targeting. We do find, however, that race has a significant effect on decision-making style—but, in our data it's correlated with income, so it's hard to know which is doing what."

END OF SESSION IV Q&A